(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
08 July 2021 (08.07.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/138247 A1

(51) International Patent Classification:
C12N 15/10 (2006.01) *C12N 15/63* (2006.01)
C12N 9/22 (2006.01)

(21) International Application Number:
PCT/US2020/067138

(22) International Filing Date:
28 December 2020 (28.12.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/955,014 30 December 2019 (30.12.2019) US
63/058,169 29 July 2020 (29.07.2020) US

(71) Applicant: **LIFEEDIT THERAPEUTICS, INC.** [US/US]; 104 T.W. Alexander Drive, Building 20, Research Triangle Park, North Carolina 27709 (US).

(72) Inventors: **CRAWLEY, Alexandra Briner**; 2216 Brisbane Woods Way, Cary, North Carolina 27518 (US). **BORDEN, Philip**; 1602 Milan Street, Durham, North Carolina 27704 (US). **BOWEN, Tyson D.**; 137 Waltons Creek Road, Morrisville, North Carolina 27560 (US). **COYLE, Michael**; 304 Copperline Drive, Apt. Q, Chapel Hill, North Carolina 27516 (US). **RACKEAR, Madison R.**; 85 Foster Street, Fl 1, New Haven, Connecticut 06511 (US). **ELICH, Tedd D.**; 5 Alumwood Place, Durham, North Carolina 27705 (US).

(74) Agent: **BUCK, B. Logan**; 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, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: RNA-GUIDED NUCLEASES AND ACTIVE FRAGMENTS AND VARIANTS THEREOF AND METHODS OF USE

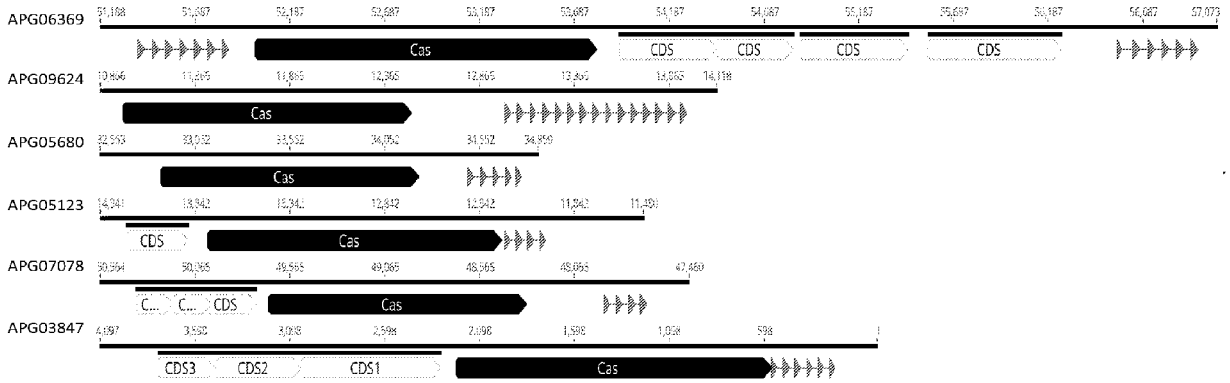


FIGURE 1

(57) Abstract: Compositions and methods for binding to a target sequence of interest are provided. The compositions find use in cleaving or modifying a target sequence of interest, detecting a target sequence of interest, and modifying the expression of a sequence of interest. Compositions comprise RNA-guided nuclease polypeptides, CRISPR RNAs, trans-activating CRISPR RNAs, guide RNAs, and nucleic acid molecules encoding the same. Vectors and host cells comprising the nucleic acid molecules are also provided. Further provided are CRISPR systems for binding a target sequence of interest, wherein the CRISPR system comprises an RNA-guided nuclease polypeptide and one or more guide RNAs.

WO 2021/138247 A1

[Continued on next page]

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

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))*

RNA-GUIDED NUCLEASES 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.62/955,014, filed December 30, 2019, and U.S. Provisional Application No. 63/058,169, filed July 29, 2020, each of which application 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_1180WO_(0077_8)_SL is 558,899 bytes in size, was created on December 17, 2020, 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, 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 break that is repaired via error-prone non-homologous end-joining (NHEJ) to introduce a mutation at a specific genomic location. Alternatively, heterologous DNA may be introduced into the genomic site via homology-directed repair. RNA-guided nucleases (RGNs) can also be used for base editing when fused with a deaminase or for detecting specific nucleotide sequences.

BRIEF SUMMARY OF THE INVENTION

Compositions and methods for binding a target sequence of interest are provided. The compositions find use in cleaving or modifying a target sequence of interest, detection of a target sequence of interest, and modifying the expression of a sequence of interest. Compositions comprise RNA-guided nuclease (RGN)

polypeptides, CRISPR RNAs (crRNAs), trans-activating CRISPR RNAs (tracrRNAs), guide RNAs (gRNAs), nucleic acid molecules encoding the same, vectors and host cells comprising the nucleic acid molecules, and kits comprising an RGN, gRNA, and detector single-stranded DNA. Also provided are CRISPR systems for binding a target sequence of interest, wherein the CRISPR system comprises an RNA-guided nuclease polypeptide and one or more guide RNAs. Thus, methods disclosed herein are drawn to binding a target sequence of interest, and in some embodiments, cleaving or modifying the target sequence of interest. The target sequence of interest can be modified, for example, as a result of non-homologous end joining, homology-directed repair with an introduced donor sequence, or base editing. Further provided are methods and kits for detecting a target DNA sequence of a DNA molecule using detector single-stranded DNA.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the bacterial genomic loci of representative RGNs of the invention.

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 and the associated drawings. 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 embodiments. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

I. Overview

RNA-guided nucleases (RGNs) allow for the targeted manipulation of specific site(s) 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 non-homologous end joining and homologous recombination, for example. The compositions and methods described herein are useful for creating single- or double-stranded breaks in polynucleotides, modifying polynucleotides, detecting a particular site within a polynucleotide, or modifying the expression of a particular gene.

The RNA-guided nucleases disclosed herein can alter gene expression by modifying a target sequence. In specific embodiments, the RNA-guided nucleases are 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. The 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. 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. Thus, provided herein are methods for using the RNA-guided nucleases to modify a target sequence in the DNA of host cells. For example, RNA-guided nucleases can be used to modify a target sequence at a genomic locus of eukaryotic cells or prokaryotic cells.

II. RNA-guided nucleases

Provided herein are RNA-guided nucleases. The term RNA-guided nuclease (RGN) refers to a polypeptide that binds to a particular target nucleotide sequence in a sequence-specific manner and is directed to the target nucleotide sequence by a guide RNA molecule that is complexed with the polypeptide and hybridizes with the target sequence. Although an RNA-guided nuclease can be capable of cleaving the target sequence upon binding, the term "RNA-guided nuclease" also encompasses nuclease-dead RNA-guided nucleases that are capable of binding to, but not cleaving, a target sequence. Cleavage of a target sequence by an RNA-guided nuclease can result in a single- or double-stranded break. RNA-guided nucleases only capable of cleaving a single strand of a double-stranded nucleic acid molecule are referred to herein as nickases.

The RGNs of the invention are members of the Class 2 CRISPR-Cas systems. More specifically, they are members of the Type V CRISPR-Cas systems. Type V CRISPR-Cas systems are broadly defined as systems that contain a single effector nuclease that is responsible for using the guide RNA to target dsDNA (double-stranded DNA); additionally, the single effector nuclease contains a split RuvC nuclease domain that is responsible for the catalytic activity (Jinek et al 2014, Science doi:10.1126/science.1247997; 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; each of which is incorporated herein by reference in its entirety). 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 Type V-A signature protein is Cas12a. It is 1,000-1,400 amino acids in length and has several domains in addition to the RuvC domain, including a wedge domain with recognition lobes (Yamano et al (2016) *Cell* 165:949-962). In contrast, the Type V-U systems are smaller in size (500-700 amino acids in length) compared to most other Type V systems. The V-U's also possess a split RuvC domain and a positively charged bridge helix (Shmakov et al 2017). The V-U proteins often do not have accessory Cas proteins encoded with the effector protein, while Cas12a co-localizes with cas1, cas2 and occasionally cas4 (Shmakov et al 2017). Based on these differences between the Type V-U systems and other Type V members, it was suggested by Shmakov et al (2017) that upon determination of functionality, the Type V-U systems should receive a new type/subtype designation.

For example, Cas14 enzymes are 400-700 amino acids in length (Harrington et al 2018). Upon first publication, these systems were touted as separate Cas enzymes from the canonical Cas12 effector protein for Type Vs. Later publications from Yan et al., have dubbed Cas14a, -b, and -c as subtype V-F within the Type V nomenclature. Cas14a and b are most closely related to c2c10, which is Type V-U3. Cas14c is

most closely related to c2c8 and c2c9, which are Types V-U2 and V-U4, respectively (Harrington et al 2018; Yan et al 2018). The genomic loci of the Cas14 RGNs are associated with accessory Cas proteins, and the tracrRNAs are encoded between the Cas14 and the repeat-spacer arrays. These systems cannot process their own guide RNAs, unlike Cas12a which is capable of processing individual guides from a single transcript containing multiple guide RNAs (Harrington et al 2018).

All RGNs of the invention contain a split RuvC domain, with the exception of APG06369. However, many of the RGNs of the invention have unique locus arrangements, which suggests that these RGNs are novel to the Class 2 CRISPR-Cas system of classification. None of the loci from which the RGNs of the invention are derived (see Table 1 in Example 1) contain Cas1 or Cas2.

As disclosed herein, APG07339, APG09624, APG03003, APG05405, APG09777, APG05680, APG02119, APG03285, APG04998, and APG07078 are standalone Cas effectors that are not encoded with accessory genes and may require a tracrRNA in addition to the crRNA. Based on the disclosures herein, these CRISPR-Cas systems need to receive a new classification. Additionally, phylogenetic analysis reveals that these RGNs can be grouped into three different subtypes. One subtype contains APG07078. The second subtype contains APG05680 and APG03285. The third sub-type contains APG07339, APG09624, APG03003, APG05405, APG09777, APG02119, and APG04998.

APG06369 is a unique effector nuclease that lacks a distinguishable RuvC domain and sits in a never-before seen CRISPR locus with non-canonical accessory genes. APG06369 has four accessory genes (the four accessory proteins are set forth as SEQ ID NOs: 178-181), none of which possess an annotated domain or function. APG06369 is a unique Cas protein.

Phylogenetically, APG03847, APG05625, APG03759, APG05123, and APG03524 form a clade of unique RuvC containing effector nucleases. These RGNs possess up to three accessory genes: one is an HNH endonuclease, one is an HTH transcriptional regulator, and the third has no known function or domains. The accessory proteins for APG03847 are set forth as SEQ ID NOs: 182, 183, and 184. The accessory proteins for APG05625 are set forth as SEQ ID NOs: 185, 186, and 187. The accessory proteins for APG03524 are set forth as SEQ ID NOs: 188, 189, and 190. The accessory proteins for APG03759 and APG05123 are set forth as SEQ ID NOs: 191 and 192, respectively. They have a unique CRISPR repeat arrangement at their loci, where the repeats associated with APG03847, APG05625, APG03759, APG05123, and APG03524 are flush with coding sequences for the numerous proteins. This is an extremely unusual feature for CRISPR-Cas systems, and suggests a form of CRISPR expression that does not require the leader sequence. Such a form of CRISPR expression is unlike any system known to date.

The RNA-guided nucleases disclosed herein include the RNA-guided nucleases shown in Table 1, the amino acid sequences of which are set forth as SEQ ID NOs: 1 to 109, and active fragments or variants thereof that retain the ability to bind to a target nucleotide sequence in an RNA-guided sequence-specific manner. In some embodiments, such an active fragment or variant of the RGN is capable of cleaving a single- or double-stranded target sequence. In some embodiments, an active variant of the RGN of the invention comprises an amino acid sequence having at least 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 any one of the amino acid sequences set forth as SEQ ID NOs: 1 to 109. In certain embodiments, an active fragment of the RGN of the invention comprises at least 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050 or more contiguous amino acid residues of any one of the amino acid sequences set forth as SEQ ID NOs: 1 to 109. RNA-guided nucleases provided herein can comprise at least one nuclease domain (*e.g.*, DNase, RNase domain) and at least one RNA recognition and/or RNA binding domain to interact with guide RNAs. Further domains that can be found in RNA-guided nucleases provided herein include, but are not limited to DNA binding domains, helicase domains, protein-protein interaction domains, and dimerization domains. In specific embodiments, the RNA-guided nucleases provided herein can comprise at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to one or more of a DNA binding domain, helicase domain, protein-protein interaction domain, and dimerization domain.

In various embodiments, a target nucleotide sequence is bound by an RNA-guided nuclease provided herein and hybridizes with the guide RNA associated with the RNA-guided nuclease. The target sequence can then be subsequently cleaved by the RNA-guided nuclease if the polypeptide possesses nuclease activity. The terms "cleave" or "cleavage" refer to the hydrolysis of at least one phosphodiester bond within the backbone of a target nucleotide sequence that can result in either single-stranded or double-stranded breaks within the target sequence. In various embodiments, the presently disclosed RGNs can cleave nucleotides within a polynucleotide, functioning as an endonuclease or can be an exonuclease, removing successive nucleotides from the end (the 5' and/or the 3' end) of a polynucleotide. In some embodiments, the disclosed RGNs can cleave nucleotides of a target sequence within any position of a polynucleotide and thus function as both an endonuclease and exonuclease. The cleavage of a target polynucleotide by the presently disclosed RGNs can result in staggered breaks or blunt ends.

In some embodiments, the RGN requires the expression or presence of at least one RGN accessory protein in order to bind to and/or cleave a target polynucleotide. In some of these embodiments, the RGN requires at least one RGN accessory protein set forth as SEQ ID NOs: 178-192 or an active variant or fragment thereof. In particular embodiments wherein the RGN is APG06369 (SEQ ID NO: 11) or a variant or fragment thereof, at least one RGN accessory protein set forth as SEQ ID NOs: 178-181 or an active variant or fragment thereof is required for activity. In some of those embodiments wherein the RGN is APG03847 (SEQ ID NO: 12) or a variant or fragment thereof, at least one RGN accessory protein set forth as SEQ ID NOs: 182-184 or an active variant or fragment thereof is required for activity. In certain embodiments wherein the RGN is APG05625 (SEQ ID NO: 13) or a variant or fragment thereof, at least one RGN accessory protein set forth as SEQ ID NOs: 185-187 or an active variant or fragment thereof is required for activity. In some embodiments wherein the RGN is APG03524 (SEQ ID NO: 16) or a variant or fragment thereof, at least one RGN accessory protein set forth as SEQ ID NOs: 188-190 or an active variant or fragment thereof is required for activity. In particular embodiments wherein the RGN is APG03759 (SEQ ID NO: 14) or a variant or fragment thereof, the RGN accessory protein set forth as SEQ

ID NO: 191 or an active variant or fragment thereof is required for activity. In certain embodiments wherein the RGN is APG05123 (SEQ ID NO: 15) or a variant or fragment thereof, the RGN accessory protein set forth as SEQ ID NO: 192 or an active variant or fragment thereof is required for activity.

In some embodiments, the presently disclosed RNA-guided nucleases can be wild-type sequences derived from bacterial or archaeal species. In some embodiments, the RNA-guided nucleases can be variants or fragments of wild-type polypeptides. The wild-type RGN can be modified to alter nuclease activity or alter PAM specificity, for example. In some embodiments, the RNA-guided nuclease is not naturally-occurring.

In certain embodiments, the RNA-guided nuclease functions as a nickase, only cleaving a single strand of the target nucleotide sequence. Such RNA-guided nucleases have a single functioning nuclease domain. In particular embodiments, the nickase is capable of cleaving the positive strand or negative strand. In some of these embodiments, additional nuclease domains have been mutated such that the nuclease activity is reduced or eliminated.

In some embodiments, the RNA-guided nuclease lacks nuclease activity altogether and is referred to herein as nuclease-dead or nuclease inactive. 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 herein by reference in its entirety.

RNA-guided nucleases that lack nuclease activity can be used to deliver a fused polypeptide, polynucleotide, or small molecule payload to a particular genomic location. In some of these embodiments, the RGN polypeptide or guide RNA can be fused to a detectable label to allow for detection of a particular sequence. As a non-limiting example, a nuclease-dead RGN can be fused to a detectable label (*e.g.*, fluorescent protein) and targeted to a particular sequence associated with a disease to allow for detection of the disease-associated sequence.

In some embodiments, nuclease-dead RGNs can be targeted to particular genomic locations to alter the expression of a desired sequence. In some embodiments, the binding of a nuclease-dead RNA-guided nuclease to a target sequence results in the reduction in expression of the target sequence or a gene under transcriptional control by the target sequence by interfering with the binding of RNA polymerase or transcription factors within the targeted genomic region. In other embodiments, the RGN (*e.g.*, a nuclease-dead RGN) or its complexed guide RNA further comprises an expression modulator that, upon binding to a target sequence, serves to either repress or activate the expression of the target sequence or a gene under transcriptional control by the target sequence. In some of these embodiments, the expression modulator modulates the expression of the target sequence or regulated gene through epigenetic mechanisms.

In some embodiments, the nuclease-dead RGNs or an RGN with only nickase activity can be targeted to particular genomic locations to modify the sequence of a target polynucleotide through fusion to a base-editing polypeptide, for example a deaminase polypeptide or active variant or fragment thereof, that directly chemically modifies (e.g., deaminates) a nucleobase, resulting in conversion from one nucleobase to

another. The base-editing polypeptide can be fused to the RGN at its N-terminal or C-terminal end. Additionally, the base-editing polypeptide may be fused to the RGN via a peptide linker. A non-limiting example of a deaminase polypeptide that is useful for such compositions and methods include a cytidine deaminase or an adenine deaminase (such as the adenosine base editor described in Gaudelli *et al.* (2017) *Nature* 551:464-471, U.S. Publ. Nos. 2017/0121693 and 2018/0073012, International Publ. No. WO 2018/027078, or any of the deaminases disclosed in International Publ. No. WO 2020/139873, and U.S. Provisional Appl. Nos. 62/785,391 (filed December 27, 2018), 62/932,169 (filed November 7, 2019), and 63/077,089 (filed September 11, 2020), each of which is herein incorporated by reference in its entirety. Further, it is known in the art that certain fusion proteins between an RGN and a base-editing enzyme may also comprise at least one uracil stabilizing polypeptide that increases the mutation rate of a cytidine, deoxycytidine, or cytosine to a thymidine, deoxythymidine, or thymine in a nucleic acid molecule by a deaminase. Non-limiting examples of uracil stabilizing polypeptides include those disclosed in U.S. Provisional Appl. No. 63/052,175, filed July 15, 2020, and a uracil glycosylase inhibitor (UGI) domain (SEQ ID NO: 137), which may increase base editing efficiency. In particular embodiments, the present disclosure provides a fusion protein comprising an RGN described herein or a variant thereof, a deaminase, and optionally at least one uracil stabilizing polypeptides, such as UGI. In certain embodiments, the RGN that is fused to the base-editing polypeptide is a nickase that cleaves the DNA strand that is not acted upon by the base-editing polypeptide (e.g., deaminase). RNA-guided nucleases that are fused to a polypeptide or domain can be separated or joined by a linker. 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 a gRNA binding domain of an RNA guided nuclease and a base-editing polypeptide, such as a deaminase. In some embodiments, a linker joins a nuclease-dead RGN 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 various embodiments, the present disclosure provides the presently disclosed RNA-guided nucleases comprising at least one nuclear localization signal (NLS) to enhance transport of the RGN to the nucleus of a cell. Nuclear localization signals are known in the art and generally comprise a stretch of basic amino acids (see, *e.g.*, Lange *et al.*, *J. Biol. Chem.* (2007) 282:5101-5105). In particular embodiments, the RGN comprises 2, 3, 4, 5, 6 or more nuclear localization signals. The nuclear localization signal(s) can be a heterologous NLS. Non-limiting examples of nuclear localization signals useful for the presently disclosed RGNs are the nuclear localization signals of SV40 Large T-antigen, nucleopasmin, and c-Myc (*see, e.g.*, Ray et al. (2015) *Bioconjug Chem* 26(6):1004-7). In particular embodiments, the RGN comprises the NLS

sequence set forth as SEQ ID NO: 149 or 150. The RGN can comprise one or more NLS sequences at its N-terminus, C- terminus, or both the N-terminus and C-terminus. For example, the RGN can comprise two NLS sequences at the N-terminal region and four NLS sequences at the C-terminal region.

Other localization signal sequences known in the art that localize polypeptides to particular subcellular location(s) can also be used to target the RGNs, including, but not limited to, plastid localization sequences, mitochondrial localization sequences, and dual-targeting signal sequences that target to both the plastid and mitochondria (see, *e.g.*, Nassoury and Morse (2005) *Biochim Biophys Acta* 1743:5-19; Kunze and Berger (2015) *Front Physiol* dx.doi.org/10.3389/fphys.2015.00259; Herrmann and Neupert (2003) *IUBMB Life* 55:219-225; Soll (2002) *Curr Opin Plant Biol* 5:529-535; Carrie and Small (2013) *Biochim Biophys Acta* 1833:253-259; Carrie *et al.* (2009) *FEBS J* 276:1187-1195; Silva-Filho (2003) *Curr Opin Plant Biol* 6:589-595; Peeters and Small (2001) *Biochim Biophys Acta* 1541:54-63; Murcha *et al.* (2014) *J Exp Bot* 65:6301-6335; Mackenzie (2005) *Trends Cell Biol* 15:548-554; Glaser *et al.* (1998) *Plant Mol Biol* 38:311-338).

In certain embodiments, the presently disclosed RNA-guided nucleases comprise at least one cell-penetrating domain that facilitates cellular uptake of the RGN. 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.

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 RNA-guided nuclease.

In certain embodiments, the presently disclosed RGNs are fused to an effector domain, such as a cleavage domain, a deaminase domain, or an expression modulator domain, either directly or indirectly via a linker peptide. Such a domain can be located at the N-terminus, the C-terminus, or an internal location of the RNA-guided nuclease. In some of these embodiments, the RGN component of the fusion protein is a nuclease-dead RGN.

In some embodiments, the RGN fusion protein comprises a cleavage domain, which is any domain that is capable of cleaving a polynucleotide (*i.e.*, RNA, DNA, or RNA/DNA hybrid) and includes, but is not limited to, restriction endonucleases and homing endonucleases, such as Type IIS endonucleases (*e.g.*, *Fok*I) (*see, e.g.*, Belfort *et al.* (1997) *Nucleic Acids Res.* 25:3379-3388; Linn *et al.* (eds.) Nucleases, Cold Spring Harbor Laboratory Press, 1993).

In some embodiments, the RGN fusion protein comprises a deaminase domain that deaminates a nucleobase, resulting in conversion from one nucleobase to another, and includes, but is not limited to, a cytidine deaminase or an adenine deaminase base editor (see, *e.g.*, Gaudelli *et al.* (2017) *Nature* 551:464-

471, U.S. Publ. Nos. 2017/0121693 and 2018/0073012, U.S. Patent No. 9,840,699, International Publ. No. WO/2018/027078, International Appl. No. PCT/US2019/068079, and U.S. Provisional Appl. Nos. 62/785,391 (filed December 27, 2018) and 62/932,169 (filed November 7, 2019)).

In some embodiments, the effector domain of the RGN fusion protein can be an expression modulator domain, which is a domain that either serves to upregulate or downregulate transcription. The expression modulator domain can be an epigenetic modification domain, a transcriptional repressor domain or a transcriptional activation domain.

In some of these embodiments, the expression modulator of the RGN fusion protein comprises an epigenetic modification domain that covalently modifies DNA or histone proteins to alter histone structure and/or chromosomal structure without altering the DNA sequence, leading to changes in gene expression (*i.e.*, upregulation or downregulation). Non-limiting examples of epigenetic modifications include acetylation or methylation of lysine residues, arginine methylation, serine and threonine phosphorylation, and lysine ubiquitination and sumoylation of histone proteins, and methylation and hydroxymethylation of cytosine residues in DNA. Non-limiting examples of epigenetic modification domains include histone acetyltransferase domains, histone deacetylase domains, histone methyltransferase domains, histone demethylase domains, DNA methyltransferase domains, and DNA demethylase domains.

In some embodiments, the expression modulator of the fusion protein comprises a transcriptional repressor domain, which interacts with transcriptional control elements and/or transcriptional regulatory proteins, such as RNA polymerases and transcription factors, to reduce or terminate transcription of at least one gene. Transcriptional repressor domains are known in the art and include, but are not limited to, Sp1-like repressors, IκB, and Krüppel associated box (KRAB) domains.

In some embodiments, the expression modulator of the fusion protein comprises a transcriptional activation domain, which interacts with transcriptional control elements and/or transcriptional regulatory proteins, such as RNA polymerases and transcription factors, to increase or activate transcription of at least one gene. Transcriptional activation domains are known in the art and include, but are not limited to, a herpes simplex virus VP16 activation domain and an NFAT activation domain.

In some embodiments, the presently disclosed RGN polypeptides comprise a detectable label or a purification tag. The detectable label or purification tag can be located at the N-terminus, the C-terminus, or an internal location of the RNA-guided nuclease, either directly or indirectly via a linker peptide. In some of these embodiments, the RGN component of the fusion protein is a nuclease-dead RGN. In other embodiments, the RGN component of the fusion protein is an RGN with nickase activity.

A detectable label is a molecule that can be visualized or otherwise observed. The detectable label may be fused to the RGN as a fusion protein (*e.g.*, fluorescent protein) or may be a small molecule conjugated to the RGN polypeptide that can be detected visually or by other means. Detectable labels that can be fused to the presently disclosed RGNs as a fusion protein include any detectable protein domain, including but not limited to, a fluorescent protein or a protein domain that can be detected with a specific antibody. Non-limiting examples of fluorescent proteins include green fluorescent proteins (*e.g.*, GFP,

EGFP, ZsGreen1) and yellow fluorescent proteins (*e.g.*, YFP, EYFP, ZsYellow1). Non-limiting examples of small molecule detectable labels include radioactive labels, such as ^{3}H and ^{35}S.

In some embodiments the presently disclosed RGN polypeptides comprise a purification tag, which is any molecule that can be utilized to isolate a protein or fused protein from a mixture (e.g., biological sample, culture medium). Non-limiting examples of purification tags include biotin, myc, maltose binding protein (MBP), and glutathione-S-transferase (GST).

III. Guide RNA

The present disclosure provides guide RNAs and polynucleotides encoding the same. 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 RNA-guided nuclease to the target nucleotide sequence. Thus, an RGN's 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 embodiments wherein the RGN has nickase or nuclease activity, also cleave the target nucleotide sequence. In some embodiments, a guide RNA comprises a CRISPR RNA (crRNA) and in some embodiments, a trans-activating CRISPR RNA (tracrRNA). Native guide RNAs that comprise both a crRNA and a tracrRNA generally comprise two separate RNA molecules that hybridize to each other through the repeat sequence of the crRNA and the anti-repeat sequence of the tracrRNA.

In some embodiments, native direct repeat sequences within a CRISPR array range in length from 28 to 37 base pairs. In some embodiments, native direct repeat sequences within a CRISPR array range in length from about 23 bp to about 55 bp (e.g., from 23 bp to 55bp). In some embodiments, spacer sequences within a CRISPR array range from about 32 to about 38 bp in length. In some embodiments, spacer sequences within a CRISPR array range from about 21 bp to about 72 bp (e.g., from 21 bp to 72 bp). In some embodiments a CRISPR array disclosed herein comprises less than 50 units of the CRISPR repeat-spacer sequence. The CRISPRs are transcribed as part of a long transcript termed the primary CRISPR transcript, which comprises much of the CRISPR array. The primary CRISPR transcript is cleaved by Cas proteins to produce crRNAs or in some cases, to produce pre-crRNAs that are further processed by additional Cas proteins into mature crRNAs. Mature crRNAs comprise a spacer sequence and a CRISPR repeat sequence. In some embodiments in which pre-crRNAs are processed into mature (or processed) crRNAs, maturation involves the removal of about one to about six or more 5', 3', or 5' and 3' nucleotides. For the purposes of genome editing or targeting a particular target nucleotide sequence of interest, these nucleotides that are removed during maturation of the pre-crRNA molecule are not necessary for generating or designing a guide RNA. The consensus repeat sequence for each of the presently disclosed RGN proteins (SEQ ID NOs: 1-109) is disclosed in SEQ ID NOs: 201-309, respectively. The processed crRNA repeat sequence for each of APG07339 (SEQ ID NO: 1), APG09624 (SEQ ID NO: 2), APG03003 (SEQ ID NO: 3), APG05405 (SEQ ID NO: 4), APG09777 (SEQ ID NO: 5), APG05680 (SEQ ID NO: 6), APG06369

(SEQ ID NO: 11), APG03847 (SEQ ID NO: 12), APG05625 (SEQ ID NO: 13), and APG03524 (SEQ ID NO: 16) is disclosed in SEQ ID NOs: 110-119, respectively.

A CRISPR RNA (crRNA) 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. 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 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 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 CRISPR repeat sequence is about 21 nucleotides in length. In some embodiments, the degree of complementarity between a CRISPR repeat sequence and its 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 particular embodiments, the CRISPR repeat sequence comprises any one of the nucleotide sequences of SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309, or an active variant or fragment thereof that when comprised within a guide RNA, is capable of directing the sequence-specific binding of an associated RNA-guided nuclease provided herein to a target sequence of interest. In certain embodiments, an active CRISPR repeat sequence variant of a wild-type sequence comprises a nucleotide sequence having at least 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 any one of the nucleotide sequences set forth as SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309. In certain embodiments, an active CRISPR repeat sequence fragment of a wild-type sequence comprises at least 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 contiguous nucleotides of any one of the nucleotide sequences set forth as SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309.

In certain embodiments, the crRNA is not naturally-occurring. In some of these embodiments, the specific CRISPR repeat sequence is not linked to the engineered spacer sequence in nature and the CRISPR repeat sequence is considered heterologous to the spacer sequence. In certain embodiments, the spacer sequence is an engineered sequence that is not naturally occurring.

In some embodiments, the guideRNA further comprises a tracrRNA molecule. A trans-activating 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 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, 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 100 nucleotides in length. In certain embodiments, the tracrRNA is about 96 nucleotides in length.

In particular embodiments, the tracrRNA comprises any one of the nucleotide sequences of SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148, or an active variant or fragment thereof that when comprised within a guide RNA is capable of directing the sequence-specific binding of an associated RNA-guided nuclease provided herein to a target sequence of interest. In certain embodiments, an active tracrRNA sequence variant of a wild-type sequence comprises a nucleotide sequence having at least 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 any one of the nucleotide sequences set forth as SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148. In certain embodiments, an active tracrRNA sequence fragment of a wild-type sequence comprises at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more contiguous nucleotides of any one of the nucleotide sequences set forth as SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148.

Two polynucleotide sequences can be considered to be substantially complementary when the two sequences hybridize to each other under stringent conditions. Likewise, an RGN is considered to bind to a particular target sequence within a sequence-specific manner if the guide RNA bound to the RGN binds to the target sequence under stringent conditions. By "stringent conditions" or "stringent hybridization conditions" is intended conditions under which the two polynucleotide sequences will hybridize 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 Tm 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 Tm can be approximated from the equation of Meinkoth and Wahl (1984) Anal. Biochem. 138:267-284: Tm = 81.5°C + 16.6 (log M) + 0.41 (%GC) - 0.61 (% 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 (Tm) 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 (Tm); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point (Tm); 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 (Tm). Using the equation, hybridization and wash compositions, and desired Tm, 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).

The term "sequence specific" can also refer to the binding of a target sequence at a greater frequency than binding to a randomized background sequence.

In some embodiments, e.g., wherein the guide RNA comprises both a crRNA and a tracrRNA, the guide RNA can be a single guide RNA or a dual-guide RNA system. A single guide RNA comprises the crRNA and 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 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 linker nucleotide sequence is the nucleotide sequence set forth as SEQ ID NO: 136. In other embodiments, the linker nucleotide sequence is at least 6 nucleotides in length.

The single guide RNA or dual-guide RNA can be synthesized chemically or via *in vitro* transcription. Assays for determining sequence-specific binding between an RGN and a guide RNA are known in the art and include, but are not limited to, *in vitro* binding assays between an expressed RGN and the guide RNA, which can be tagged with a detectable label (e.g., biotin) and used in a pull-down detection assay in which the guide RNA:RGN complex is captured via the detectable label (e.g., with streptavidin beads). A control guide RNA with an unrelated sequence or structure to the guide RNA can be used as a negative control for non-specific binding of the RGN to RNA. In certain embodiments, the guide RNA is any one of SEQ ID NOs: 129 to 135 and 310, wherein the spacer sequence can be any sequence and is indicated as a poly-N sequence.

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). In certain embodiments, cleavage of a double-stranded target sequence is dependent upon the presence of a PAM, whereas cleavage of a single-stranded target sequence is PAM-independent. A protospacer adjacent motif 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 5' of the target sequence for the presently disclosed RGNs. Generally, the PAM is a consensus sequence of about 3-4 nucleotides, but in particular embodiments, can be 1, 2, 3, 4, 5, 6, 7, 8, 9, or more nucleotides in length. In some embodiments, the PAM is 5' of the target sequence and is T-rich.

In some embodiments, the RGN binds to a guide sequence comprising a CRISPR repeat sequence set forth in any one of SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309, or an active variant or fragment thereof, and a tracrRNA sequence set forth in any one of SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148, respectively, or an active variant or fragment thereof. The RGN systems are described further in Example 1 and Table 1 of the present specification.

It is well-known in the art that PAM sequence specificity for a given nuclease enzyme is affected by enzyme concentration (*see, e.g.*, Karvelis *et al.* (2015) *Genome Biol* 16:253), which may be modified by altering the promoter used to express the RGN, or the amount of ribonucleoprotein complex delivered to the cell, organelle, or embryo.

In those embodiments wherein binding and cleavage by the RGN is dependent upon a PAM sequence, 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. In some embodiments, the cleavage site may be over 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides from the PAM in either the 5' or 3' direction. In some embodiments, the cleavage site is 4 nucleotides away from the PAM. In other embodiments, the cleavage site is at least 15 nucleotides away from the PAM. 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.

IV. Nucleotides Encoding RNA-guided nucleases, CRISPR RNA, and/or tracrRNA

The present disclosure provides polynucleotides comprising the presently disclosed CRISPR RNAs, tracrRNAs, and/or sgRNAs and polynucleotides comprising a nucleotide sequence encoding the presently disclosed RNA-guided nucleases, CRISPR RNAs, tracrRNAs, and/or sgRNAs. Presently disclosed polynucleotides include those comprising or encoding a CRISPR repeat sequence comprising any one of the nucleotide sequences of SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309, or an active variant or fragment thereof that when comprised within a guide RNA is capable of directing the sequence-specific binding of an associated RNA-guided nuclease to a target sequence of interest. Also disclosed are polynucleotides comprising or encoding a tracrRNA comprising any one of the nucleotide sequences of SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148, or an active variant or fragment thereof that when comprised within a guide RNA is capable of directing the sequence-specific binding of an associated RNA-guided nuclease to a target sequence of interest. Polynucleotides are also provided that encode an RNA-guided nuclease comprising any one of the amino acid sequences set forth as SEQ ID NOs: 1 to 109, and active fragments or variants thereof that retain the ability to bind to a target nucleotide sequence in an RNA-guided sequence-specific manner.

The use of the term "polynucleotide" is not intended to limit the present disclosure to polynucleotides comprising DNA, though such DNA polynucleotides are contemplated. 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. These include, e.g., peptide nucleic acids (PNAs), PNA-DNA chimers, locked nucleic acids (LNAs), and phosphothiorate linked sequences. The polynucleotides disclosed herein also encompass all forms of sequences including, but not limited to, single-stranded forms, double-stranded forms, DNA-RNA hybrids, triplex structures, stem-and-loop structures, circular forms (e.g., including circular RNA), and the like.

In some embodiments, the nucleic acid molecules encoding RGNs are 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 RGNs, crRNAs, tracrRNAs, and/or sgRNAs provided herein are in some embodiments provided in expression cassettes for *in vitro* expression or expression in a cell, organelle, embryo, or organism of interest. The cassette may include 5' and 3' regulatory sequences operably linked to a polynucleotide encoding an RGN, crRNA, tracrRNAs, and/or sgRNAs 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 an RGN, crRNA, tracrRNAs, and/or sgRNAs) 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 RGN can be present on one expression cassette, whereas the nucleotide sequence encoding a crRNA, tracrRNA, or complete guide RNA can be on a separate 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. The expression cassette may additionally contain a selectable marker gene.

The expression cassette may include in the 5'-3' direction of transcription, a transcriptional (and, in some embodiments, translational) initiation region (*i.e.*, a promoter), an RGN-, crRNA-, tracrRNA-and/or sgRNA- 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; 8,586832; 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 an RGN, crRNA, and/or tracrRNA 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 RGNs, crRNAs, tracrRNAs, and/or sgRNAs 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 RGN, crRNA, tracrRNA, and/or sgRNA 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 RGN 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 RGN, crRNA, tracrRNA, and/or sgRNA 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 the RGN polypeptide can further comprise a sequence encoding a crRNA and/or a tracrRNA, or the crRNA and tracrRNA combined to create a guide RNA. The sequence(s) encoding the crRNA and/or tracrRNA can be operably linked to at least one transcriptional control sequence for expression of the crRNA and/or tracrRNA in the organism or host cell of interest. For example, the polynucleotide encoding the crRNA and/or tracrRNA 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 RGNs, crRNA, tracrRNA, and/or sgRNA 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 an RGN, crRNA, and/or tracrRNA 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 an RGN, crRNA, and/or tracrRNA includes retroviral, lentiviral, adenoviral, and adeno-associated viral mediated introduction and expression, as well as the use of Caulimoviruses, Geminiviruses, and RNA plant viruses.

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. Nail. 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 RGNs, crRNAs, and/or tracrRNAs can also be used to transform any prokaryotic species, including but not limited to, archaea and bacteria (*e.g.*, *Bacillus* sp., *Klebsiella* sp. *Streptomyces* sp., *Rhizobium* sp., *Escherichia* sp., *Pseudomonas* sp., *Salmonella* sp., *Shigella* sp., *Vibrio* sp., *Yersinia* sp., *Mycoplasma* sp., Agrobacterium, *Lactobacillus* sp.).

The polynucleotides encoding the RGNs, crRNAs, and/or tracrRNAs 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.

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 components of a CRISPR system 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. 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, 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. Viral. 66:2731-2739 (1992); Johann et al., J. Viral. 66:1635-1640 (1992); Sommnerfelt et al., Viral. 176:58-59 (1990); Wilson et al., J. Viral. 63:2374-2378 (1989); Miller et al., *1.* Viral. 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, *1.* 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., *1.* Viral. 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 ψJ2 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. 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, Huhl, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panel, PC-3, TFl, CTLL-2, CIR, Rat6, CVI, RPTE, AlO, T24, 182, A375, ARH-77, Calul, SW480, SW620, SKOV3, SK-UT, CaCo2, P388Dl, SEM-K2, WEHI- 231, HB56, TIB55, lurkat, *145.01,* LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep 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-Ll, 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-I cells, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10Tl/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-Kl, CHO-K2, CHO-T, CHO Dhfr-/-, COR-L23, COR-L23/CPR, COR-L235010, CORL23/ R23, COS-7, COV-434, CML Tl, 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, lurkat, *lY* cells, K562 cells, Ku812, KCL22, KGl, KYOl, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-l0A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCKII, MDCKII, MOR/ 0.2R, MONO-MAC 6, MTD-lA, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, OPCN/OPCT cell lines, Peer, PNT-lA/ PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 cells, Sf-9, SkBr3, T2, T-47D, T84, THPl 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 establish a new cell line comprising one or more vector-derived sequences. In some embodiments, a cell transiently transfected with the components of a CRISPR system as described herein (such as by transient

transfection of one or more vectors, or transfection with RNA), and modified through the activity of a CRISPR 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 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 a mammal, such as a mouse, rat, or rabbit.

V. Variants and Fragments of Polypeptides and Polynucleotides

The present disclosure provides active variants and fragments of a naturally-occurring (*i.e.*, wild-type) RNA-guided nuclease, the amino acid sequences of which are set forth as SEQ ID NOs: 1 to 109, as well as active variants and fragments of naturally-occurring CRISPR repeats, such as the sequences set forth as SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309, and active variant and fragments of naturally-occurring tracrRNAs, such as the sequences set forth as SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148, and polynucleotides encoding the same. Also provided are active variants and fragments of naturally-occurring RGN accessory proteins, such as the sequences set forth as SEQ ID NOs: 178-192.

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 RGN polypeptides, such as those disclosed herein, will retain sequence-specific, RNA-guided DNA-binding activity. In particular embodiments, fragments and variants of naturally-occurring RGN polypeptides, such as those disclosed herein, will retain nuclease activity (single-stranded or double-stranded).

Fragments and variants of naturally-occurring CRISPR repeats, such as those disclosed herein, will retain the ability, when part of a guide RNA (comprising a tracrRNA), to bind to and guide an RNA-guided nuclease (complexed with the guide RNA) to a target nucleotide sequence in a sequence-specific manner.

Fragments and variants of naturally-occurring tracrRNAs, such as those disclosed herein, will retain the ability, when part of a guide RNA (comprising a CRISPR RNA), to guide an RNA-guided nuclease (complexed with the guide RNA) to a target nucleotide sequence in a sequence-specific manner.

Fragments and variants of naturally-occurring RGN accessory proteins, such as those disclosed herein, will retain the ability, when part of an RGN system (i.e., RGN protein and guide RNA), to allow for the RGN system to bind to a target nucleotide sequence in a sequence-specific manner.

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.*, binding to and directing an RGN in a sequence-specific manner to a target nucleotide sequence when comprised within a guideRNA). "Fragments" or "biologically active portions" include polypeptides comprising a sufficient number of contiguous amino acid residues to retain the biological activity (*i.e.*, binding to a target nucleotide sequence in a sequence-specific manner when complexed with a guide RNA). Fragments of the RGN proteins 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 an RGN protein can be a polypeptide that comprises, for example, 10, 25, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600 or more contiguous amino acid residues of SEQ ID NOs: 1 to 109. Such biologically active portions can be prepared by recombinant techniques and evaluated for sequence-specific, RNA-guided DNA-binding activity. A biologically active fragment of a CRISPR repeat sequence can comprise at least 8 contiguous amino acids of any one of SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309. A biologically active portion of a CRISPR repeat sequence can be a polynucleotide that comprises, for example, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 contiguous nucleotides of any one of SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309. A biologically active portion of a tracrRNA can be a polynucleotide that comprises, for example, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80 or more contiguous nucleotides of any one of SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148. Fragments of the RGN accessory proteins 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 an RGN accessory protein can be a polypeptide that comprises, for example, 10, 25, 50, 100, 150, 200, or more contiguous amino acid residues of SEQ ID NOs: 178 to 192. Such biologically active portions can be prepared by recombinant techniques and evaluated for biological 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 an RNA-guided nuclease polypeptide 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 one of SEQ ID NOs: 1 to 109.

A biologically active variant of an RGN 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 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600 amino acids or more from either the N or C terminus of the polypeptide.

In certain embodiments, the presently disclosed polynucleotides encode an RNA-guided nuclease accessory polypeptide 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 one of SEQ ID NOs: 178 to 192.

A biologically active variant of an RGN accessory 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 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200 amino acids or more from either the N or C terminus of the polypeptide.

In certain embodiments, the presently disclosed polynucleotides comprise or encode a CRISPR repeat comprising a nucleotide 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 any one of the nucleotide sequences set forth as SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309.

The presently disclosed polynucleotides can comprise or encode a tracrRNA comprising a nucleotide 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 any one of the nucleotide sequences set forth as SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148.

Biologically active variants of a CRISPR repeat or tracrRNA of the invention may differ by as few as about 1-15 nucleotides, 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 nucleotide. In specific embodiments, the polynucleotides can comprise a 5' or 3' truncation, which can comprise at least a deletion of 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80 nucleotides or more from either the 5' or 3' end of the polynucleotide.

It is recognized that modifications may be made to the RGN polypeptides, CRISPR repeats, and tracrRNAs 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 RGN proteins. Alternatively, modifications may be made that improve the activity of the RGN.

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 RGN proteins disclosed herein (e.g., SEQ ID NOs: 1 to 109) is manipulated to create a new RGN protein 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 RGN sequences provided herein and other known RGN 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; Crameri *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; Crameri *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 acid 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 (on the world wide web at 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 RGN polypeptides or ribonucleoproteins comprising the RGN polypeptides of the present invention, including those having any one of the amino acid sequences set forth as SEQ ID NOs: 1 to 109 or active variants or fragments thereof, or the RGN accessory proteins of the present invention, including those having any one of the amino acid sequences set forth as SEQ ID NOs: 178 to 192 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 RGN polypeptides or ribonucleoproteins. Thus, this disclosure provides kits comprising antibodies that specifically bind to the polypeptides or ribonucleoproteins described herein, including, for example, polypeptides having the sequence of any one of SEQ ID NOs: 1 to 109 or 178 to 192.

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

The present disclosure provides a system for binding a target sequence of interest, wherein the system comprises at least one guide RNA or a nucleotide sequence encoding the same, and at least one RNA-guided nuclease or a nucleotide sequence encoding the same. The guide RNA hybridizes to the target sequence of interest and also forms a complex with the RGN polypeptide, thereby directing the RGN polypeptide to bind to the target sequence. In some of these embodiments, the RGN comprises an amino acid sequence of any one of SEQ ID NOs: 1 to 109, or an active variant or fragment thereof. In various embodiments, the guide RNA comprises a CRISPR repeat sequence comprising the nucleotide sequence of any one of SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309, or an active variant or fragment thereof. In particular embodiments, the guide RNA comprises a tracrRNA comprising a nucleotide sequence of any one of SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148, or an active variant or fragment thereof. The guide RNA of the system can be a single guide RNA or a dual-guide RNA. In particular embodiments, the system comprises an RNA-guided nuclease that is heterologous to the guideRNA, wherein the RGN and guideRNA are not found complexed to one another (i.e., bound to one another) in nature.

In some embodiments, the system further comprises at least one RGN accessory protein in order to bind to and/or cleave a target polynucleotide. In some of these embodiments, the system further comprises at least one RGN accessory protein set forth as SEQ ID NOs: 178-192 or an active variant or fragment thereof. In particular embodiments wherein the RGN is APG06369 (SEQ ID NO: 11) or a variant or fragment thereof, the system can further comprise at least one RGN accessory protein set forth as SEQ ID NOs: 178-181 or an active variant or fragment thereof. In some of those embodiments wherein the RGN is APG03847 (SEQ ID NO: 12) or a variant or fragment thereof, the system can further comprise at least one RGN accessory protein set forth as SEQ ID NOs: 182-184 or an active variant or fragment thereof. In certain embodiments wherein the RGN is APG05625 (SEQ ID NO: 13) or a variant or fragment thereof, the system can further comprise at least one RGN accessory protein set forth as SEQ ID NOs: 185-187 or an active variant or fragment thereof. In some embodiments wherein the RGN is APG03524 (SEQ ID NO: 16) or a variant or fragment thereof, the system can further comprise at least one RGN accessory protein set forth as SEQ ID NOs: 188-190 or an active variant or fragment thereof. In particular embodiments wherein the RGN is APG03759 (SEQ ID NO: 14) or a variant or fragment thereof, the system can further comprise the RGN accessory protein set forth as SEQ ID NO: 191 or an active variant or fragment thereof. In certain embodiments wherein the RGN is APG05123 (SEQ ID NO: 15) or a variant or fragment thereof, the system can further comprise the RGN accessory protein set forth as SEQ ID NO: 192 or an active variant or fragment thereof.

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 an RNA-guided

nuclease as the protein component. Such ribonucleoprotein complexes can be purified from a cell or organism that naturally expresses an RGN polypeptide and has been engineered to express a particular guide RNA that is specific for a target sequence of interest. Alternatively, the ribonucleoprotein complex can be purified from a cell or organism that has been transformed with polynucleotides that encode an RGN polypeptide and a guide RNA and cultured under conditions to allow for the expression of the RGN polypeptide and guide RNA. Thus, methods are provided for making an RGN polypeptide or an RGN ribonucleoprotein complex. Such methods comprise culturing a cell comprising a nucleotide sequence encoding an RGN polypeptide, and in some embodiments a nucleotide sequence encoding a guide RNA, under conditions in which the RGN polypeptide (and in some embodiments, the guide RNA) is expressed. The RGN polypeptide or RGN ribonucleoprotein can then be purified from a lysate of the cultured cells.

Methods for purifying an RGN polypeptide or RGN 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 RGN polypeptide 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, 10xHis, biotin carboxyl carrier protein (BCCP), and calmodulin. Generally, the tagged RGN polypeptide or RGN ribonucleoprotein complex is purified using immobilized metal affinity chromatography. It will be appreciated that other similar methods known in the art may be used, including other forms of chromatography or for example immunoprecipitation, either alone or in combination.

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.

Particular methods provided herein for binding and/or cleaving a target sequence of interest involve the use of an *in vitro* assembled RGN ribonucleoprotein complex. *In vitro* assembly of an RGN ribonucleoprotein complex can be performed using any method known in the art in which an RGN polypeptide is contacted with a guide RNA under conditions to allow for binding of the RGN polypeptide to the guide RNA. As used herein, "contact", "contacting", "contacted," refer to placing the components of a desired reaction together under conditions suitable for carrying out the desired reaction. The RGN polypeptide can be purified from a biological sample, cell lysate, or culture medium, produced via *in vitro*

translation, or chemically synthesized. The guide RNA can be purified from a biological sample, cell lysate, or culture medium, transcribed *in vitro*, or chemically synthesized. The RGN polypeptide and guide RNA can be brought into contact in solution (*e.g.*, buffered saline solution) to allow for *in vitro* assembly of the RGN ribonucleoprotein complex.

VIII. Methods of Binding, Cleaving, or Modifying a Target Sequence

The present disclosure provides methods for binding, cleaving, and/or modifying a target nucleotide sequence of interest. The methods include delivering a system comprising at least one guide RNA or a polynucleotide encoding the same, and at least one RGN 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 RGN comprises any one of the amino acid sequences of SEQ ID NOs: 1 to 109, or an active variant or fragment thereof. In various embodiments, the guide RNA comprises a CRISPR repeat sequence comprising any one of the nucleotide sequences of SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309, or an active variant or fragment thereof. In particular embodiments, the guide RNA comprises a tracrRNA comprising any one of the nucleotide sequences of SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148, or an active variant or fragment thereof. The guide RNA of the system can be a single guide RNA or a dual-guide RNA. The RGN of the system may be nuclease dead RGN, have nickase activity, or may be a fusion polypeptide. In some embodiments, the fusion polypeptide comprises a base-editing polypeptide, for example a cytidine deaminase or an adenosine deaminase. In other embodiments, the RGN fusion protein comprises a reverse transcriptase. In other embodiments, the RGN fusion protein comprises a polypeptide that recruits members of a functional nucleic acid repair complex, such as a member of the nucleotide excision repair (NER) or transcription coupled-nucleotide excision repair (TC-NER) pathway (Wei *et al.*, 2015, *PNAS USA* 112(27):E3495-504 ; Troelstra *et al.*, 1992, *Cell* 71:939-953; Marnef *et al.*, 2017, *J Mol Biol* 429(9):1277-1288), as described in U.S. Provisional Application No. 62/966,203, which was filed on January 27, 2020, and is incorporated herein by reference in its entirety. In some embodiments, the RGN fusion protein comprises CSB (van den Boom *et al.*, 2004, *J Cell Biol* 166(1):27-36; van Gool *et al.*, 1997, *EMBO J* 16(19):5955-65; an example of which is set forth as SEQ ID NO: 138), which is a member of the TC-NER (nucleotide excision repair) pathway and functions in the recruitment of other members. In further embodiments, the RGN fusion protein comprises an active domain of CSB, such as the acidic domain of CSB which comprises amino acid residues 356-394 of SEQ ID NO: 138 (Teng *et al.*, 2018, *Nat Commun* 9(1):4115).

In particular embodiments, the RGN and/or guide RNA is heterologous to the cell, organelle, or embryo to which the RGN and/or guide RNA (or polynucleotide(s) encoding at least one of the RGN and guide RNA) are introduced.

In some embodiments, the method further requires delivering at least one RGN accessory protein or polynucleotide(s) encoding the same in order for the RGN to bind to and/or cleave a target polynucleotide. In some of these embodiments, the method further requires delivering at least one RGN accessory protein set

forth as SEQ ID NOs: 178-192 or an active variant or fragment thereof, or polynucleotide(s) encoding the same. In particular embodiments wherein the RGN is APG06369 (SEQ ID NO: 11) or a variant or fragment thereof, the method further comprises delivering at least one RGN accessory protein set forth as SEQ ID NOs: 178-181 or an active variant or fragment thereof, or polynucleotide(s) encoding the same. In some of those embodiments wherein the RGN is APG03847 (SEQ ID NO: 12) or a variant or fragment thereof, the method further comprises delivering at least one RGN accessory protein set forth as SEQ ID NOs: 315-317 or an active variant or fragment thereof, or polynucleotide(s) encoding the same. In certain embodiments wherein the RGN is APG05625 (SEQ ID NO: 13) or a variant or fragment thereof, the method further comprises delivering at least one RGN accessory protein set forth as SEQ ID NOs: 185-187 or an active variant or fragment thereof, or polynucleotide(s) encoding the same. In some embodiments wherein the RGN is APG03524 (SEQ ID NO: 16) or a variant or fragment thereof, the method further comprises delivering at least one RGN accessory protein set forth as SEQ ID NOs: 188-190 or an active variant or fragment thereof, or polynucleotide(s) encoding the same. In particular embodiments wherein the RGN is APG03759 (SEQ ID NO: 14) or a variant or fragment thereof, the method further comprises delivering the RGN accessory protein set forth as SEQ ID NO: 191 or an active variant or fragment thereof, or polynucleotide(s) encoding the same. In certain embodiments wherein the RGN is APG05123 (SEQ ID NO: 15) or a variant or fragment thereof, the method further comprises delivering the RGN accessory protein set forth as SEQ ID NO: 192 or an active variant or fragment thereof, or polynucleotide(s) encoding the same.

In those embodiments wherein the method comprises delivering a polynucleotide encoding a guide RNA and/or an RGN polypeptide, the cell or embryo can then be cultured under conditions in which the guide RNA and/or RGN polypeptide are expressed. In various embodiments, the method comprises contacting a target sequence with an RGN ribonucleoprotein complex. The RGN ribonucleoprotein complex may comprise an RGN that is nuclease dead or has nickase activity. In some embodiments, the RGN of the ribonucleoprotein complex is a fusion polypeptide comprising a base-editing polypeptide. In certain embodiments, the method comprises introducing into a cell, organelle, or embryo comprising a target sequence an RGN ribonucleoprotein complex. The RGN ribonucleoprotein complex 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 RGN 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 RGN ribonucleoprotein complex can be introduced into a cell, organelle, or embryo using any method known in the art, including, but not limited to electroporation. Alternatively, an RGN polypeptide and/or 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 RGN to bind to the target sequence in a sequence-specific

manner. In those embodiments wherein the RGN has nuclease activity, the RGN polypeptide cleaves the target sequence of interest upon binding. The target sequence can subsequently be modified via endogenous repair mechanisms, such as non-homologous end joining, or homology-directed repair with a provided donor polynucleotide.

Methods to measure binding of an RGN polypeptide 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 single type of RGN 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.

In those embodiments wherein a donor polynucleotide is not provided, a double-stranded break introduced by an RGN polypeptide can be repaired by a non-homologous end-joining (NHEJ) repair process. Due to the error-prone nature of NHEJ, repair of the double-stranded break can result in a modification to the target sequence. As used herein, a "modification" in reference to a nucleic acid molecule refers to a change in the nucleotide sequence of the nucleic acid molecule, which can be a deletion, insertion, or substitution of one or more nucleotides, or a combination thereof. Modification of the target sequence can result in the expression of an altered protein product or inactivation of a coding sequence.

In those embodiments wherein a donor polynucleotide is present, the donor sequence in the donor polynucleotide can be integrated into or exchanged with the target nucleotide sequence during the course of repair of the introduced double-stranded break, resulting in the introduction of the exogenous donor sequence. A donor polynucleotide thus comprises a donor sequence that is desired to be introduced into a target sequence of interest. In some embodiments, the donor sequence alters the original target nucleotide sequence such that the newly integrated donor sequence will not be recognized and cleaved by the RGN. Integration of the donor sequence can be enhanced by the inclusion within the donor polynucleotide of flanking sequences, referred to herein as "homology arms" that have substantial sequence identity with the sequences flanking the target nucleotide sequence, allowing for a homology-directed repair process. In some embodiments, homology arms have a length of at least 50 base pairs, at least 100 base pairs, and up to 2000 base pairs or more, and have at least 90%, at least 95%, or more, sequence homology to their corresponding sequence within the target nucleotide sequence.

In those embodiments wherein the RGN polypeptide introduces double-stranded staggered breaks, the donor polynucleotide can comprise a donor sequence flanked by compatible overhangs, allowing for

direct ligation of the donor sequence to the cleaved target nucleotide sequence comprising overhangs by a non-homologous repair process during repair of the double-stranded break.

In those embodiments wherein the method involves the use of an RGN that is a nickase (*i.e.*, is only able to cleave a single strand of a double-stranded polynucleotide), the method can comprise introducing two RGN nickases 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.

In various embodiments, a method is provided for binding a target nucleotide sequence and detecting the target sequence, wherein the method comprises introducing into a cell, organelle, or embryo at least one guide RNA or a polynucleotide encoding the same, and at least one RGN polypeptide or a polynucleotide encoding the same, expressing the guide RNA and/or RGN polypeptide (if coding sequences are introduced), wherein the RGN polypeptide is a nuclease-dead RGN and further comprises a detectable label, and the method further comprises detecting the detectable label. The detectable label may be fused to the RGN as a fusion protein (*e.g.*, fluorescent protein) or may be a small molecule conjugated to or incorporated within the RGN polypeptide that can be detected visually or by other means.

Also provided herein are methods for modulating the expression of a target sequence or a gene of interest under the regulation of a target sequence. The methods comprise introducing into a cell, organelle, or embryo at least one guide RNA or a polynucleotide encoding the same, and at least one RGN polypeptide or a polynucleotide encoding the same, expressing the guide RNA and/or RGN polypeptide (if coding sequences are introduced), wherein the RGN polypeptide is a nuclease-dead RGN. In some of these embodiments, the nuclease-dead RGN is a fusion protein comprising an expression modulator domain (*i.e.*, epigenetic modification domain, transcriptional activation domain or a transcriptional repressor domain) as described herein.

The present disclosure also provides methods for binding and/or modifying a target nucleotide sequence 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 polypeptide comprises an RGN of the invention and a base-editing polypeptide, for example a cytidine deaminase or an adenosine deaminase, or a polynucleotide encoding the fusion polypeptide, to the target sequence or a cell, organelle, or embryo comprising the target sequence.

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 single RGN polypeptide in combination with multiple, distinct guide RNAs, which can target multiple, distinct sequences within a single gene and/or multiple genes. Also encompassed herein are methods wherein multiple, distinct guide RNAs are introduced in combination with multiple, distinct RGN polypeptides. These guide RNAs and guide RNA/RGN polypeptide systems can target multiple, distinct sequences within a single gene and/or multiple genes.

In one aspect, the invention provides kits containing any one or more of the elements disclosed in the above methods and compositions. In some embodiments, the kit comprises a vector system and instructions for using the kit. In some embodiments, the vector system comprises (a) a first regulatory element operably linked to a DNA sequence encoding the crRNA sequence and one or more insertion sites for inserting a guide sequence upstream of the encoded crRNA sequence, wherein when expressed, the guide sequence directs sequence-specific binding of a CRISPR complex to a target sequence in a eukaryotic cell, wherein the CRISPR complex comprises a CRISPR enzyme complexed with (a) the guide RNA polynucleotide; and/or (b) a second regulatory element operably linked to an enzyme coding sequence encoding said CRISPR enzyme comprising a nuclear localization sequence.

In some embodiments, the kit comprises one or more oligonucleotides corresponding to a guide sequence for insertion into a vector so as to operably link the guide sequence and a regulatory element. In some embodiments, the kit comprises a homologous recombination template polynucleotide. In one aspect, the invention provides methods for using one or more elements of a CRISPR system. The CRISPR complex of the invention provides an effective means for modifying a target polynucleotide. The CRISPR complex of the invention has a wide variety of utility including modifying (e.g., deleting, inserting, translocating, inactivating, activating, base editing) a target polynucleotide in a multiplicity of cell types. As such the CRISPR complex of the invention has a broad spectrum of applications in, e.g., gene therapy, drug screening, disease diagnosis, and prognosis. An exemplary CRISPR complex comprises a CRISPR enzyme complexed with a guide sequence hybridized to a target sequence within the target polynucleotide.

IX. Target Polynucleotides

In one aspect, the invention provides for methods of modifying a target polynucleotide in a eukaryotic cell, which may be *in vivo*, *ex vivo* or *in vitro*. In some embodiments, the method comprises sampling a cell or population of cells from a human or non-human animal or plant (including microalgae) and modifying the cell or cells. Culturing may occur at any stage *ex vivo*. The cell or cells may even be re-introduced into the non-human animal or plant (including micro-algae).

Using natural variability, plant breeders combine most useful genes for desirable qualities, such as yield, quality, uniformity, hardiness, and resistance against pests. These desirable qualities also include growth, day length preferences, temperature requirements, initiation date of floral or reproductive development, fatty acid content, insect resistance, disease resistance, nematode resistance, fungal resistance, herbicide resistance, tolerance to various environmental factors including drought, heat, wet, cold, wind, and adverse soil conditions including high salinity The sources of these useful genes include native or foreign varieties, heirloom varieties, wild plant relatives, and induced mutations, e.g., treating plant material with mutagenic agents. Using the present invention, plant breeders are provided with a new tool to induce mutations. Accordingly, one skilled in the art can analyze the genome for sources of useful genes, and in varieties having desired characteristics or traits employ the present invention to induce the rise of useful

genes, with more precision than previous mutagenic agents and hence accelerate and improve plant breeding programs.

The target polynucleotide of an RGN system can be any polynucleotide endogenous or exogenous to the eukaryotic cell. For example, the target polynucleotide can be a polynucleotide residing in the nucleus of the eukaryotic cell. The target polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., a regulatory polynucleotide or a junk DNA). In some embodiments, the target sequence is associated with a PAM (protospacer adjacent motif); that is, a short sequence recognized by the CRISPR complex. The precise sequence and length requirements for the PAM differ depending on the CRISPR enzyme used (and in some embodiments, the RGN does not require a PAM sequence), but PAMs are typically 2-5 base pair sequences adjacent the protospacer (that is, the target sequence).

The target polynucleotide of a CRISPR complex may include a number of disease-associated genes and polynucleotides as well as signaling biochemical pathway-associated genes and polynucleotides. Examples of target polynucleotides include a sequence associated with a signaling biochemical pathway, e.g., a signaling biochemical pathway-associated gene or polynucleotide. Examples of target polynucleotides include a disease associated gene or polynucleotide. A "disease-associated" gene or polynucleotide refers to any gene or polynucleotide which is yielding transcription or translation products at an abnormal level or in an abnormal form in cells derived from a disease-affected tissues compared with tissues or cells of a non-disease control. It may be a gene that becomes expressed at an abnormally high level; it may be a gene that becomes expressed at an abnormally low level, where the altered expression correlates with the occurrence and/or progression of the disease. A disease-associated gene also refers to a gene possessing mutation(s) or genetic variation that is directly responsible or is in linkage disequilibrium with a gene(s) that is responsible for the etiology of a disease (e.g., a causal mutation). The transcribed or translated products may be known or unknown, and further may be at a normal or abnormal level. Examples of disease-associated genes and polynucleotides are available from McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, Md.) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, Md.), available on the World Wide Web.

Although CRISPR systems are particularly useful for their relative ease in targeting to genomic sequences of interest, there still remains an issue of what the RGN can do to address a causal mutation. One approach is to produce a fusion protein between an RGN (preferably an inactive or nickase variant of the RGN) and a base-editing enzyme or the active domain of a base editing enzyme, such as a cytidine deaminase or an adenosine deaminase base editor (U.S. Patent No. 9,840, 699, herein incorporated by reference). In some embodiments, the methods comprise contacting a DNA molecule with (a) a fusion protein comprising an RGN of the invention and a base-editing polypeptide such as a deaminase; and (b) a gRNA targeting the fusion protein of (a) to a target nucleotide sequence of the DNA strand; 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 nucleobase. In some embodiments, the target DNA sequence comprises a sequence associated with a disease or disorder, and wherein the deamination of the nucleobase

results in a sequence that is not associated with a disease or disorder. 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 nucleobase results in an allele that improves the trait and increases the agronomic value of the plant.

In some embodiments, the DNA sequence comprises a T→C or A→G point mutation associated with a disease or disorder, and wherein the deamination of the mutant C or G 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.

X. Pharmaceutical Compositions and Methods of Treatment

Pharmaceutical compositions comprising the presently disclosed RGN polypeptides and active variants and fragments thereof, as well as polynucleotides encoding the same, the presently disclosed gRNAs or polynucleotides encoding the same, the presently disclosed systems, or cells comprising any of the RGN polypeptides or RGN-encoding polynucleotides, gRNA or gRNA-encoding polynucleotides, or the RGN systems, and a pharmaceutically acceptable carrier are provided.

A pharmaceutical composition is a composition that is employed to prevent, reduce in intensity, cure or otherwise treat a target condition or disease that comprises an active ingredient (i.e., RGN polypeptides, RGN-encoding polynucleotides, gRNA, gRNA-encoding polynucleotides, RGN systems, or cells comprising any one of these) and a pharmaceutically acceptable carrier.

As used herein, a "pharmaceutically acceptable carrier" refers to a material that does not cause significant irritation to an organism and does not abrogate the activity and properties of the active ingredient (i.e., RGN polypeptides, RGN-encoding polynucleotides, gRNA, gRNA-encoding polynucleotides, RGN systems, or cells comprising any one of these). Carriers must be of sufficiently high purity and of sufficiently low toxicity to render them suitable for administration to a subject being treated. The carrier can be inert, or it can possess pharmaceutical benefits. In some embodiments, a pharmaceutically acceptable carrier comprises 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 pharmaceutically acceptable carrier is not naturally-occurring. In some embodiments, the pharmaceutically acceptable carrier and the active ingredient are not found together in nature.

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. See, *e.g.*, Remington, The Science and Practice of Pharmacy (21st ed. 2005). Suitable formulations include, for example, powders, pastes, ointments, jellies, waxes, oils, lipids, lipid (cationic or anionic) containing vesicles (such as LIPOFECTIN vesicles), lipid nanoparticles, DNA conjugates, anhydrous absorption pastes, oil-in-water and water-in-oil emulsions, emulsions carbowax (polyethylene glycols of various molecular weights), semi-solid gels, and semi-solid mixtures containing carbowax. 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.

In some embodiments wherein cells comprising or modified with the presently disclosed RGN, gRNAs, RGN systems or polynucleotides encoding the same are administered to a subject, the cells are administered as a suspension with a pharmaceutically acceptable carrier. One of skill in the art will recognize that a pharmaceutically acceptable carrier to be used in a cell composition will not include buffers, compounds, cryopreservation agents, preservatives, or other agents in amounts that substantially interfere with the viability of the cells to be delivered to the subject. A formulation comprising cells can include e.g., osmotic buffers that permit cell membrane integrity to be maintained, and optionally, nutrients to maintain cell viability or enhance engraftment upon administration. Such formulations and suspensions are known to those of skill in the art and/or can be adapted for use with the cells described herein using routine experimentation.

A cell composition can also be emulsified or presented as a liposome composition, provided that the emulsification procedure does not adversely affect cell viability. The cells and any other active ingredient can be mixed with excipients that are pharmaceutically acceptable and compatible with the active ingredient, and in amounts suitable for use in the therapeutic methods described herein.

Additional agents included in a cell composition can include pharmaceutically acceptable salts of the components therein. Pharmaceutically acceptable salts include the acid addition salts (formed with the free amino groups of the polypeptide) that are formed with inorganic acids, such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, tartaric, mandelic and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases, such as, for example, sodium, potassium, ammonium, calcium or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine and the like.

Physiologically tolerable and pharmaceutically acceptable carriers are well known in the art. Exemplary liquid carriers are sterile aqueous solutions that contain no materials in addition to the active ingredients and water, or contain a buffer such as sodium phosphate at physiological pH value, physiological saline or both, such as phosphate-buffered saline. Still further, aqueous carriers can contain more than one buffer salt, as well as salts such as sodium and potassium chlorides, dextrose, polyethylene glycol and other solutes. Liquid compositions can also contain liquid phases in addition to and to the exclusion of water.

Exemplary of such additional liquid phases are glycerin, vegetable oils such as cottonseed oil, and water-oil emulsions. The amount of an active compound used in the cell compositions that is effective in the treatment of a particular disorder or condition can depend on the nature of the disorder or condition, and can be determined by standard clinical techniques.

The presently disclosed RGN polypeptides, guide RNAs, RGN systems or polynucleotides encoding the same can be formulated with pharmaceutically acceptable excipients such as carriers, solvents, stabilizers, adjuvants, diluents, etc., depending upon the particular mode of administration and dosage form. In some embodiments, these pharmaceutical compositions are formulated to achieve a physiologically compatible pH, and range from a pH of about 3 to a pH of about 11, about pH 3 to about pH 7, depending on the formulation and route of administration. In some embodiments, the pH can be adjusted to a range from about pH 5.0 to about pH 8. In some embodiments, the compositions can comprise a therapeutically effective amount of at least one compound as described herein, together with one or more pharmaceutically acceptable excipients. In some embodiments, the compositions comprise a combination of the compounds described herein, or include a second active ingredient useful in the treatment or prevention of bacterial growth (for example and without limitation, anti-bacterial or anti-microbial agents), or include a combination of reagents of the present disclosure.

Suitable excipients include, for example, carrier molecules that include large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Other exemplary excipients can include antioxidants (for example and without limitation, ascorbic acid), chelating agents (for example and without limitation, EDTA), carbohydrates (for example and without limitation, dextrin, hydroxyalkylcellulose, and hydroxyalkylmethylcellulose), stearic acid, liquids (for example and without limitation, oils, water, saline, glycerol and ethanol), wetting or emulsifying agents, pH buffering substances, and the like.

In some embodiments, the formulations are provided in unit-dose or multi-dose containers, for example sealed ampules and vials, and may be stored in a freeze-dried (lyophilized) condition requiring the addition of the sterile liquid carrier, for example, saline, water-for-injection, a semi-liquid foam, or gel, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets of the kind previously described. In some embodiments, the active ingredient is dissolved in a buffered liquid solution that is frozen in a unit-dose or multi-dose container and later thawed for injection or kept/stabilized under refrigeration until use.

The therapeutic agent(s) may be contained in controlled release systems. In order to prolong the effect of a drug, it often is desirable to slow the absorption of the drug from subcutaneous, intrathecal, or intramuscular injection. This may be accomplished by the use of a liquid suspension of crystalline or amorphous material with poor water solubility. The rate of absorption of the drug then depends upon its rate of dissolution which, in turn, may depend upon crystal size and crystalline form. Alternatively, delayed absorption of a parenterally administered drug form is accomplished by dissolving or suspending the drug in an oil vehicle. In some embodiments, the use of a long-term sustained release implant may be particularly

suitable for treatment of chronic conditions. Long-term sustained release implants are well-known to those of ordinary skill in the art.

Methods of treating a disease in a subject in need thereof are provided herein. The methods comprise administering to a subject in need thereof an effective amount of a presently disclosed RGN polypeptide or active variant or fragment thereof or a polynucleotide encoding the same, a presently disclosed gRNA or a polynucleotide encoding the same, a presently disclosed RGN system, or a cell modified by or comprising any one of these compositions.

In some embodiments, the treatment comprises *in vivo* gene editing by administering a presently disclosed RGN polypeptide, gRNA, or RGN system or polynucleotide(s) encoding the same. In some embodiments, the treatment comprises *ex vivo* gene editing wherein cells are genetically modified *ex vivo* with a presently disclosed RGN polypeptide, gRNA, or RGN system or polynucleotide(s) encoding the same and then the modified cells are administered to a subject. In some embodiments, the genetically modified cells originate from the subject that is then administered the modified cells, and the transplanted cells are referred to herein as autologous. In some embodiments, the genetically modified cells originate from a different subject (i.e., donor) within the same species as the subject that is administered the modified cells (i.e., recipient), and the transplanted cells are referred to herein as allogeneic. In some examples described herein, the cells can be expanded in culture prior to administration to a subject in need thereof.

In some embodiments, the disease to be treated with the presently disclosed compositions is one that can be treated with immunotherapy, such as with a chimeric antigen receptor (CAR) T cell. Such diseases include but are not limited to cancer.

In some embodiments, the disease to be treated with the presently disclosed compositions is 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. In some embodiments, the disease to be treated with the presently disclosed compositions is associated with a causal mutation. As used herein, a "causal mutation" refers to a particular nucleotide, nucleotides, or nucleotide sequence in the genome that contributes to the severity or presence of a disease or disorder in a subject. The correction of the causal mutation leads to the improvement of at least one symptom resulting from a disease or disorder. In some embodiments, the causal mutation is adjacent to a PAM site recognized by an RGN disclosed herein. The causal mutation can be corrected with a presently disclosed RGN or a fusion polypeptide comprising a presently disclosed RGN and a base-editing polypeptide (i.e., a base editor). Non-limiting examples of diseases associated with a causal mutation include cystic fibrosis, Hurler syndrome, Friedreich's Ataxia, Huntington's Disease, and sickle cell disease. Additional non-limiting examples of disease-associated genes and mutations are available from McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, Md.) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, Md.), available on the World Wide Web.

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 presently disclosed composition 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 provided herein is to restore the function of a dysfunctional gene via genome editing. The presently disclosed RGN polypeptides and systems comprising the same can be validated for gene editing-based human therapeutics in vitro, *e.g.*, by correcting a disease associated mutation in human cell culture.

As used herein, "treatment" or "treating," or "palliating" or "ameliorating" are used interchangeably. These terms refer to an approach for obtaining beneficial or desired results including but not limited to a therapeutic benefit and/or a prophylactic benefit. By therapeutic benefit is meant any therapeutically relevant improvement in or effect on one or more diseases, conditions, or symptoms under treatment. For prophylactic benefit, the compositions may be administered to a subject at risk of developing a particular disease, condition, or symptom, or to a subject reporting one or more of the physiological symptoms of a disease, even though the disease, condition, or symptom may not have yet been manifested.

The term "effective amount" or "therapeutically effective amount" refers to the amount of an agent that is sufficient to effect beneficial or desired results. The therapeutically effective amount may vary depending upon one or more of: the subject and disease condition being treated, the weight and age of the subject, the severity of the disease condition, the manner of administration and the like, which can readily be determined by one of ordinary skill in the art. The specific dose may vary depending on one or more of: the particular agent chosen, the dosing regimen to be followed, whether it is administered in combination with other compounds, timing of administration, and the delivery system in which it is carried.

The term "administering" refers to the placement of an active ingredient into a subject, by a method or route that results in at least partial localization of the introduced active ingredient at a desired site, such as a site of injury or repair, such that a desired effect(s) is produced. In those embodiments wherein cells are administered, the cells can be administered by any appropriate route that results in delivery to a desired location in the subject where at least a portion of the implanted cells or components of the cells remain viable. The period of viability of the cells after administration to a subject can be as short as a few hours, e.g., twenty-four hours, to a few days, to as long as several years, or even the life time of the patient, i.e., long-term engraftment. For example, in some aspects described herein, an effective amount of photoreceptor cells or retinal progenitor cells is administered via a systemic route of administration, such as an intraperitoneal or intravenous route.

In some embodiments, the administering comprises administering by viral delivery. In some embodiments, the administering comprises administering by electroporation. In some embodiments, the administering comprises administering by nanoparticle delivery. In some embodiments, the administering

comprises administering by liposome delivery. Any effective route of administration can be used to administer an effective amount of a pharmaceutical composition described herein. In some embodiments, the administering comprises administering by a method selected from the group consisting of: intravenously, subcutaneously, intramuscularly, orally, rectally, by aerosol, parenterally, ophthalmicly, pulmonarily, transdermally, vaginally, otically, nasally, and by topical administration, or any combination thereof. In some embodiments, for the delivery of cells, administration by injection or infusion is used.

As used herein, the term "subject" refers to any individual for whom diagnosis, treatment or therapy is desired. In some embodiments, the subject is an animal. In some embodiments, the subject is a mammal. In some embodiments, the subject is a human being.

The efficacy of a treatment can be determined by the skilled clinician. However, a treatment is considered an "effective treatment," if any one or all of the signs or symptoms of a disease or disorder are altered in a beneficial manner (e.g., decreased by at least 10%), or other clinically accepted symptoms or markers of disease are improved or ameliorated. Efficacy can also be measured by failure of an individual to worsen as assessed by hospitalization or need for medical interventions (e.g., progression of the disease is halted or at least slowed). Methods of measuring these indicators are known to those of skill in the art. Treatment includes: (1) inhibiting the disease, e.g., arresting, or slowing the progression of symptoms; or (2) relieving the disease, e.g., causing regression of symptoms; and (3) preventing or reducing the likelihood of the development of symptoms.

A. Modifying causal mutations using base-editing

In some embodiments, RGNs of the invention are used to modify causal mutations using base-editing. An example of a genetically inherited disease which could be corrected using an approach that relies on an RGN-base editor fusion protein of the invention is Hurler Syndrome. Hurler Syndrome, also known as MPS-1, is the result of a deficiency of α-L-iduronidase (IDUA) resulting in a lysosomal storage disease characterized at the molecular level by the accumulation of dermatan sulfate and heparan sulfate in lysosomes. This disease is generally an inherited genetic disorder caused by mutations in the IDUA gene encoding α-L-iduronidase. Common IDUA mutations are W402X and Q70X, both nonsense mutations resulting in premature termination of translation. Such mutations are well addressed by precise genome editing (PGE) approaches, since reversion of a single nucleotide, for example by a base-editing approach, would restore the wild-type coding sequence and result in protein expression controlled by the endogenous regulatory mechanisms of the genetic locus. Additionally, since heterozygotes are known to be asymptomatic, a PGE therapy that targets one of these mutations would be useful to a large proportion of patients with this disease, as only one of the mutated alleles needs to be corrected (Bunge et al. (1994) Hum. Mol. Genet. 3(6): 861-866, herein incorporated by reference).

Current treatments for Hurler Syndrome include enzyme replacement therapy and bone marrow transplants (Vellodi et al. (1997) Arch. Dis. Child. 76(2): 92-99; Peters et al. (1998) Blood 91(7): 2601-2608, herein incorporated by reference). While enzyme replacement therapy has had a dramatic effect on

the survival and quality of life of Hurler Syndrome patients, this approach requires costly and time-consuming weekly infusions. Additional approaches include the delivery of the IDUA gene on an expression vector or the insertion of the gene into a highly expressed locus such as that of serum albumin (U.S. Patent No. 9,956,247, herein incorporated by reference). However, these approaches do not restore the original IDUA locus to the correct coding sequence. A genome-editing strategy would have a number of advantages, most notably that regulation of gene expression would be controlled by the natural mechanisms present in healthy individuals. Additionally, using base editing does not necessitate causing a double stranded DNA breaks, which could lead to large chromosomal rearrangements, cell death, or oncogenecity by the disruption of tumor suppression mechanisms. A general strategy may be directed toward using RGN-base editor fusion proteins of the invention to target and correct certain disease-causing mutations in the human genome. It will be appreciated that similar approaches to target diseases that can be corrected by base-editing may also be pursued. It will be further appreciated that similar approaches to target disease-causing mutations in other species, particularly common household pets or livestock, can also be deployed using the RGNs of the invention. Common household pets and livestock include dogs, cats, horses, pigs, cows, sheep, chickens, donkeys, snakes, ferrets, and fish including salmon and shrimp.

B. Modifying causal mutations by targeted deletion

RGNs of the invention could also be useful in human therapeutic approaches where the causal mutation is more complicated. For example, some diseases such as Friedreich's Ataxia and Huntington's Disease are the result of a significant increase in repeats of a three nucleotide motif at a particular region of a gene, which affects the ability of the expressed protein to function or to be expressed. Friedreich's Ataxia (FRDA) is an autosomal recessive disease resulting in progressive degeneration of nervous tissue in the spinal cord. Reduced levels of the frataxin (FXN) protein in the mitochondria cause oxidative damages and iron deficiencies at the cellular level. The reduced FXN expression has been linked to a GAA triplet expansion within the intron 1 of the somatic and germline FXN gene. In FRDA patients, the GAA repeat frequently consists of more than 70, sometimes even more than 1000 (most commonly 600-900) triplets, whereas unaffected individuals have about 40 repeats or less (Pandolfo et al. (2012) Handbook of Clinical Neurology 103: 275-294; Campuzano et al. (1996) Science 271: 1423-1427; Pandolfo (2002) Adv. Exp. Med. Biol. 516: 99-118; all herein incorporated by reference).

The expansion of the trinucleotide repeat sequence causing Friedreich's Ataxia (FRDA) occurs in a defined genetic locus within the FXN gene, referred to as the FRDA instability region. RNA guided nucleases (RGNs) may be used for excising the instability region in FRDA patient cells. This approach requires 1) an RGN and guide RNA sequence that can be programmed to target the allele in the human genome; and 2) a delivery approach for the RGN and guide sequence. Many nucleases used for genome editing, such as the commonly used Cas9 nuclease from *S. pyogenes* (SpCas9), are too large to be packaged into adeno-associated viral (AAV) vectors, especially when considering the length of the SpCas9 gene and

the guide RNA in addition to other genetic elements required for functional expression cassettes. This makes an approach using SpCas9 more difficult.

Certain RNA guided nucleases of the invention are well suited for packaging into an AAV vector along with a guide RNA. Packing two guide RNAs would likely require a second vector, but this approach still compares favorably to what would be required of a larger nuclease such as SpCas9, which may require splitting the protein sequence between two vectors. The present invention encompasses a strategy using RGNs of the invention in which a region of genomic instability is removed. Such a strategy is applicable to other diseases and disorders which have a similar genetic basis, such as Huntington's Disease. Similar strategies using RGNs of the invention may also be applicable to similar diseases and disorders in non-human animals of agronomic or economic importance, including dogs, cats, horses, pigs, cows, sheep, chickens, donkeys, snakes, ferrets, and fish including salmon and shrimp.

C. Modifying causal mutations by targeted mutagenesis

RGNs of the invention could also be to introduce disruptive mutations that may result in a beneficial effect. Genetic defects in the genes encoding hemoglobin, particularly the beta globin chain (the HBB gene), can be responsible for a number of diseases known as hemoglobinopathies, including sickle cell anemia and thalassemias.

In adult humans, hemoglobin is a heterotetramer comprising two alpha (α)-like globin chains and two beta (β)-like globin chains and 4 heme groups. In adults the α2β2 tetramer is referred to as Hemoglobin A (HbA) or adult hemoglobin. Typically, the alpha and beta globin chains are synthesized in an approximate 1:1 ratio and this ratio seems to be critical in terms of hemoglobin and red blood cell (RBC) stabilization. In a developing fetus, a different form of hemoglobin, fetal hemoglobin (HbF), is produced which has a higher binding affinity for oxygen than Hemoglobin A such that oxygen can be delivered to the baby's system via the mother's blood stream. Fetal hemoglobin also contains two α globin chains, but in place of the adult β-globin chains, it has two fetal gamma (γ)-globin chains (i.e., fetal hemoglobin is α2γ2). The regulation of the switch from production of gamma- to beta-globin is quite complex, and primarily involves a down-regulation of gamma globin transcription with a simultaneous up-regulation of beta globin transcription. At approximately 30 weeks of gestation, the synthesis of gamma globin in the fetus starts to drop while the production of beta globin increases. By approximately 10 months of age, the newborn's hemoglobin is nearly all α2β2 although some HbF persists into adulthood (approximately 1-3% of total hemoglobin). In the majority of patients with hemoglobinopathies, the genes encoding gamma globin remain present, but expression is relatively low due to normal gene repression occurring around parturition as described above.

Sickle cell disease is caused by a V6E mutation in the β globin gene (HBB) (a GAG to GTG at the DNA level), where the resultant hemoglobin is referred to as "hemoglobinS" or "HbS." Under lower oxygen conditions, HbS molecules aggregate and form fibrous precipitates. These aggregates cause the abnormality or 'sickling' of the RBCs, resulting in a loss of flexibility of the cells. The sickling RBCs are no longer able to squeeze into the capillary beds and can result in vaso-occlusive crisis in sickle cell patients.

In addition, sickled RBCs are more fragile than normal RBCs, and tend towards hemolysis, eventually leading to anemia in the patient.

Treatment and management of sickle cell patients is a life-long proposition involving antibiotic treatment, pain management and transfusions during acute episodes. One approach is the use of hydroxyurea, which exerts its effects in part by increasing the production of gamma globin. Long term side effects of chronic hydroxyurea therapy are still unknown, however, and treatment gives unwanted side effects and can have variable efficacy from patient to patient. Despite an increase in the efficacy of sickle cell treatments, the life expectancy of patients is still only in the mid to late 50's and the associated morbidities of the disease have a profound impact on a patient's quality of life.

Thalassemias (alpha thalassemias and beta thalassemia) are also diseases relating to hemoglobin and typically involve a reduced expression of globin chains. This can occur through mutations in the regulatory regions of the genes or from a mutation in a globin coding sequence that results in reduced expression or reduced levels or functional globin protein. Treatment of thalassemias usually involves blood transfusions and iron chelation therapy. Bone marrow transplants are also being used for treatment of people with severe thalassemias if an appropriate donor can be identified, but this procedure can have significant risks.

One approach that has been proposed for the treatment of both sickle cell disease (SCD) and beta thalassemias is to increase the expression of gamma globin so that HbF functionally replaces the aberrant adult hemoglobin. As mentioned above, treatment of SCD patients with hydroxyurea is thought to be successful in part due to its effect on increasing gamma globin expression (DeSimone (1982) Proc Nat'l Acad Sci USA 79(14):4428-31; Ley, et al., (1982) N. Engl. J. Medicine, 307: 1469-1475; Ley, et al., (1983) Blood 62: 370-380; Constantoulakis et al., (1988) Blood 72(6):1961-1967, all herein incorporated by reference). Increasing the expression of HbF involves identification of genes whose products play a role in the regulation of gamma globin expression. One such gene is BCL11A. BCL11A encodes a zinc finger protein that expressed in adult erythroid precursor cells, and down-regulation of its expression leads to an increase in gamma globin expression (Sankaran et at (2008) Science 322: 1839, herein incorporated by reference). Use of an inhibitory RNA targeted to the BCL11A gene has been proposed (e.g., U.S. Patent Publication 2011/0182867, herein incorporated by reference) but this technology has several potential drawbacks, including that complete knock down may not be achieved, delivery of such RNAs may be problematic, and the RNAs must be present continuously, requiring multiple treatments for life.

RGNs of the invention may be used to target the BCL11A enhancer region to disrupt expression of BCL11A, thereby increasing gamma globin expression. This targeted disruption can be achieved by non-homologous end joining (NHEJ), whereby an RGN of the invention targets to a particular sequence within the BCL11A enhancer region, makes a double-stranded break, and the cell's machinery repairs the break, typically simultaneously introducing deleterious mutations. Similar to what is described for other disease targets, RGNs of the invention may have advantages over other known RGNs due to their relatively small size, which enables packaging expression cassettes for the RGN and its guide RNA into a single AAV vector

for *in vivo* delivery. Similar strategies using RGNs of the invention may also be applicable to similar diseases and disorders in both humans and in non-human animals of agronomic or economic importance.

XI. Cells Comprising a Polynucleotide Genetic Modification

Provided herein are cells and organisms comprising a target sequence of interest that has been modified using a process mediated by an RGN, crRNA, and/or tracrRNA as described herein. In some of these embodiments, the RGN comprises any one of the amino acid sequences of SEQ ID NOs: 1 to 109, or an active variant or fragment thereof. In various embodiments, the guide RNA comprises a CRISPR repeat sequence comprising any one of the nucleotide sequences of SEQ ID NOs: 110 to 119, 139, 141, 143, 146, and 201 to 309, or an active variant or fragment thereof. In particular embodiments, the guide RNA comprises a tracrRNA comprising any one of the nucleotide sequences of SEQ ID NOs: 120 to 128, 140, 142, 145, 147, and 148, or an active variant or fragment thereof. The guide RNA of the system can be a single guide RNA or a dual-guide RNA.

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

The chromosomal modification of the cell, organism, organelle, or embryo 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 embodiments wherein the chromosomal modification 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 some embodiments, the chromosomal modification of a cell, organism, organelle, or embryo can produce a "knock in", which results from the chromosomal integration of a nucleotide sequence that encodes a protein. In some of these embodiments, the coding sequence is integrated into the chromosome such that the chromosomal sequence encoding the wild-type protein is inactivated, but the exogenously introduced protein is expressed.

In some embodiments, the chromosomal modification 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 encoded by the altered chromosomal sequence 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 some embodiments, the chromosomal modification can result in an altered expression pattern of a protein. As a non-limiting example, chromosomal alterations 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.

The cells that have been modified can be grown into an 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 modified strain or different strains, and the resulting hybrid having the genetic modification. The present invention provides genetically modified seed. Progeny, variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the genetic modification. Further provided is a processed plant product or byproduct that retains the genetic modification, including for example, soymeal.

The methods provided herein may be used for modification 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.).

The methods provided herein can also be used to genetically modify any prokaryotic species, including but not limited to, archaea and bacteria (*e.g.*, *Bacillus* sp., *Klebsiella* sp. *Streptomyces* sp., *Rhizobium* sp., *Escherichia* sp., *Pseudomonas* sp., *Salmonella* sp., *Shigella* sp., *Vibrio* sp., *Yersinia* sp., *Mycoplasma* sp., Agrobacterium, *Lactobacillus* sp.).

The methods provided herein can be used to genetically modify any eukaryotic species or cells therefrom, including but not limited to animals (*e.g.*, mammals, insects, fish, birds, and reptiles), fungi, amoeba, algae, and yeast. In some embodiments, the cell that is modified by the presently disclosed methods include cells of hematopoietic origin, such as cells of the immune system (i.e., immune cells) including but not limited to B cells, T cells, natural killer (NK) cells, stem cells including pluripotent stem cells and induced pluripotent stem cells, chimeric antigen receptor T (CAR-T) cells, monocytes, macrophages, and dendritic cells.

Cells that have been modified may be introduced into an organism. These cells could have originated from the same organism (e.g., person) in the case of autologous cellular transplants, wherein the cells are modified in an *ex vivo* approach. Alternatively, the cells originated from another organism within

the same species (e.g., another person) in the case of allogeneic cellular transplants.

XII. Kits and Methods for Detecting Target DNA or Cleaving a Population of Single-Stranded DNA

The presently disclosed RGNs, particularly, APG09624 and APG05405 (set forth as SEQ ID NOs: 2 and 4), can promiscuously cleave non-targeted single-stranded DNA (ssDNA) once activated by detection of a target DNA. Thus, provided herein are compositions and methods for detecting a target DNA (double-stranded or single-stranded) in a sample.

Methods of detecting a target DNA of a DNA molecule comprise contacting a sample with an RGN (or a polynucleotide encoding the same), a guide RNA (or a polynucleotide encoding the same) capable of hybridizing with the RGN and a target DNA sequence in a DNA molecule, and a detector single-stranded DNA (detector ssDNA) that does not hybridize with the guide RNA, followed by measuring a detectable signal produced by cleavage of the ssDNA by the RGN, thereby detecting the target DNA sequence of the DNA molecule. In some embodiments, the method can comprise a step of amplification of the nucleic acid molecules within a sample, either before or simultaneously with contact with the RGN and guideRNA. In some of these embodiments, specific sequences to which the guide RNA will hybridize can be amplified in order to increase sensitivity of a detection method.

In those embodiments wherein a sample is contacted with a polynucleotide encoding an RGN polypeptide and/or a polynucleotide encoding the guide RNA, the sample comprises intact cells and the polynucleotides are introduced into the cells in which they are then expressed. In some of these embodiments, at least one of the polynucleotides further comprises a promoter that is operably linked to the nucleotide sequence encoding the RGN polypeptide and/or guide RNA.

In some embodiments, the desired target may exist as RNA, such as the genome or part of a genome of an RNA virus, such as for example a coronavirus. In some embodiments, the coronavirus may be a SARS-like coronavirus. In further embodiments, the coronavirus may be SARS-CoV-2, SARS-CoV, or a bat SARS-like coronavirus such as bat-SL-CoVZC45 (accession MG772933). In embodiments where the target exists as RNA, the target may be reverse-transcribed into a DNA molecule which can be effectively targeted by the RGN. Reverse-transcription may be followed by an amplification step, such as RT-PCR methods known in the art, which involve thermocycling, or may be by isothermal methods such as RT-LAMP (reverse transcription loop-mediated isothermal amplification) (Notomi et al., *Nucleic Acids Res* 28: E63, (2000)).

The nucleic acid amplification can occur before the sample is contacted with the RGN, guide RNA, and detector ssDNA or amplification can occur simultaneously with the contacting step.

In certain embodiments, the method involves contacting a sample with an RGN and more than one guide RNA. The guide RNAs, each capable of hybridizing with the RGN, can bind to unique target sequences of a single DNA molecule in order to amplify the detectable signal and lead to the detection of that DNA molecule.

These compositions and methods involve the use of a detector ssDNA that does not hybridize with the guideRNA and is a non-target ssDNA. In some embodiments, the detector ssDNA comprises a detectable label that provides a detectable signal after cleavage of the detector ssDNA. A non-limiting example is a detector ssDNA that comprises a fluorophore/quencher pair wherein the fluorophore does not fluoresce when the detector ssDNA is whole (i.e., uncleaved) as its signal is suppressed by the presence of the quencher in close proximity. Cleavage of the detector ssDNA results in removal of the quencher and the fluorescent label can then be detected. Non-limiting examples of fluorescent labels or dyes include Cy5, fluorescein (e.g., FAM, 6 FAM, 5(6) FAM, FITC), Cy3, Alexa Fluor® dyes, and Texas Red. Non-limiting examples of quenchers include Iowa Black®FQ, Iowa Black® RQ, a Qx1 quencher, an ATT0 quencher, and a QSY dye. In some embodiments, the detector ssDNA comprises a second quencher, such as an internal quencher like ZEN™, TAO™, and Black Hole Quencher®, which can lower background and increase signal detection.

In other embodiments, the detector ssDNA comprises a detectable label that provides a detectable signal before cleavage of the detector ssDNA and cleavage of the ssDNA inhibits or prevents detection of the signal. A non-limiting example of such a scenario is a detector ssDNA that comprises a fluorescence resonance energy transfer (FRET) pair. FRET is a process by which radiationless transfer of energy occurs from an excited state of a first (donor) fluorophore to a second (acceptor) fluorophore in close proximity. The emission spectrum of the donor fluorophore overlaps with the excitation spectrum of the acceptor fluorophore. Thus, the acceptor fluorophore will fluoresce when the detector ssDNA is whole (i.e., uncleaved) and the acceptor fluorophore will no longer fluoresce when the detector ssDNA is cleaved because the donor and acceptor fluorophore will no longer be in close proximity to one another. FRET donor and acceptor fluorophores are known in the art and include, but are not limited to cyan fluorescent protein (CFP)/green fluorescent protein (GFP), Cy3/Cy5, and GFP/yellow fluorescent protein (YFP).

In some embodiments, the detector ssDNA has a length of from about 2 nucleotides to about 30 nucleotides, including but not limited to about 2, about 3, about 4, about 5, 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 nucleotides, about 26 nucleotides, about 27 nucleotides, about 28 nucleotides, about 29 nucleotides, and about 30 nucleotides.

The sample in which a target DNA can be detected using these compositions and methods comprising a detector ssDNA include any sample comprising or believed to comprise a nucleic acid (e.g., DNA or RNA molecule). The sample can be derived from any source including a synthetic combination of purified nucleic acids or a biological sample such as respiratory swab (e.g., nasopharyngeal swab) extracts, a cell lysate, a patient sample, cells, tissues, saliva, blood, serum, plasma, urine, aspirate, biopsy samples, cerebral spinal fluid, or organism (e.g., bacteria, virus).

The contacting of the sample with the RGN, guide RNA, and detector ssDNA can include contacting *in vitro*, *ex vivo*, or *in vivo*. In some embodiments, the detector ssDNA and/or the RGN and/or guide RNA is immobilized on for example, a lateral flow device, wherein the sample contacts the

immobilized detector ssDNA and/or RGN and/or guide RNA. In some embodiments, antibodies against antigen moieties on the detector ssDNA are immobilized on, for example, a lateral flow device in a manner that allows differentiation of cleaved detector ssDNA from intact detector ssDNA. Also provided are devices (e.g., lateral flow, microfluidic), such as those described in International Publ. No. WO 2020/028729, which is herein incorporated by reference in its entirety, that comprise an immobilized detector ssDNA. The RGN and guide RNA can be added to a sample prior to, simultaneous with, or after the addition of the sample to the device and when the target DNA is present within the sample, the RGN will cleave the target DNA as well as the detector ssDNA, leading to the increase or reduction of a detectable signal that can be measured to detect the presence of the target DNA sequence. Alternatively, the RGN and/or guide RNA is immobilized on the device (e.g., lateral flow, microfluidic) and the sample and the detector ssDNA are added to the device. The detector ssDNA can be added to the sample before, during or after addition of the sample to the device. Another alternative device (e.g., lateral flow, microfluidic) comprises an immobilized antibody (or antibodies) against antigen moieties on the detector ssDNA and the sample, detector ssDNA, RGN and guide RNA are added to the device. In some embodiments, the methods can further comprise determining the amount of the target DNA present in the sample. The measurement of the detectable signal in the test sample can be compared to a reference measurement (e.g., a measurement of a reference sample or series thereof comprising a known amount of target DNA).

Non-limiting examples of applications of the compositions and methods include single-nucleotide polymorphism (SNP) detection, cancer screening, detection of a bacterial infection, detection of antibiotic resistance, and detection of a viral infection.

The detectable signal produced by cleavage of the ssDNA by the RGN can be measured using any suitable method known in the art including but not limited to measuring fluorescent signal, a visual analysis of bands on a gel, a colorimetric change, and the presence or absence of an electrical signal.

The present invention provides kits for detecting a target DNA of a DNA molecule in a sample, wherein the kit comprises an RGN polypeptide of the invention (or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide), a guide RNA (or a polynucleotide comprising a nucleotide sequence encoding the guide RNA) capable of hybridizing with the RGN and a target DNA sequence in a DNA molecule, and a detector ssDNA that does not hybridize with the guide RNA. In those embodiments wherein the target to be detected is an RNA, the kit can further comprise a reverse transcriptase. In those embodiments wherein nucleic acid amplification is used, the kit comprising the RGN and guide RNA (or polynucleotides encoding the same), and detector ssDNA can further comprise nucleic acid amplification reagents (e.g., DNA polymerase, nucleotides, buffer). In those embodiments wherein the kit comprises a polynucleotide encoding an RGN polypeptide and/or a polynucleotide encoding the guide RNA, the polynucleotides are introduced into a cell in which they are then expressed. In some of these embodiments at least one of the polynucleotides further comprises a promoter that is operably linked to the nucleotide sequence encoding the RGN polypeptide and/or guide RNA. In certain embodiments, the kit comprises more than one guide RNA (or polynucleotide(s) encoding more than one guide RNA) each capable of

hybridizing with the RGN. The guide RNAs can bind to unique target sequences of a single DNA molecule in order to amplify the detectable signal and lead to the detection of that DNA molecule.

Elements of the kit may be provided individually or in combinations, and may be provided in any suitable container, such as a vial, a bottle, or a tube. In some embodiments, the kit includes instructions in one or more languages. In some embodiments, a kit comprises one or more reagents for use in a process utilizing one or more of the elements described herein. Reagents may be provided in any suitable container. For example, a kit may provide one or more reaction or storage buffers. Reagents may be provided in a form that is usable in a particular assay, or in a form that requires addition of one or more other components before use (e.g. in concentrate or lyophilized form). A buffer can be any buffer, including but not limited to a sodium carbonate buffer, a sodium bicarbonate buffer, a borate buffer, a Tris buffer, a MOPS buffer, a HEPES buffer, and combinations thereof. In some embodiments, the buffer is alkaline. In some embodiments, the buffer has a pH from about 7 to about 10.

Also provided herein are methods of cleaving single-stranded DNAs by contacting a population of nucleic acids, wherein the population comprises a target DNA sequence of a DNA molecule and a plurality of non-target ssDNAs with an RGN and a guide RNA capable of hybridizing with the RGN and the target DNA sequence. In some of these embodiments, the population of nucleic acids are within a cell lysate. In some of these embodiments, the non-target ssDNAs are foreign to the cell and in some of these embodiments, the non-target ssDNAs are viral DNAs. In particular embodiments, the target DNA sequence is a viral sequence. The method can be performed *in vitro*, *in vivo*, or *ex vivo*. For example, the method could be performed *in vivo* wherein a subject is administered an RGN polypeptide and a guide RNA or one or more polynucleotides comprising a nucleotide sequence that encodes the RGN polypeptide and/or the guide RNA and the binding and cleavage of a viral target DNA sequence by the RGN can result in the cleavage of non-target viral ssDNAs within the infected cell.

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 herein 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 embodiments.

Non-limiting embodiments include:

1. A nucleic acid molecule comprising a polynucleotide encoding an RNA-guided nuclease (RGN) polypeptide, wherein said polynucleotide comprises a nucleotide sequence encoding an RGN

polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109;

wherein said RGN polypeptide is capable of binding a target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to said target DNA sequence, and

wherein said polynucleotide encoding an RGN polypeptide is operably linked to a promoter heterologous to said polynucleotide.

2. The nucleic acid molecule of embodiment 1, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109.

3. The nucleic acid molecule of embodiment 1, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109.

4. The nucleic acid molecule of any one of embodiments 1-3, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

5. The nucleic acid molecule of embodiment 4, wherein said RGN polypeptide is capable of cleaving said target DNA sequence upon binding.

6. The nucleic acid molecule of any one of embodiments 1-3, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

7. The nucleic acid molecule of embodiment 6, wherein said RGN polypeptide is capable of cleaving said target DNA sequence upon binding.

8. The nucleic acid molecule of embodiment 7, wherein cleavage by said RGN polypeptide generates a double-stranded break.

9. The nucleic acid molecule of embodiment 7, wherein cleavage by said RGN polypeptide generates a single-stranded break.

10. The nucleic acid molecule of any one of embodiments 1-9, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

11. The nucleic acid molecule of embodiment 10, wherein the base-editing polypeptide is a deaminase.

12. The nucleic acid molecule of any one of embodiments 1-11, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

13. The nucleic acid molecule of any one of embodiments 1-12, wherein the RGN polypeptide comprises one or more nuclear localization signals.

14. The nucleic acid molecule of any one of embodiments 1-13, wherein the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

15. A vector comprising the nucleic acid molecule of any one of embodiments 1-14.

16. The vector of embodiment 15, further comprising at least one nucleotide sequence encoding an RGN accessory protein selected from the group consisting of:

a) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

17. The vector of embodiment 16, wherein said RGN accessory protein is selected from the group consisting of:

a) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

18. The vector of embodiment 16, wherein said RGN accessory protein is selected from the group consisting of:

a) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

19. The vector of any one of embodiments 15-18, further comprising at least one nucleotide sequence encoding said gRNA capable of hybridizing to said target DNA sequence.

20. The vector of embodiment 19, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said gRNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

21. The vector of embodiment 19, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said gRNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

22. The vector of embodiment 19, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said gRNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

23. The vector of embodiment 19, wherein said gRNA comprises a tracrRNA.

24. The vector of embodiment 23, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence

identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

25. The vector of embodiment 23, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

26. The vector of embodiment 23, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ

ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

27. The vector of any one of embodiments 23-26, where said gRNA is a single guide RNA.

28. The vector of any one of embodiments 23-26, wherein said gRNA is a dual-guide RNA.

29. A cell comprising the nucleic acid molecule of any one of embodiments 1-14 or the vector of any one of embodiments 15-28.

30. A method for making an RGN polypeptide comprising culturing the cell of embodiment 29 under conditions in which the RGN polypeptide is expressed.

31. A method for making an RGN polypeptide comprising introducing into a cell a heterologous nucleic acid molecule comprising a nucleotide sequence encoding an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109;

wherein said RGN polypeptide binds a target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to said target DNA sequence;

and culturing said cell under conditions in which the RGN polypeptide is expressed.

32. The method of embodiment 31, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109.

33. The method of embodiment 31, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109.

34. The method of any one of embodiments 30-33, further comprising purifying said RGN polypeptide.

35. The method of any one of embodiments 30-33, wherein said cell further expresses one or more guide RNAs that binds to said RGN polypeptide to form an RGN ribonucleoprotein complex.

36. The method of embodiment 35, further comprising purifying said RGN ribonucleoprotein complex.

37. An isolated RNA-guided nuclease (RGN) polypeptide, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109; and

wherein said RGN polypeptide is capable of binding a target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to said target DNA sequence.

38. The isolated RGN polypeptide of embodiment 37, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109.

39. The isolated RGN polypeptide of embodiment 37, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109.

40. The isolated RGN polypeptide of any one of embodiments 37-39, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

41. The isolated RGN polypeptide of embodiment 40, wherein said RGN polypeptide is capable of cleaving said target DNA sequence upon binding.

42. The isolated RGN polypeptide of any one of embodiments 37-39, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

43. The isolated RGN polypeptide of embodiment 42, wherein said RGN polypeptide is capable of cleaving said target DNA sequence upon binding.

44. The isolated RGN polypeptide of embodiment 43, wherein cleavage by said RGN polypeptide generates a double-stranded break.

45. The isolated RGN polypeptide of embodiment 43, wherein cleavage by said RGN polypeptide generates a single-stranded break.

46. The isolated RGN polypeptide of any one of embodiments 37-45, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

47. The isolated RGN polypeptide of embodiment 46, wherein the base-editing polypeptide is a deaminase.

48. The isolated RGN polypeptide of any one of embodiments 37-47, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

49. The isolated RGN polypeptide of any one of embodiments 37-48, wherein the RGN polypeptide comprises one or more nuclear localization signals.

50. A nucleic acid molecule comprising a polynucleotide encoding a CRISPR RNA (crRNA), wherein said crRNA comprises a spacer sequence and a CRISPR repeat sequence, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 110 to 119;

wherein a guide RNA comprising:

a) said crRNA; or

b) said crRNA and a trans-activating CRISPR RNA (tracrRNA) capable of hybridizing to said CRISPR repeat sequence of said crRNA;

is capable of hybridizing to a target DNA sequence of a DNA molecule in a sequence specific manner through the spacer sequence of said crRNA when said guide RNA is bound to an RNA-guided nuclease (RGN) polypeptide, and

wherein said polynucleotide encoding a crRNA is operably linked to a promoter heterologous to said polynucleotide.

51. The nucleic acid molecule of embodiment 50, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 95% sequence identity to any one of SEQ ID NOs: 110 to 119.

52. The nucleic acid molecule of embodiment 50, wherein said CRISPR repeat sequence comprises a nucleotide sequence having 100% sequence identity to any one of SEQ ID NOs: 110 to 119.

53. A vector comprising the nucleic acid molecule of any one of embodiments 50-52.

54. The vector of embodiment 53, wherein said vector further comprises a polynucleotide encoding said tracrRNA.

55. The vector of embodiment 54, wherein said CRISPR repeat sequence has at least 90% sequence identity to SEQ ID NO: 110, and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 120.

56. The vector of embodiment 54, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 110, and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 120.

57. The vector of embodiment 54, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 110, and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 120.

58. The vector of any one of embodiments 55-57, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1.

59. The vector of embodiment 58, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1.

60. The vector of embodiment 58, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1.

61. The vector of embodiment 54, wherein said CRISPR repeat sequence has at least 90% sequence identity to SEQ ID NO: 111, and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 121.

62. The vector of embodiment 54, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 111, and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 121.

63. The vector of embodiment 54, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 111, and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 121.

64. The vector of any one of embodiments 61-63, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2.

65. The vector of embodiment 64, wherein RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2.

66. The vector of embodiment 64, wherein RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2.

67. The vector of embodiment 54, wherein said CRISPR repeat sequence has at least 90% sequence identity to SEQ ID NO: 112, and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 122.

68. The vector of embodiment 54, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 112, and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 122.

69. The vector of embodiment 54, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 112, and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 122.

70. The vector of any one of embodiments 67-69, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3.

71. The vector of embodiment 70, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3.

72. The vector of embodiment 70, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3.

73. The vector of embodiment 54, wherein said CRISPR repeat sequence has at least 90% sequence identity to SEQ ID NO: 113, and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 123.

74. The vector of embodiment 54, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 113, and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 123.

75. The vector of embodiment 54, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 113, and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 123.

76. The vector of any one of embodiments 73-75, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4.

77. The vector of embodiment 76, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4.

78. The vector of embodiment 76, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4.

79. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 114, and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 124.

80. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 114, and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 124.

81. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 114, and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 124.

82. The vector of any one of embodiments 79-81, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5.

83. The vector of embodiment 82, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5.

84. The vector of embodiment 82, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5.

85. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 115, and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 125.

86. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 115, and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 125.

87. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 115, and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 125.

88. The vector of any one of embodiments 85-87, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6.

89. The vector of embodiment 88, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6.

90. The vector of embodiment 88, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6.

91. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 117, and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 126.

92. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 117, and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 126.

93. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 117, and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 126.

94. The vector of any one of embodiments 91-93, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12.

95. The vector of embodiment 94, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12.

96. The vector of embodiment 94, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12.

97. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 118, and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 127.

98. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 118, and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 127.

99. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 118, and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 127.

100. The vector of any one of embodiments 97-99, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13.

101. The vector of embodiment 100, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13.

102. The vector of embodiment 100, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13.

103. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 119, and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 128.

104. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 119, and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 128.

105. The vector of embodiment 54, wherein said CRISPR repeat sequence comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 119, and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 128.

106. The vector of any one of embodiments 103-105, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

107. The vector of embodiment 106, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

108. The vector of embodiment 106, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

109. The vector of any one of embodiments 54-109, wherein said polynucleotide encoding said crRNA and said polynucleotide encoding said tracrRNA are operably linked to the same promoter and are encoded as a single guide RNA.

110. The vector of any one of embodiments 54-109, wherein said polynucleotide encoding said crRNA and said polynucleotide encoding said tracrRNA are operably linked to separate promoters.

111. A nucleic acid molecule comprising a polynucleotide encoding a trans-activating CRISPR RNA (tracrRNA) comprising a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 120 to 128;

wherein a guide RNA comprising:

a) said tracrRNA; and

b) a crRNA comprising a spacer sequence and a CRISPR repeat sequence, wherein said tracrRNA is capable of hybridizing with said CRISPR repeat sequence of said crRNA;

is capable of hybridizing to a target DNA sequence in a sequence specific manner through the spacer sequence of said crRNA when said guide RNA is bound to an RNA-guided nuclease (RGN) polypeptide, and

wherein said polynucleotide encoding a tracrRNA is operably linked to a promoter heterologous to said polynucleotide.

112. The nucleic acid molecule of embodiment 111, wherein said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to any one of SEQ ID NOs: 120 to 128.

113. The nucleic acid molecule of embodiment 111, wherein said tracrRNA comprises a nucleotide sequence having 1000% sequence identity to any one of SEQ ID NOs: 120 to 128.

114. A vector comprising the nucleic acid molecule of any one of embodiments 111-113.

115. The vector of embodiment 114, wherein said vector further comprises a polynucleotide encoding said crRNA.

116. The vector of embodiment 115, wherein said crRNA comprises a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110 and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 120.

117. The vector of embodiment 116, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 110 and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 120.

118. The vector of embodiment 116, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 110 and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 120.

119. The vector of any one of embodiments 116-118, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1.

120. The vector of embodiment 119, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1.

121. The vector of embodiment 119, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1.

122. The vector of embodiment 115, wherein said crRNA comprises a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111 and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 121.

123. The vector of embodiment 122, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 111 and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 121.

124. The vector of embodiment 122, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 111 and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 121.

125. The vector of any one of embodiments 122-124, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2.

126. The vector of embodiment 125, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2.

127. The vector of embodiment 125, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2.

128. The vector of embodiment 115, wherein said crRNA comprises a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112 and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 122.

129. The vector of embodiment 128, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 112 and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 122.

130. The vector of embodiment 128, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 112 and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 122.

131. The vector of any one of embodiments 128-130, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3.

132. The vector of embodiment 131, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3.

133. The vector of embodiment 131, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3.

134. The vector of embodiment 115, wherein said crRNA comprises a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113 and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 123.

135. The vector of embodiment 134, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 113 and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 123.

136. The vector of embodiment 134, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 113 and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 123.

137. The vector of any one of embodiments 134-136, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4.

138. The vector of embodiment 137, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4.

139. The vector of embodiment 137, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4.

140. The vector of embodiment 115, wherein said crRNA comprises a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114 and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 124.

141. The vector of embodiment 140, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 114 and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 124.

142. The vector of embodiment 140, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 114 and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 124.

143. The vector of any one of embodiments 140-142, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5.

144. The vector of embodiment 143, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5.

145. The vector of embodiment 143, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5.

146. The vector of embodiment 115, wherein said crRNA comprises a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115 and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 125.

147. The vector of embodiment 146, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 115 and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 125.

148. The vector of embodiment 146, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 115 and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 125.

149. The vector of any one of embodiments 146-148, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6.

150. The vector of embodiment 149, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6.

151. The vector of embodiment 149, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6.

152. The vector of embodiment 115, wherein said crRNA comprises a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117 and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 126.

153. The vector of embodiment 152, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 117 and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 126.

154. The vector of embodiment 152, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 117 and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 126.

155. The vector of any one of embodiments embodiment 152-154, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

156. The vector of embodiment 155, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

157. The vector of embodiment 155, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

158. The vector of embodiment 115, wherein said crRNA comprises a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118 and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 127.

159. The vector of embodiment 158, wherein said CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118 and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 127.

160. The vector of embodiment 158, wherein said CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118 and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 127.

161. The vector of any one of embodiments 158-160, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13.

162. The vector of embodiment 161, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13.

163. The vector of embodiment 161, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13.

164. The vector of embodiment 115, wherein said crRNA comprises a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119 and said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 128.

165. The vector of embodiment 164, wherein said CRISPR repeat sequence has at least 95% sequence identity to SEQ ID NO: 119 and said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to SEQ ID NO: 128.

166. The vector of embodiment 164, wherein said CRISPR repeat sequence has 100% sequence identity to SEQ ID NO: 119 and said tracrRNA comprises a nucleotide sequence having 100% sequence identity to SEQ ID NO: 128.

167. The vector of any one of embodiments 164-166, wherein said vector further comprises a polynucleotide encoding said RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

168. The vector of embodiment 167, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

169. The vector of embodiment 167, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

170. The vector of any one of embodiments 115-169, wherein said polynucleotide encoding said crRNA and said polynucleotide encoding said tracrRNA are operably linked to the same promoter and are encoded as a single guide RNA.

171. The vector of any one of embodiments 115-169, wherein said polynucleotide encoding said crRNA and said polynucleotide encoding said tracrRNA are operably linked to separate promoters.

172. A system for binding a target DNA sequence of a DNA molecule, said system comprising:

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

b) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109 or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide;

wherein at least one of said nucleotide sequence encoding the one or more guide RNAs and said nucleotide sequence encoding the RGN polypeptide is 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 RGN polypeptide in order to direct said RGN polypeptide to bind to said target DNA sequence of the DNA molecule.

173. A system for binding a target DNA sequence of a DNA molecule, said system comprising:

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

b) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109;

wherein the one or more guide RNAs are capable of hybridizing to the target DNA sequence, and

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

174. The system of embodiment 173, wherein at least one of said nucleotides sequences encoding the one or more guide RNAs is operably linked to a promoter heterologous to said nucleotide sequence.

175. The system of any one of embodiments 172-174, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109.

176. The system of any one of embodiments 172-174, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109.

177. The system of any one of embodiments 172-176, wherein said RGN polypeptide and said one or more guide RNAs are not found complexed to one another in nature.

178. The system of any one of embodiments 172-177, wherein said target DNA sequence is a eukaryotic target DNA sequence.

179. The system of any one of embodiments 172-178, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

180. The system of any one of embodiments 172-178, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said one or more guide

RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

181. The system of any one of embodiments 172-178, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

182. The system of any one of embodiments 172-181, wherein said one or more guide RNAs comprise a tracrRNA.

183. The system of embodiment 182, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least

90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

184. The system of embodiment 182, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

185. The system of embodiment 182, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100%

sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

186. The system of any one of embodiments 182-185, wherein said one or more guide RNAs are a single guide RNA (sgRNA).

187. The system of any one of embodiments 182-185, wherein said one or more guide RNAs are a dual-guide RNA.

188. The system of any one of embodiments 172-187, wherein said system further comprises at least one RGN accessory protein or a polynucleotide comprising a nucleotide sequence encoding the same selected from the group consisting of:

a) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

189. The system of embodiment 188, wherein said at least one RGN accessory protein is selected from the group consisting of:

a) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

190. The system of embodiment 188, wherein said at least one RGN accessory protein is selected from the group consisting of:

a) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

191. The system of any one of embodiments 172-190, wherein the target DNA sequence is within a cell.

192. The system of embodiment 191, wherein the cell is a eukaryotic cell.

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

194. The system of embodiment 192, wherein the eukaryotic cell is a mammalian cell.

195. The system of embodiment 194, wherein the mammalian cell is a human cell.

196. The system of embodiment 195, wherein the human cell is an immune cell.

197. The system of embodiment 196, wherein the human cell is a stem cell.

198. The system of embodiment 197, wherein the stem cell is an induced pluripotent stem cell.

199. The system of embodiment 192, wherein the eukaryotic cell is an insect cell.

200. The system of embodiment 191, wherein the cell is a prokaryotic cell.

201. The system of any one of embodiments 172-200, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

202. The system of embodiment 201, wherein when transcribed the one or more guide RNAs is capable of hybridizing to the target DNA sequence and the guide RNA is capable of forming a complex with the RGN polypeptide to direct cleavage of the target DNA sequence.

203. The system of any one of embodiments 172-200, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

204. The system of embodiment 203, wherein when transcribed the one or more guide RNAs is capable of hybridizing to the target DNA sequence and the guide RNA is capable of forming a complex with the RGN polypeptide to direct cleavage of the target DNA sequence.

205. The system of embodiment 204, wherein said RGN polypeptide is capable of generating a double-stranded break.

206. The system of embodiment 204, wherein said RGN polypeptide is capable of generating a single-stranded break.

207. The system of any one of embodiments 172-206, wherein the RGN polypeptide is operably linked to a base-editing polypeptide.

208. The system of embodiment 207, wherein said base-editing polypeptide is a deaminase.

209. The system of any one of embodiments 172-208, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

210. The system of any one of embodiments 172-209, wherein the RGN polypeptide comprises one or more nuclear localization signals.

211. The system of any one of embodiments 172-210, wherein the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

212. The system of any one of embodiments 172-211, wherein polynucleotides comprising the nucleotide sequences encoding the one or more guide RNAs and the polynucleotide comprising the nucleotide sequence encoding an RGN polypeptide are located on one vector.

213. The system of any one of embodiments 172-212, wherein said system further comprises one or more donor polynucleotides or one or more polynucleotides comprising one or more nucleotide sequences encoding the one or more donor polynucleotides.

214. A pharmaceutical composition comprising the nucleic acid molecule of any one of embodiments 1-14, 50-52, and 111-113, the vector of any one of embodiments 15-28, 53-110, and 114-171, the cell of embodiment 29, the isolated RGN polypeptide of any one of embodiments 37-49, or the system of any one of embodiments 172-213, and a pharmaceutically acceptable carrier.

215. A method for binding a target DNA sequence of a DNA molecule comprising delivering a system according to any one of embodiments 172-213, to said target DNA sequence or a cell comprising the target DNA sequence.

216. The method of embodiment 215, wherein said RGN polypeptide or said guide RNA further comprises a detectable label, thereby allowing for detection of said target DNA sequence.

217. The method of embodiment 215, wherein said guide RNA or said RGN polypeptide further comprises an expression modulator, thereby modulating expression of said target DNA sequence or a gene under transcriptional control by said target DNA sequence.

218. A method for cleaving a target DNA sequence of a DNA molecule comprising delivering a system according to any one of embodiments 172-213, to said target DNA sequence or a cell comprising the target DNA sequence.

219. The method of embodiment 218, wherein said modified target DNA sequence comprises insertion of heterologous DNA into the target DNA sequence.

220. The method of embodiment 218, wherein said modified target DNA sequence comprises deletion of at least one nucleotide from the target DNA sequence.

221. The method of embodiment 218, wherein said modified target DNA sequence comprises mutation of at least one nucleotide in the target DNA sequence.

222. A method for binding a target DNA sequence of a DNA molecule, said method comprising:

a) assembling an RNA-guided nuclease (RGN) ribonucleotide complex *in vitro* by combining:

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

ii) an RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109;

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

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

wherein the one or more guide RNAs hybridize to the target DNA sequence, thereby directing said RGN polypeptide to bind to said target DNA sequence.

223. The method of embodiment 222, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

224. The method of embodiment 222, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

225. The method of any one of embodiments 222-224, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

226. The method of any one of embodiments 222-225, wherein said RGN polypeptide or said guide RNA further comprises a detectable label, thereby allowing for detection of said target DNA sequence.

227. The method of any one of embodiments 222-225, wherein said guide RNA or said RGN polypeptide further comprises an expression modulator, thereby allowing for the modulation of expression of said target DNA sequence.

228. A method for cleaving and/or modifying a target DNA sequence of a DNA molecule, comprising contacting the DNA molecule with:

a) an RNA-guided nuclease (RGN) polypeptide, wherein said RGN comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109; and

b) one or more guide RNAs capable of targeting the RGN of (a) to the target DNA sequence;

wherein the one or more guide RNAs hybridize to the target DNA sequence, thereby directing said RGN polypeptide to bind to said target DNA sequence and cleavage and/or modification of said target DNA sequence occurs.

229. The method of embodiment 228, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

230. The method of embodiment 228, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

231. The method of embodiment 230, wherein cleavage by said RGN polypeptide generates a double-stranded break.

232. The method of embodiment 230, wherein cleavage by said RGN polypeptide generates a single-stranded break.

233. The method of any one of embodiments 228-232, wherein said RGN polypeptide is operably linked to a base-editing polypeptide.

234. The method of embodiment 233, wherein said base-editing polypeptide comprises a deaminase.

235. The method of any one of embodiments 228-234, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

236. The method of any one of embodiments 228-235, wherein said modified target DNA sequence comprises insertion of heterologous DNA into the target DNA sequence.

237. The method of any one of embodiments 228-235, wherein said modified target DNA sequence comprises deletion of at least one nucleotide from the target DNA sequence.

238. The method of any one of embodiments 228-235, wherein said modified target DNA sequence comprises mutation of at least one nucleotide in the target DNA sequence.

239. The method of any one of embodiments 222-238, wherein said RGN comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109.

240. The method of any one of embodiments 222-238, wherein said RGN comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109.

241. The method of any one of embodiments 222-240, wherein said RGN polypeptide and said one or more guide RNAs are not found complexed to one another in nature.

242. The method of any one of embodiments 222-241, wherein said target DNA sequence is a eukaryotic target DNA sequence.

243. The method of any one of embodiments 222-242, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

244. The method of any one of embodiments 222-242, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

245. The method of any one of embodiments 222-242, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

246. The method of any one of embodiments 222-242, wherein said one or more guide RNAs comprise a tracrRNA.

247. The method of embodiment 246, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least

90% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

248. The method of embodiment 246, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

249. The method of embodiment 246, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100%

sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

250. The method of any one of embodiments 246-249, wherein said one or more guide RNAs are a single guide RNA (sgRNA).

251. The method of any one of embodiments 246-249, wherein said one or more guide RNAs are a dual-guide RNA.

252. The method of any one of embodiments 222-251, wherein said method further comprises contacting the DNA molecule with one or more RGN accessory proteins selected from the group consisting of:

a) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

253. The method of embodiment 252, wherein said one or more RGN accessory proteins are selected from the group consisting of:

a) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

254. The method of embodiment 252, wherein said one or more RGN accessory proteins are selected from the group consisting of:

a) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

255. The method of any one of embodiments 215-254, wherein the target DNA sequence is within a cell.

256. The method of embodiment 255, wherein the cell is a eukaryotic cell.

257. The method of embodiment 256, wherein the eukaryotic cell is a plant cell.

258. The method of embodiment 256, wherein the eukaryotic cell is a mammalian cell.

259. The method of embodiment 258, wherein the mammalian cell is a human cell.

260. The method of embodiment 259, wherein the human cell is an immune cell.

261. The method of embodiment 260, wherein the human cell is a stem cell.

262. The method of embodiment 261, wherein the stem cell is an induced pluripotent stem cell.

263. The method of embodiment 256, wherein the eukaryotic cell is an insect cell.

264. The method of embodiment 255, wherein the cell is a prokaryotic cell.

265. A cell comprising a modified target DNA sequence according to the method of any one of embodiments 228-254.

266. The cell of embodiment 265, wherein the cell is a eukaryotic cell.

267. The cell of embodiment 266, wherein the eukaryotic cell is a plant cell.

268. A plant comprising the cell of embodiment 267.

269. A seed comprising the cell of embodiment 267.

270. The cell of embodiment 266, wherein the eukaryotic cell is a mammalian cell.

271. The cell of embodiment 270, wherein the mammalian cell is a human cell.

272. The cell of embodiment 271, wherein the human cell is an immune cell.

273. The cell of embodiment 272, wherein the human cell is a stem cell.

274. The cell of embodiment 273, wherein the stem cell is an induced pluripotent stem cell.

275. The cell of embodiment 266, wherein the eukaryotic cell is an insect cell.

276. The cell of embodiment 265, wherein the cell is a prokaryotic cell.

277. A pharmaceutical composition comprising the cell of any one of embodiments 266 and 270-274, and a pharmaceutically acceptable carrier.

278. A kit for detecting a target DNA sequence of a DNA molecule in a sample, the kit comprising:

a) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109 or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide, wherein said RGN polypeptide is capable of binding and cleaving said target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA capable of hybridizing to said target DNA sequence;

b) said guide RNA or a polynucleotide comprising a nucleotide sequence encoding said guide RNA; and

c) a detector single-stranded DNA (ssDNA) that does not hybridize with the guide RNA.

279. The kit of embodiment 278, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109.

280. The kit of embodiment 278, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109.

281. The kit of any one of embodiments 278-280, wherein at least one of said nucleotide sequence encoding the guide RNA and said nucleotide sequence encoding the RGN polypeptide is operably linked to a promoter heterologous to said nucleotide sequence.

282. The kit of any one of embodiments 278-281, wherein said RGN polypeptide and said one or more guide RNAs are not found complexed to one another in nature.

283. The kit of any one of embodiments 278-282, wherein said target DNA sequence is a eukaryotic target DNA sequence.

284. The kit of any one of embodiments 278-283, wherein said detector ssDNA comprises a fluorophore/quencher pair.

285. The kit of any one of embodiments 278-283, wherein said detector ssDNA comprises a fluorescence resonance energy transfer (FRET) pair.

286. The kit of any one of embodiments 278-285, wherein said kit further comprises at least one RGN accessory protein selected from the group consisting of:

a) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

287. The kit of embodiment 286, wherein said at least one RGN accessory protein is selected from the group consisting of:

a) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

288. The kit of embodiment 286, wherein said at least one RGN accessory protein is selected from the group consisting of:

a) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

289. The kit of any one of embodiments 278-288, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

290. The kit of any one of embodiments 278-288, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

291. The kit of any one of embodiments 278-288, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

292. The kit of any one of embodiments 278-288, wherein said one or more guide RNAs comprise a tracrRNA.

293. The kit of embodiment 292, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least

90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

294. The kit of embodiment 292, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least

95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

295. The kit of embodiment 292, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100%

sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

296. The kit of any one of embodiments 292-295, wherein said one or more guide RNAs are a single guide RNA (sgRNA).

297. The kit of any one of embodiments 292-295, wherein said one or more guide RNAs are a dual-guide RNA.

298. The kit of any one of embodiments 278-297, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

299. The kit of any one of embodiments 278-297, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

300. The kit of embodiment 299, wherein cleavage by said RGN polypeptide generates a double-stranded break.

301. The kit of embodiment 299, wherein cleavage by said RGN polypeptide generates a single-stranded break.

302. The kit of any one of embodiments 278-301, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

303. A method of detecting a target DNA sequence of a DNA molecule in a sample, the method comprising:

a) contacting the sample with:

i) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109, wherein said RGN polypeptide is capable of binding and cleaving said target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA capable of hybridizing to said target DNA sequence;

ii) said guide RNA; and

iii) a detector single-stranded DNA (ssDNA) that does not hybridize with the guide RNA; and

b) measuring a detectable signal produced by cleavage of the detector ssDNA by the RGN, thereby detecting the target DNA sequence.

304. The method of embodiment 303, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109.

305. The method of embodiment 303, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109.

306. The method of any one of embodiments 303-305, wherein said sample comprises DNA molecules from a cell lysate.

307. The method of any one of embodiments 303-305, wherein said sample comprises cells.

308. The method of embodiment 307, wherein said cells are eukaryotic cells.

309. The method of any one of embodiments 303-305, wherein the DNA molecule comprising the target DNA sequence is produced by reverse-transcription of an RNA template molecule present in a sample comprising RNA.

310. The method of embodiment 309, wherein the RNA template molecule is an RNA virus.

311. The method of embodiment 310, wherein the RNA virus is a coronavirus.

312. The method of embodiment 311, wherein the coronavirus is a bat SARS-like coronavirus, SARS-CoV, or SARS-CoV-2.

313. The method of any one of embodiments 309-312, wherein the samples comprising RNA is derived from a sample comprising cells.

314. The method of any one of embodiments 303-313, wherein said detector ssDNA comprises a fluorophore/quencher pair.

315. The method of any one of embodiments 303-313, wherein said detector ssDNA comprises a fluorescence resonance energy transfer (FRET) pair.

316. The method of any one of embodiments 303-315, wherein said method further comprises amplifying nucleic acids in the sample prior to or together with the contacting of step a).

317. The method of any one of embodiments 303-316, wherein said method further comprises contacting the sample with one or more RGN accessory proteins selected from the group consisting of:

a) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

318. The method of embodiment 317, wherein said one or more RGN accessory proteins is selected from the group consisting of:

a) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

319. The method of embodiment 317, wherein said one or more RGN accessory proteins is selected from the group consisting of:

a) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

320. A method of cleaving single-stranded DNAs (ssDNAs), the method comprising contacting a population of nucleic acids, wherein said population comprises a DNA molecule comprising a target DNA sequence and a plurality of non-target ssDNAs with:

a) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109, wherein said RGN polypeptide is capable of binding and cleaving said target DNA sequence in an RNA-guided sequence specific manner when bound to a guide RNA capable of hybridizing to said target DNA sequence; and

b) said guide RNA;

wherein the RGN polypeptide cleaves non-target ssDNAs of said plurality.

321. The method of embodiment 320, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109.

322. The method of embodiment 320, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109.

323. The method of any one of embodiments 320-322, wherein said population of nucleic acids are within a cell lysate.

324. The method of any one of embodiments 320-323, wherein the DNA molecule comprising the target DNA sequence is produced by reverse-transcription of an RNA template molecule.

325. The method of any one of embodiments 320-324, wherein said method further comprises contacting the population with one or more RGN accessory proteins selected from the group consisting of:

a) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 90% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 90% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

326. The method of embodiment 325, wherein said one or more RGN accessory proteins is selected from the group consisting of:

a) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having at least 95% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having at least 95% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

327. The method of embodiment 325, wherein said one or more RGN accessory proteins is selected from the group consisting of:

a) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 178-181, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11;

b) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 182-184, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

c) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 185-187, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13;

d) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 191, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 14;

e) an RGN accessory protein having 100% sequence identity to SEQ ID NO: 192, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 15; and

f) at least one RGN accessory protein having 100% sequence identity to any one of SEQ ID NOs: 188-190, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

328. The method of any one of embodiments 303-327, wherein said RGN polypeptide and said guide RNA are not found complexed to one another in nature.

329. The method of any one of embodiments 303-328, wherein said target DNA sequence is a eukaryotic target DNA sequence.

330. The method of any one of embodiments 303-329, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said guide RNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

331. The method of any one of embodiments 303-329, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said guide RNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

332. The method of any one of embodiments 303-329, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said guide RNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

333. The method of any one of embodiments 303-329, wherein said guide RNA comprises a tracrRNA.

334. The method of embodiment 333, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90%

sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

335. The method of embodiment 333, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

336. The method of embodiment 333, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to

SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

337. The method of any one of embodiments 333-336, wherein said guide RNA is a single guide RNA (sgRNA).

338. The method of any one of embodiments 333-336, wherein said guide RNA is a dual-guide RNA.

339. The method of any one of embodiments 303-338, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

340. The method of any one of embodiments 303-338, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

341. The method of embodiment 340, wherein cleavage of said target DNA sequence by said RGN polypeptide generates a double-stranded break.

342. The method of embodiment 340, wherein cleavage of said target DNA sequence by said RGN polypeptide generates a single-stranded break.

343. The method of any one of embodiments 303-342, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

344. 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) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109, or a polynucleotide encoding said RGN polypeptide, wherein said polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and

b) a guide RNA (gRNA), wherein the gRNA comprises a CRISPR repeat sequence comprising a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 110 to 119, 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 RGN and gRNA target to the genomic location of the causal mutation and modify the genomic sequence to remove the causal mutation.

345. The method of embodiment 344, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109 and said CRISPR repeat sequence comprises a nucleotide sequence having at least 95% sequence identity to any one of SEQ ID NOs: 110 to 119.

346. The method of embodiment 344, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109 and said CRISPR repeat sequence comprises a nucleotide sequence having 100% sequence identity to any one of SEQ ID NOs: 110 to 119.

347. The method of any one of embodiments 344-346, wherein the RGN is operably linked to a base-editing polypeptide.

348. The method of embodiment 347, wherein the base-editing polypeptide is a deaminase.

349. The method of any one of embodiments 344-348, wherein the cell is an animal cell.

350. The method of any one of embodiments 344-348, wherein the cell is a mammalian cell.

351. The method of embodiment 349, wherein the cell is derived from a dog, cat, mouse, rat, rabbit, horse, cow, pig, or human.

352. The method of embodiment 349, wherein the genetically inherited disease is caused by a single nucleotide polymorphism.

353. The method of embodiment 351, wherein the genetically inherited disease is Hurler syndrome.

354. The method of embodiment 353, wherein the gRNA further comprises a spacer sequence that targets a region proximal to the causal single nucleotide polymorphism.

355. A method for producing a genetically modified cell with a deletion in a disease-causing genomic region of instability, the method comprising introducing into the cell:

a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1 to 109, or a polynucleotide encoding said RGN polypeptide, wherein said polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and

b) a first guide RNA (gRNA), wherein the first gRNA comprises a CRISPR repeat sequence comprising a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 110 to 119, or a polynucleotide encoding said first gRNA, wherein said polynucleotide encoding the first gRNA is operably linked to a promoter to enable expression of the first gRNA in the cell, and further wherein the first gRNA comprises a spacer sequence that targets the 5'flank of the genomic region of instability; and

c) a second guide RNA (gRNA), wherein the second gRNA comprises a CRISPR repeat sequence comprising a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 110 to 119, or a polynucleotide encoding said second gRNA, wherein said polynucleotide encoding the second gRNA is operably linked to a promoter to enable expression of the second gRNA in the cell, and further wherein said second gRNA comprises a spacer sequence that targets the 3'flank of the genomic region of instability;

whereby the RGN and the first and second gRNAs target to the genomic region of instability and at least a portion of the genomic region of instability is removed.

356. The method of embodiment 355, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1 to 109 and said CRISPR repeat sequence of said first gRNA and said second gRNA comprises a nucleotide sequence having at least 95% sequence identity to any one of SEQ ID NOs: 110 to 119.

357. The method of embodiment 355, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1 to 109 and said CRISPR repeat sequence of said first gRNA and said second gRNA comprises a nucleotide sequence having 100% sequence identity to any one of SEQ ID NOs: 110 to 119.

358. The method of any one of embodiments 355-357, wherein the cell is an animal cell.

359. The method of any one of embodiments 355-357, wherein the cell is a mammalian cell.

360. The method of embodiment 359, wherein the cell is derived from a dog, cat, mouse, rat, rabbit, horse, cow, pig, or human.

361. The method of embodiment 358, wherein the genetically inherited disease is Friedrich's Ataxia or Huntington's Disease.

362. The method of embodiment 358, wherein the spacer sequence of the first gRNA further targets a region within or proximal to the genomic region of instability.

363. The method of embodiment 362, wherein the spacer sequence of the second gRNA further targets a region within or proximal to the genomic region of instability.

364. A method for producing a genetically modified mammalian hematopoietic progenitor cell having decreased BCL11A mRNA and protein expression, the method comprising introducing into an isolated human hematopoietic progenitor cell:

a) an RNA-guided nuclease (RGN) polypeptide, wherein the RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-109, or a polynucleotide encoding said RGN polypeptide, wherein said polynucleotide encoding the RGN polypeptide is operably linked to a promoter to enable expression of the RGN polypeptide in the cell; and

b) a guide RNA (gRNA), wherein the gRNA comprises a CRISPR repeat sequence comprising a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 110-119, 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 RGN and gRNA are expressed in the cell and cleave at the BCL11A enhancer region, resulting in genetic modification of the human hematopoietic progenitor cell and reducing the mRNA and/or protein expression of BCL11A.

365. The method of embodiment 364, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-109 and said CRISPR repeat sequence comprises a nucleotide sequence having at least 95% sequence identity to any one of SEQ ID NOs: 110-119.

366. The method of embodiment 364, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-109 and said CRISPR repeat sequence comprises a nucleotide sequence having 100% sequence identity to any one of SEQ ID NOs: 110-119.

367. The method of any one of embodiments 364-366, wherein the gRNA further comprises a spacer sequence that targets a region within or proximal to the BCL11A enhancer region.

368. The method of any one of embodiments 344-367, wherein the guide RNA, first guide RNA, and second guide RNA comprises a tracrRNA.

369. The method of embodiment 368, wherein said tracrRNA comprises a nucleotide sequence having at least 90% sequence identity to any one of SEQ ID NOs: 120 to 128.

370. The method of embodiment 368, wherein said tracrRNA comprises a nucleotide sequence having at least 95% sequence identity to any one of SEQ ID NOs: 120 to 128.

371. The method of embodiment 368, wherein said tracrRNA comprises a nucleotide sequence having 100% sequence identity to any one of SEQ ID NOs: 120 to 128.

372. A method of treating a disease, said method comprising administering to a subject in need thereof an effective amount of a pharmaceutical composition of embodiment 214 or 277.

373. The method of embodiment 372, wherein said disease is associated with a causal mutation and said effective amount of said pharmaceutical composition corrects said causal mutation.

374. Use of the nucleic acid molecule of any one of embodiments 1-14, 50-52, and 111-113, the vector of any one of embodiments 15-28, 53-110, and 114-171, the cell of any one of embodiments 29, 266, and 270-274, the isolated RGN polypeptide of any one of embodiments 37-49, or the system of any one of embodiments 172-213 for the treatment of a disease in a subject.

375. The use of embodiment 374, wherein said disease is associated with a causal mutation and said treating comprises correcting said causal mutation.

376. Use of the nucleic acid molecule of any one of embodiments 1-14, 50-52, and 111-113, the vector of any one of embodiments 15-28, 53-110, and 114-171, the cell of any one of embodiments 29, 266, and 270-274, the isolated RGN polypeptide of any one of embodiments 37-49, or the system of any one of embodiments 172-213 for the manufacture of a medicament useful for treating a disease.

377. The use of embodiment 376, wherein said disease is associated with a causal mutation and an effective amount of said medicament corrects said causal mutation.

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

EXPERIMENTAL

Example 1. Identification of RNA-guided nucleases

CRISPR-associated sequences with sequence similarity to transposases were identified in genomes of interest. CRISPR repeats were identified by minCED (mining CRISPRs in Environmental Datasets) with the minimum number of repeats in array set to two. Only putative RNA-guided nucleases that co-localize with repeats on the same contig were considered for further investigation. Several increasingly stringent cutoffs for distance between repeats and the putative *cas* gene on a contig (100kb, 50kb, 20kb, 10kb) were used. The final filter of 5 kb was selected.

In the literature, CRISPR repeats in active systems are between 27 and 47 nucleotides in length. This feature was used to filter and remove non-CRISPR repetitive features. Part of the acquisition of new spacer sequences in a CRISPR array requires that the first nucleotide of the repeat be a G in order to provide the proper chemistry for array expansion. As a part of this step, the consensus repeat sequence was predicted, as well as the orientation of the repeat-spacer array. This filter was included to prioritize likely functional RGNs. The minimum number of repeats required in an array was increased to three.

Some proteins, mainly DNA-binding proteins, have repeating amino acids in their primary structure that can be falsely detected as CRISPR loci. Putative RGNs whose repetitive features occurred internally in a protein were discarded. Only intergenic repeats were considered for further analysis.

To determine clusters of homologs, the proteins and consensus repeat sequences were aligned and categorized into clusters based on their phylogeny. Repeats and proteins tend to co-cluster phylogenetically, lending weight to the concept of clusters of homologs. The protein cluster information is also displayed in Table 1. Proteins were clustered at 95% identity using cdhit.

The relatedness of strains can be traced by comparing their spacer content. If systems have the same repeat sequence and similar protein sequences, they are said to be related and their spacer content can provide information about their shared history. Homologs in the same clusters tend to share conserved ancestral spacers. Divergent strains have entirely unique spacer content from one another. Clonal isolates will completely share the exact same spacer content. Systems whose repeat sequences are conserved, but whose spacer content varies between genome can promote systems that are more likely to be active and these were prioritized.

109 distinct CRISPR-associated RNA-guided nucleases (RGNs) were identified and are described in Table 1 below. Table 1 provides the name of each RGN, its amino acid sequence, the source from which it was derived, and processed crRNA and tracrRNA sequences.

The locus surrounding each putative nuclease was searched for potential accessory genes that might be needed for CRISPR immunity. Several putative nucleases appeared to be in operonic structures with potential accessory genes, but no loci contained homologs to *cas1*, *cas2*, *cas4* or other known cas genes (Figure 1). Additionally, several systems contained repeats that were flush with the end of the nuclease and lacked the expected leader sequence upstream of the CRISPR repeats, suggesting a novel mechanism of CRISPR RNA expression.

Table 1: RGN system information

RGN ID	SEQ ID NO.	Source	Prot Cluster	crRNA repeat seq (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA (SEQ ID NO)	Consensus Repeat (SEQ ID NO.)
APG07339	1	*Bacillus* sp.	4	110	120	129	201
APG09624	2	*Bacillus* sp.	4	111	121	130	202
APG03003	3	*Bacillus* sp.	4	112	122		203
APG05405	4	*Bacillus* sp.	4	113	123	131	204
APG09777	5	*Bacillus* sp.	4	114	124	132	205
APG05680	6	*Bacillus* sp.	16	115	125	310	206
APG02119	7	*Bacillus* sp.	4				207
APG03285	8	*Bacillus* sp.	15				208
APG04998	9	*Bacillus* sp.	11				209
APG07078	10	*Bacillus* sp.	13				210
APG06369	11	*Gordonia sp.*	0	116			211
APG03847	12	*Micrococcus sp.*	2	117	126	133	212
APG05625	13	*Micrococcus sp.*	2	118	127	134	213
APG03759	14	*Paeniglutamicibacter sp.*	5				214
APG05123	15	*Streptomyces sp.*	3				215
APG03524	16	*Micrococcus sp.*	2	119	128	135	216
APG05361	17	*Bacillus* sp.	1				217

RGN ID	SEQ ID NO.	Source	Prot Cluster	crRNA repeat seq (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA (SEQ ID NO)	Consensus Repeat (SEQ ID NO.)
APG04303	18	*Bacillus* sp.	14				218
APG04291	19	*Bacillus* sp.	14				219
APG01006	20	*Bacillus* sp.	14				220
APG06547	21	*Bacillus* sp.	22				221
APG01699	22	*Bacillus* sp.	14				222
APG06155	23	*Bacillus* sp.	14				223
APG09116	24	*Bacillus* sp.	4				224
APG09403	25	*Bacillus* sp.	23				225
APG08954	26	*Bacillus* sp.	14				226
APG02589	27	*Bacillus* sp.	14				227
APG04061	28	*Bacillus* sp.	14				228
APG08773	29	*Bacillus* sp.	12				229
APG02836	30	*Bacillus* sp.	14				230
APG09123	31	*Bacillus* sp.	14				231
APG04288	32	*Bacillus* sp.	14				232
APG06873	33	*Bacillus* sp.	12				233
APG02381	34	*Bacillus* sp.	14				234
APG06947	35	*Bacillus* sp.	27				ND
APG04677	36	*Bacillus* sp.	14				236
APG07253	37	*Bacillus* sp.	14				237
APG08319	38	*Bacillus* sp.	24				ND
APG04362	39	*Bacillus* sp.	14				239
APG00992	40	*Bacillus* sp.	14				240
APG04193	41	*Bacillus* sp.	17				241
APG08201	42	*Bacillus* sp.	17				242
APG01031	43	*Bacillus* sp.	14				243
APG06773	44	*Bacillus* sp.	14				244
APG08945	45	*Bacillus* sp.	14				245
APG03214	46	*Bacillus* sp.	4				246
APG09942	47	*Bacillus* sp.	4				247
APG01836	48	*Bacillus* sp.	14				248
APG06336	49	*Bacillus* sp.	14				249
APG08839	50	*Bacillus* sp.	4				250
APG02684	51	*Bacillus* sp.	20				ND
APG05281	52	*Bacillus* sp.	14				252
APG01046	53	*Bacillus* sp.	14				253
APG01240	54	*Bacillus* sp.	14				254
APG05981	55	*Bacillus* sp.	14				255
APG04054	56	*Bacillus* sp.	14				256

RGN ID	SEQ ID NO.	Source	Prot Cluster	crRNA repeat seq (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA (SEQ ID NO)	Consensus Repeat (SEQ ID NO.)
APG07301	57	*Bacillus* sp.	14				257
APG07284	58	*Bacillus* sp.	14				258
APG06812	59	*Bacillus* sp.	14				259
APG08143	60	*Bacillus* sp.	14				260
APG08031	61	*Bacillus* sp.	14				261
APG03966	62	*Bacillus* sp.	14				262
APG05371	63	*Bacillus* sp.	14				263
APG04324	64	*Bacillus* sp.	14				264
APG01233	65	*Bacillus* sp.	14				265
APG00823	66	*Bacillus* sp.	14				266
APG06704	67	*Bacillus* sp.	14				267
APG02228	68	*Bacillus* sp.	14				268
APG08636	69	*Bacillus* sp.	6				269
APG02665	70	*Bacillus* sp.	6				270
APG01832	71	*Bacillus* sp.	6				271
APG03625	72	*Bacillus* sp.	14				272
APG07479	73	*Bacillus* sp.	14				273
APG02608	74	*Bacillus* sp.	14				274
APG04337	75	*Bacillus* sp.	12				275
APG01431	76	*Bacillus* sp.	14				276
APG05423	77	*Bacillus* sp.	14				277
APG01452	78	*Bacillus* sp.	14				278
APG05729	79	*Bacillus* sp.	19				279
APG01946	80	*Bacillus* sp.	14				280
APG02414	81	*Bacillus* sp.	14				281
APG01839	82	*Bacillus* sp.	14				282
APG00752	83	*Bacillus* sp.	14				283
APG02156	84	*Bacillus* sp.	14				284
APG08789	85	*Bacillus* sp.	10				285
APG07736	86	*Bacillus* sp.	26				286
APG01573	87	*Bacillus* sp.	14				287
APG07722	88	*Bacillus* sp.	18				288
APG08071	89	*Bacillus* sp.	8				289
APG01280	90	*Bacillus* sp.	14				290
APG07455	91	*Bacillus* sp.	14				291
APG05150	92	*Bacillus* sp.	7				292
APG09405	93	*Bacillus* sp.	14				293
APG09583	94	*Bacillus* sp.	4				294
APG09909	95	*Bacillus* sp.	14				295

RGN ID	SEQ ID NO.	Source	Prot Cluster	crRNA repeat seq (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA (SEQ ID NO)	Consensus Repeat (SEQ ID NO.)
APG07075	96	*Bacillus* sp.	14				296
APG05892	97	*Bacillus* sp.	14				297
APG09648	98	*Bacillus* sp.	25				298
APG02311	99	*Bacillus* sp.	21				299
APG03906	100	*Bacillus* sp.	14				300
APG01953	101	*Bacillus* sp.	9				301
APG00903	102	*Bacillus* sp.	14				302
APG01543	103	*Bacillus* sp.	22				303
APG07261	104	*Bacillus* sp.	15				ND
APG05635	105	*Bacillus* sp.	13				305
APG02962	106	*Bacillus* sp.	15				306
APG07448	107	*Bacillus* sp.	15				307
APG05378	108	*Bacillus* sp.	16				308
APG01852	109	*Bacillus* sp.	16				309

ND=Not Determined

Example 2: Protein Analysis

Nuclease domains were predicted by searching for domains in the interpro database. Split RuvC nuclease domains were predicted using hmm nuclease domain profiles built on split RuvC domains from known Cas proteins. The predicted nuclease residues can be found in Table 2 (ND = Not Determined).

Table 2: Nuclease residues

RGN ID	RuvC-I	RuvC-II	RuvC-III
APG07339	D297	E395	D477
APG09624	D297	E395	D477
APG03003	D297	E395	D477
APG05405	D297	E395	D477
APG09777	D297	E395	D477
APG05680	D232	E331	D407
APG02119	D297	E395	D477
APG03285	D232	E331	D407
APG04998	D281	E379	D456
APG07078	D233	E332	D408
APG06369	ND	ND	ND
APG03847	D282	E382	D463
APG05625	D282	E382	D463
APG03759	D257	E357	D438

RGN ID	RuvC-I	RuvC-II	RuvC-III
APG05123	D248	E356	D437
APG03524	D282	E382	D463
APG05361	D244	D362	D443
APG04303	D233	E332	D408
APG04291	D233	E332	D408
APG01006	D233	E332	D408
APG06547	D183	E277	D359
APG01699	D233	E326	D402
APG06155	D233	E332	D408
APG09116	D297	E395	D477
APG09403	D183	E277	D359
APG08954	D233	E326	D402
APG02589	D236	E329	D405
APG04061	D233	E332	D408
APG08773	D267	E363	D440
APG02836	D233	E332	D408
APG09123	D233	E332	D408
APG04288	D233	E332	D408
APG06873	D267	E363	D440
APG02381	D233	E332	D408
APG06947	D2	E95	D184
APG04677	D233	E326	D402
APG07253	D233	E332	D408
APG08319	D183	E267	D350
APG04362	D233	E332	D408
APG00992	D233	E332	D408
APG04193	D232	E331	D407
APG08201	D232	E331	D407
APG01031	D226	E325	D401
APG06773	D233	E332	D408
APG08945	D233	E332	D408
APG03214	D297	E395	D477
APG09942	D297	E395	D477
APG01836	D233	E332	D408
APG06336	D233	E332	D408
APG08839	D297	E395	D477
APG02684	D221	E305	D387
APG05281	D34	E133	D209
APG01046	D233	E332	D408
APG01240	D233	E326	D402
APG05981	D233	E332	D408
APG04054	D233	E332	D408

RGN ID	RuvC-I	RuvC-II	RuvC-III
APG07301	D233	E332	D408
APG07284	D233	E326	D402
APG06812	D233	E326	D402
APG08143	D233	E332	D408
APG08031	D233	E332	D408
APG03966	D233	E332	D408
APG05371	D233	E332	D408
APG04324	D233	E326	D402
APG01233	D233	E332	D408
APG00823	D233	E332	D408
APG06704	D233	E332	D408
APG02228	D233	E332	D408
APG08636	D267	E365	D442
APG02665	D267	E365	D442
APG01832	D267	E365	D442
APG03625	D233	E332	D408
APG07479	D233	E332	D408
APG02608	D233	E332	D408
APG04337	D267	E365	D442
APG01431	D233	E332	D408
APG05423	D233	E332	D408
APG01452	D233	E332	D408
APG05729	D231	E330	D409
APG01946	D233	E332	D408
APG02414	D226	E325	D401
APG01839	D236	E329	D405
APG00752	D226	E325	D401
APG02156	D236	E329	D405
APG08789	D280	E378	D455
APG07736	D52	E145	D221
APG01573	D226	E319	D395
APG07722	D232	E331	D407
APG08071	D281	E379	D456
APG01280	D233	E332	D408
APG07455	D233	E332	D408
APG05150	D267	E365	D442
APG09405	D233	E326	D402
APG09583	D297	E395	D477
APG09909	D233	E332	D408
APG07075	D233	E326	D402
APG05892	D226	E319	D395
APG09648	D86	E170	D253

RGN ID	RuvC-I	RuvC-II	RuvC-III
APG02311	D221	E305	D387
APG03906	D233	E332	D408
APG01953	D281	E379	D456
APG00903	D233	E332	D408
APG01543	D183	E277	D359
APG07261	D232	E331	D407
APG05635	D233	E332	D408
APG02962	D232	E331	D407
APG07448	D232	E331	D407
APG05378	D232	E331	D407
APG01852	D232	E331	D407

Example 3: Guide RNA Prediction and Confirmation

Cultures of bacteria that natively express the RNA-guided nuclease system under investigation were grown to mid-log phase (OD600 of ~0.600), pelleted, and flash frozen. RNA was isolated from the pellets using a mirVANA miRNA Isolation Kit (Life Technologies, Carlsbad, CA), and sequencing libraries were prepared from the isolated RNA using an NEBNext Small RNA Library Prep kit (NEB, Beverly, MA). The library prep was fractionated on a 6% polyacrylamide gel to capture the RNA species less than 200nt to detect crRNAs and tracrRNAs, respectively. Deep sequencing (75 bp paired-end) was performed on a Next Seq 500 (High Output kit) by a service provider (MoGene, St. Louis, MO). Reads were quality trimmed using Cutadapt and mapped to reference genomes using Bowtie2. A custom RNAseq pipeline was written in python to detect the crRNA and tracrRNA transcripts. Processed crRNA boundaries were determined by sequence coverage of the native repeat spacer array. The anti-repeat portion of the tracrRNA was identified using permissive BLASTn parameters. RNA sequencing depth confirmed the boundaries of the processed tracrRNA by identifying the transcript containing the anti-repeat. Manual curation of RNAs was performed using secondary structure prediction by RNAfold, an RNA folding software. sgRNA cassettes were prepared by DNA synthesis and were generally designed as follows (5'->3'): the processed tracrRNA, operably linked at its 3' end to a 4 bp noncomplementary linker (AAAG; SEQ ID NO: 136), operably linked at its 3' end to the processed repeat portion of the crRNA, operably linked at its 3' end to a 20-30 bp spacer sequence. Other 4 bp noncomplementary linkers may also be used.

For *in vitro* assays, sgRNAs and some tracrRNAs were synthesized by *in vitro* transcription of the sgRNA cassettes with a TranscriptAid T7 High Yield Transcription Kit (ThermoFisher). crRNAs and some tracrRNAs were produced synthetically.

For protein expression and purification, plasmids containing the putative RGNs fused to a C terminal His10 tag were constructed and transformed into BL21 (DE3) strains of *E. coli*. Expression was performed using Magic Media self-inducing media supplemented with kanamycin. After lysis and

clarification, the proteins were purified by immobilized metal affinity chromatography. Further purification of APG05405 was performed using Heparin chromatography.

The longer of the tracrRNAs (SEQ ID NOs: 140, 145, and 147) were produced by *in vitro* transcription (IVT) using a dsDNA template with T7 promoter upstream of the tracrRNA sequence. The template for IVT was amplified by PCR from a synthesized gBlock template (Integrated DNA Technologies). Shorter tracr and crRNAs were produced synthetically.

RNA binding was confirmed by differential scanning fluorimetry (Niesen, F.H., H. Berglund, and M. Vedadi. 2007. *Nat. Protoc.* 2: 2212–2221). Dual RNA complexes were produced by mixing an excess of crRNA with tracrRNA in Annealing Buffer (Synthego, 60 mM KCl 6 mM HEPES pH 7). The candidate effector protein and guide RNA (either dual RNA complex or sgRNA) were incubated at final concentrations of 0.5 μM effector protein and 1 μM guide RNA in phosphate buffered saline (PBS, Thermo Fisher). These were incubated for 20 minutes at room temperature and then mixed 1:1 with Sypro Orange dye solution that had been diluted in PBS. A melt curve was obtained measuring fluorescence intensity (FI) as a function of temperature and the first derivative of the melt curve (dFI/dT) was calculated. A shift in the plot of dFI/dT as a function of temperature of the putative RNP relative to the original nuclease indicates RNA binding and was used to evaluate putative guide RNA combinations. Of the original guide combinations identified, only the full length putative tracrRNA and crRNA were observed to induce a significant shift in the dFI/dT vs. temperature functions for RGNs APG09624 and APG05405. Small shifts in this function were observed for other protein/RNA combinations but were smaller in magnitude than previously observed for functional RNP formation and were not interpreted as indicative of the formation of a functional complex. Peak 1 refers to the temperature associated with the largest observed peak for the given sample. If a second peak is observed, it is indicated in the Peak 2 column. Interpretation of the data regarding the formation of a complex is indicated in the "Binding?" column Table 3. "Yes" indicates binding was observed. "N/A" indicates sufficient data was not available to determine if binding could occur.

Table 3: Binding of RGNs to guide RNA

RGN	crRNA (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	Peak 1 (°C)	Peak 2 (°C)	Binding?
APG03003	None	None	38	Not obs	N/A
APG03003	139	140	38	Not obs	N/A
APG03003	141	142	37	Not obs	N/A
APG09777	None	None	39	Not obs	N/A
APG09777	139	140	40	Not obs	N/A
APG09777	141	142	39	Not obs	N/A
APG07339	None	None	37	Not obs	N/A
APG07339	143	144	37	Not obs	N/A
APG07339	143	145	38	Not obs	N/A
APG09624	None	None	38	Not obs	N/A

RGN	crRNA (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	Peak 1 (°C)	Peak 2 (°C)	Binding?
APG09624	143	144	37	Not obs	N/A
APG09624	143	145	39	44	Yes
APG05405	None	None	36	Not obs	N/A
APG05405	143	144	37	Not obs	N/A
APG05405	143	145	43	Not obs	Yes
APG09777	None	None	38	Not obs	N/A
APG09777	146	148	39	Not obs	N/A
APG09777	146	147	39	Not obs	N/A

Example 4: Direct ssDNA Target Cleavage

Purified APG09624, APG05405, and catalytically inactivated APG05405 (dAPG05405 set forth as SEQ ID NO: 173) were incubated with single guide RNA (sgRNA) Gsg.2 (set forth as SEQ ID NO: 194) in Cutsmart buffer (New England Biolabs B7204S) at a final concentration of 200 nM nuclease and 400 nM sgRNA for 20 minutes. These were then added at a final concentration of 100 nM nuclease to solutions of 5' Cy5 labeled ssDNA referred to herein as LE111 (set forth as SEQ ID NO: 195) or LE113 (set forth as SEQ ID NO: 196) at 10 nM in 1.5X Cutsmart buffer (New England Biolabs B7204S).

Samples were quenched by adding RNase and EDTA at a final concentration of 0.1 mg/mL and 45 mM, respectively, and placed on ice at the following timepoints: 0, 40, 80, and 120 min. After quenching all samples, they were incubated at 50°C for 30 min, then 95°C for 5 min. One-fifth volume of loading buffer (1x TBE, 12% Ficoll, 7 M urea) was added to each reaction and incubated at 95°C for 15 min, and 5 µl of each reaction were analyzed on 15% TBE-urea acrylamide gel (Bio-Rad 3450092).

Quantitation of cleaved product as a function of time, nuclease, and guide RNA is shown below in Table 4. The sequence of LE111 (SEQ ID NO: 195) was used as a negative control, while the sequence of LE113 (SEQ ID NO: 196) bears a target sequence for the sgRNA that was loaded onto the nuclease.

Table 4: Cleavage of Target Oligonucleotide

Nuclease	Target oligonucleotide	Time (min)	Cleavage %
dAPG05405	LE111	0	0
dAPG05405	LE111	0	0
dAPG05405	LE111	40	0
dAPG05405	LE111	40	0
dAPG05405	LE111	80	0
dAPG05405	LE111	80	0
dAPG05405	LE111	120	0
dAPG05405	LE111	120	0
dAPG05405	LE113	0	0
dAPG05405	LE113	0	0
dAPG05405	LE113	40	0
dAPG05405	LE113	40	0

Nuclease	Target oligonucleotide	Time (min)	Cleavage %
dAPG05405	LE113	80	0
dAPG05405	LE113	80	0
dAPG05405	LE113	120	0
dAPG05405	LE113	120	0
APG09624	LE111	0	0
APG09624	LE111	0	0
APG09624	LE111	40	0
APG09624	LE111	40	0
APG09624	LE111	80	24
APG09624	LE111	80	24
APG09624	LE111	120	30
APG09624	LE111	120	26
APG09624	LE113	0	0
APG09624	LE113	0	0
APG09624	LE113	40	59
APG09624	LE113	40	65
APG09624	LE113	80	80
APG09624	LE113	80	84
APG09624	LE113	120	92
APG09624	LE113	120	93
APG05405	LE111	0	0
APG05405	LE111	0	0
APG05405	LE111	40	0
APG05405	LE111	40	0
APG05405	LE111	80	0
APG05405	LE111	80	0
APG05405	LE111	120	0
APG05405	LE111	120	0
APG05405	LE113	0	0
APG05405	LE113	0	0
APG05405	LE113	40	58
APG05405	LE113	40	49
APG05405	LE113	80	80
APG05405	LE113	80	85
APG05405	LE113	120	73
APG05405	LE113	120	88

Gel analysis revealed time-dependent formation of cleavage products primarily in the samples targeted by the sgRNA and with catalytically active nuclease, demonstrating the RNA-guided nuclease activity of these proteins and providing evidence that the critical catalytic residues are correctly defined. Some non-specific activity is observed when the APG09624 RNP is incubated with LE111, which may be due to the fact that the batches of dAPG05405 and APG05405 were purified by an additional

chromatography step, and thus the batch of APG09624 may have some level of background activity due to incomplete removal of nucleases that derived from the expression host.

Example 5: Programmable DNA binding and gene activation

Construction of RGN gene activation and gRNA Mammalian Expression Plasmids

Nuclease constructs for mammalian expression were synthesized. Human codon-optimized APG05405 with N-terminal SV40 (SEQ ID NO: 149) and C-terminal nucleoplasmin NLS sequences (SEQ ID NOs: 149 and 150, respectively), N-terminal 3xFLAG tag (SEQ ID NO: 151), and a C- or N-terminal VPR activation domain (SEQ ID NO: 154; Chavez, et al. 2015, *Nature Methods*, 12(4): 326-328) under the control of a cytomegalovirus (CMV) promoter (SEQ ID NO: 152) was produced and introduced into a mammalian expression vector. Both presumed catalytically active and catalytically inactivated ("dAPG5405") versions of APG05405 were used. Guide RNA expression constructs encoding a single gRNA, each under the control of a human RNA polymerase III U6 promoter (SEQ ID NO: 153), were also produced. Nuclease constructs are indicated in Table 5 below.

Table 5: RGN constructs

Nuclease construct	SEQ ID NO.
APG05405	155
dAPG05405	156
APG05405-VPR	157
dAPG05405-VPR	158
VPR-APG05405	159
VPR-dAPG05405	160

Transfection and Expression in Mammalian Cells

One day prior to transfection, $1x10^4$ HEK293T cells (Sigma) were plated in 96-well plates in Dulbecco's modified Eagle medium (DMEM) plus 10% (vol/vol) fetal bovine serum (Gibco) and 1% Penicillin-Streptomycin (Gibco). The next day when the cells were at 50-60% confluency, 100 ng of an RGN expression plasmid plus 100 ng of a single gRNA expression plasmid were co-transfected using 0.3 μL of Lipofectamine 3000 (Thermo Scientific) per well, following the manufacturer's instructions. After 48 hours of growth, total RNA was harvested using the Cells-to-Ct One Step kit (ThermoFisher).

Taqman Assay for Target Gene Expression

Endogenous genes were chosen which normally have low expression in HEK cells, but which can be induced upon CRISPR activation. RHOXF2 and CD2 were chosen for this purpose. TaqMan gene expression assays are performed using FAM labelled probes for RHOXF2 and CD2, and a VIC labelled probe for ACTB (all probes from ThermoFisher) as a normalization control. TaqMan assays are performed following the manufacturer's instructions in the Cells-to-CT™ One Step kit (Thermofisher) in a BioRad

CFX96 Real Time thermocycler. Background is measured in similar experiments where no gRNA is present. Fold changes in gene expression relative to background are calculated using the $2^{-\Delta\Delta Ct}$ method (Livak et al. 2001, *Methods*, 25(4):402-8), normalizing expression to ACTB transcript levels.

Table 6: Guide RNAs for target gene expression

Nuclease	Gene	Guide (SEQ ID NO)
APG05405-VPR	RHOXF2	167
APG05405-VPR	RHOXF2	168
APG05405-VPR	RHOXF2	169
APG05405-VPR	RHOXF2	170
APG05405-VPR	RHOXF2	171
APG05405-VPR	RHOXF2	172
dAPG05405-VPR	RHOXF2	167
dAPG05405-VPR	RHOXF2	168
dAPG05405-VPR	RHOXF2	169
dAPG05405-VPR	RHOXF2	170
dAPG05405-VPR	RHOXF2	171
dAPG05405-VPR	RHOXF2	172
VPR-APG05405	RHOXF2	167
VPR-APG05405	RHOXF2	168
VPR-APG05405	RHOXF2	169
VPR-APG05405	RHOXF2	170
VPR-APG05405	RHOXF2	171
VPR-APG05405	RHOXF2	172
VPR-dAPG05405	RHOXF2	167
VPR-dAPG05405	RHOXF2	168
VPR-dAPG05405	RHOXF2	169
VPR-dAPG05405	RHOXF2	170
VPR-dAPG05405	RHOXF2	171
VPR-dAPG05405	RHOXF2	172
APG05405-VPR	CD2	161
APG05405-VPR	CD2	162
APG05405-VPR	CD2	163
APG05405-VPR	CD2	164
APG05405-VPR	CD2	165
APG05405-VPR	CD2	166
dAPG05405-VPR	CD2	161
dAPG05405-VPR	CD2	162
dAPG05405-VPR	CD2	163
dAPG05405-VPR	CD2	164
dAPG05405-VPR	CD2	165

Nuclease	Gene	Guide (SEQ ID NO)
dAPG05405-VPR	CD2	166
VPR-APG05405	CD2	161
VPR-APG05405	CD2	162
VPR-APG05405	CD2	163
VPR-APG05405	CD2	164
VPR-APG05405	CD2	165
VPR-APG05405	CD2	166
VPR-dAPG05405	CD2	161
VPR-dAPG05405	CD2	162
VPR-dAPG05405	CD2	163
VPR-dAPG05405	CD2	164
VPR-dAPG05405	CD2	165
VPR-dAPG05405	CD2	166

Example 6: Programmable DNA binding and Base Editing

Oligonucleotides and PCR

All PCR reactions described below are performed using 10 µl of 2X Master Mix Phusion High-Fidelity DNA polymerase (Thermo Scientific) in a 20 µl reaction including 0.5 uM of each primer. Large genomic regions encompassing each target gene are first amplified using PCR#1 primers, using a program of: 98°C, 1 min; 30 cycles of [98°C, 10 sec; 62°C, 15 sec; 72°C, 5 min]; 72°C, 5 min; 12°C, forever. One microliter of this PCR reaction is then further amplified using primers specific for each guide (PCR#2 primers), using a program of: 98°C, 1 min; 35 cycles of [98°C, 10 sec; 67°C, 15 sec; 72°C, 30 sec]; 72°C, 5 min; 12°C, forever. Primers for PCR#2 include Nextera Read 1 and Read 2 Transposase Adapter overhang sequences for Illumina sequencing.

Construction of RGN Base Editing and gRNA Mammalian Expression Plasmids

Nuclease constructs for mammalian expression are synthesized. Human codon-optimized APG05405 with N-terminal SV40 (SEQ ID NO: 149) and C-terminal nucleoplasmin NLS sequences (SEQ ID NOs: 149 and 150, respectively), N-terminal 3xFLAG tag (SEQ ID NO: 151) and an N-terminal deaminase (for example, hAPOBEC3A; SEQ ID NO: 177) under the control of a cytomegalovirus (CMV) promoter (SEQ ID NO: 152) is produced and introduced into a mammalian expression vector. A catalytically inactivated version of APG05405 ("dAPG5405") is used. Guide RNA expression constructs encoding a single gRNA each under the control of a human RNA polymerase III U6 promoter (SEQ ID NO: 153) are also produced.

Transfection and Expression in Mammalian Cells

One day prior to transfection, $1x10^5$ HEK293T cells (Sigma) are plated in 24-well dishes in Dulbecco's modified Eagle medium (DMEM) plus 10% (vol/vol) fetal bovine serum (Gibco) and 1% Penicillin-Streptomycin (Gibco). When the cells are at 50-60% confluency, 500ng of a APG05405 expression plasmid plus 500ng of a single gRNA expression plasmid are co-transfected using 1.5uL of Lipofectamine 3000 (Thermo Scientific) per well, following the manufacturer's instructions. After 48 hours of growth, total genomic DNA is harvested using a genomic DNA isolation kit (Machery-Nagel) according to the manufacturer's instructions.

Next Generation Sequencing

Products from PCR#2 containing Illumina overhang sequences underwent library preparation following the Illumina 16S Metagenomic Sequencing Library protocol. Deep sequencing is performed on an Illumina Mi-Seq platform by a service provider (MOGene). Typically, 200,000 250 bp paired-end reads (2 x 100,000 reads) are generated per amplicon. The reads are analyzed using CRISPResso (Pinello, et al. 2016 *Nature Biotech,* 34:695-697) to calculate the rates of base editing. Output alignments are hand-curated to confirm base editing windows as well as identify insertions or deletions. Each position across the target is analyzed to determine the editing rate and specific nucleotide changes that occur at each position.

Example 7: Trans ssDNA cleavage

7.1 Determining assay conditions for trans DNA cleavage

Purified APG05405 was incubated with single guide RNA (sgRNA) in Cutsmart buffer (New England Biolabs B7204S) at a final concentration of either 50 nM nuclease and 100 nM sgRNA or 200 nM nuclease and 400 nM sgRNA for 10 min. These RNP solutions were added to solutions of ssDNA – a target or mismatched negative control ssDNA - at a final concentration of 10 nM and a reporter ssDNA probe at a final concentration of 250 nM in 1.5X Cutsmart buffer (New England Biolabs B7204S). The reporter probes (TB0125 and TB0089 set forth as SEQ ID NOs: 197 and 198, respectively) contain a fluorescent dye at the 5' end (56-FAM for TB0125 and Cy5 for TB0089), a quencher at the 3' end (3IABkFQ for TB0125 and 3IAbRQSp for TB0089), and optionally an internal quencher (the internal quencher ZEN is only present on TB0125). Cleavage of the reporter probe results in dequenching of the fluorescent dye and thus an increase in fluorescence signal. To monitor fluorescence intensity, 10 µL of each reaction was incubated in a Corning low volume 384-well microplate at 30°C in a microplate reader (CLARIOstar Plus).

A number of conditions were scouted in order to determine suitable parameters for this assay. In order to determine if there are effects of quenched probe design or fluorophore characteristics, two such reporters were included as a mixture in each reaction. They were at the same concentration as each other in any given reaction. In all cases, the control or target ssDNA concentration (LE201 or LE205 set forth as SEQ ID NOs: 199 and 200, respectively) was 10 nM. The RNP names signify the nuclease and the target as indicated in Table 7 below.

Table 7. Ribonucleoprotein complexes

Nuclease	sgRNA (SEQ ID NO)	Intended Target	RNP Name
APG05405	Gsg.1 (193)	LE201	APG05405.1
APG05405	Gsg.2 (194)	LE205	APG05405.2

The results are shown in Table 8 below.

5

Table 8. Results of trans DNA cleavage assays

				Slope (Arbitrary Units / min)	
RNP	**[RNP] (nM)**	**[reporters] (nM)**	**Target**	**Cy5 channel**	**FAM channel**
APG05405.1	25 nM	0	LE201	-0.60	-0.60
APG05405.1	25 nM	50	LE201	72.60	4.20
APG05405.1	25 nM	250	LE201	198.60	-1.20
APG05405.1	25 nM	500	LE201	294.00	-3.00
APG05405.1	25 nM	750	LE201	428.40	-16.80
APG05405.1	25 nM	1000	LE201	509.40	-19.80
APG05405.1	25 nM	0	LE205	-0.48	-0.42
APG05405.1	25 nM	50	LE205	6.60	-1.80
APG05405.1	25 nM	250	LE205	57.00	0.60
APG05405.1	25 nM	500	LE205	185.40	-6.60
APG05405.1	25 nM	750	LE205	318.60	-14.40
APG05405.1	25 nM	1000	LE205	483.60	3.60
APG05405.2	25 nM	0	LE201	0.00	-2.40
APG05405.2	25 nM	50	LE201	12.00	-3.00
APG05405.2	25 nM	250	LE201	81.00	3.60
APG05405.2	25 nM	500	LE201	232.20	7.80
APG05405.2	25 nM	750	LE201	423.60	7.80
APG05405.2	25 nM	1000	LE201	684.60	9.60
APG05405.2	25 nM	0	LE205	-0.60	-2.40
APG05405.2	25 nM	50	LE205	173.40	44.40
APG05405.2	25 nM	250	LE205	262.80	49.20
APG05405.2	25 nM	500	LE205	381.00	44.40
APG05405.2	25 nM	750	LE205	530.40	31.80
APG05405.2	25 nM	1000	LE205	678.00	21.00
APG05405.1	100 nM	0	LE201	0.36	-3.00
APG05405.1	100 nM	50	LE201	97.20	5.40
APG05405.1	100 nM	250	LE201	429.00	19.20
APG05405.1	100 nM	500	LE201	651.00	26.40
APG05405.1	100 nM	750	LE201	1176.00	29.40
APG05405.1	100 nM	1000	LE201	1601.40	29.40
APG05405.1	100 nM	0	LE205	-1.80	-0.36
APG05405.1	100 nM	50	LE205	24.00	0.24
APG05405.1	100 nM	250	LE205	142.20	3.60

				Slope (Arbitrary Units / min)	
RNP	**[RNP] (nM)**	**[reporters] (nM)**	**Target**	**Cy5 channel**	**FAM channel**
APG05405.1	100 nM	500	LE205	331.80	5.40
APG05405.1	100 nM	750	LE205	590.40	10.80
APG05405.1	100 nM	1000	LE205	952.80	9.00
APG05405.2	100 nM	0	LE201	-0.36	-3.00
APG05405.2	100 nM	50	LE201	30.60	-4.20
APG05405.2	100 nM	250	LE201	213.60	4.80
APG05405.2	100 nM	500	LE201	499.20	6.00
APG05405.2	100 nM	750	LE201	1128.00	-1.80
APG05405.2	100 nM	1000	LE201	1682.40	7.20
APG05405.2	100 nM	0	LE205	0.06	-1.80
APG05405.2	100 nM	50	LE205	389.40	72.60
APG05405.2	100 nM	250	LE205	1123.80	197.40
APG05405.2	100 nM	500	LE205	1510.20	235.20
APG05405.2	100 nM	750	LE205	2048.40	256.20
APG05405.2	100 nM	1000	LE205	2393.40	234.00

From this experiment, it was concluded that the 100 nM concentration of RNP results generally in a higher cleavage rate of the reporter probe than the 25 nM RNP concentration. In general, reporter cleavage rates are higher at higher concentrations of the reporter oligonucleotides up to 250 nM reporter concentration, with little benefit observed from further increase in reporter concentrations. Notably, for the TB0089 reporter (detected in the Cy5 channel), there are substantially higher levels of background activity that interfere with target differentiation, especially at reporter concentrations higher than 250 nM. Therefore, it was concluded that reporter concentrations higher than 250 nM would not be beneficial and that in general, the doubly quenched TB0125 probe (detected in the FAM channel) is more suitable for future experiments since it provides a higher ratio of specific to background activity at a wide range of reporter concentrations.

7.2 APG09624 trans DNA cleavage and effect of purification on non-specific activity

Purified APG05405 and APG09624 were incubated with single guide RNA (sgRNA) as indicated below in 1X Cutsmart buffer (New England Biolabs B7204S) at a final concentration of 200 nM nuclease and 400 nM sgRNA for 10 min at 37°C.

Table 9. Ribonucleoprotein complexes

Nuclease	Nuclease batch	sgRNA	Intended Target	RNP Name
APG09624	N/A	Gsg.2	LE113	APG09624.2
APG05405	A	Gsg.2	LE113	APG05405.2A
APG05405	B	Gsg.2	LE113	APG05405.2B

These RNP solutions were then added to solutions of ssDNA – a target or mismatched negative control ssDNA - at a final concentration of 10 nM and a reporter probe (TP0003; set forth as SEQ ID NO: 314) at a final concentration of 250 nM in 1.5X Cutsmart buffer (New England Biolabs B7204S). The reporter probe contains a fluorescent dye at the 5' end and a quencher at the 3' end. Cleavage of the reporter probe results in dequenching of the fluorescent dye and thus an increase in fluorescence signal. To monitor fluorescence intensity, 10 µl of each reaction was incubated in a Corning low volume 384-well microplate at 37°C in a microplate reader (CLARIOstar Plus).

Incubation with target sequences resulted in a substantial increase in fluorescence intensity as a function of time relative to the negative control. The rate of cleavage is expediently summarized as the slope of the linear portion of the fluorescence vs. time function as shown in Table 10.

Table 10. Results of trans DNA cleavage assay

RNP	ssDNA Target Sequence	Slope (Arb. Units / min)
APG09624.2	LE111	182
APG09624.2	LE113	561
APG05405.2A	LE111	197
APG05405.2A	LE113	1625
APG05405.2B	LE111	73
APG05405.2B	LE113	1054

These data show differentiation of the target sequences by the RNPs formed from both APG05405 and APG09624.

7.3 Trans DNA cleavage activation by PCR products

Oligonucleotides containing degenerate nucleotides 5' of the target were PCR amplified to produce target sequences. Target dsPCR2 and dsPCR3 contain target sequences ACTACAACAGCCACAACGTCTATATCATGG (dsPCR2) and TGGAATGGGAACTAAAGTAATGG (dsPCR3) set forth as SEQ ID NOs: 311 and 312, respectively) contained an 8 bp and 5 bp degenerate region, respectively, on the 5' side of the target encoded by the guide RNA. The oligo pairs were annealed and PCR amplified with appropriate primers. In addition, ssDNA targets were included in the experiment – oligonucleotides containing the reverse complement of the target sequences described above CCATGATATAGACGTTGTGGCTGTTGTAGT (LE205; SEQ ID NO: 200) and CCATTACTTTAGTTCCCATTCCA (LE501; SEQ ID NO: 174).

RNP solutions were formed by incubation of nuclease and sgRNA at 0.5 µM and 1 µM, respectively in 1X NEBuffer 2 (New England Biolabs) and incubated at room temperature for 20 minutes.

Table 11. Ribonucleoprotein complexes

sgRNA	Intended Targets	RNP Name
Gsg.3 (SEQ ID NO: 175)	dsPCR3, LE501	APG05405.3
Gsg.2 (SEQ ID NO: 194)	dsPCR2, LE205	APG05405.2

The cleavage reaction was performed in 1.5X NEBuffer 2 with 1.5 μM reporter with a 5' TEX615 label and a 3' Iowa Black FQ quencher and 100 nM of the respective PCR product or ssDNA oligonucleotide LE501 or LE205 (set forth as SEQ ID NOs: 174 and 200, respectively; LE501 comprised a 5' FAM fluorophore). Cleavage of the reporter probe results in dequenching of the fluorescent dye and thus an increase in fluorescence signal. To monitor fluorescence intensity, 10 μl of each reaction was incubated in a Corning low volume 384-well microplate at 37°C in a microplate reader (CLARIOstar Plus). The results of kinetic analysis are shown in Table 12.

Table 12. Results of trans DNA cleavage assay

Target	RNP	Slope (Arb. Units / min)
dsPCR3	APG05405.3	780
	APG05405.2	102
dsPCR2	APG05405.3	78
	APG05405.2	768
LE501	APG05405.3	624
	APG05405.2	120
LE205	APG05405.3	90
	APG05405.2	1434

These results demonstrate target sequence specific activation of non-sequence specific cleavage of ssDNA. Both dsDNA PCR products and target ssDNA oligonucleotides were able to induce this activity.

7.4 PAM determination by induced non-specific ssDNA cleavage

Isolation of the ternary complex containing a DNA target (including a PAM library) with the RNP and sequencing of the DNA recovered from it can be used to identify the PAM sequence, if the given system requires a PAM for DNA binding, modification or cleavage. The complex can be captured by a number of methods, such as immuno-pulldown, capture with immobilized metal affinity resin (such as Ni-NTA agarose), or isolation by size exclusion chromatography.

Alternatively, a parallel library of DNA fragments with distinct PAM sequences adjacent to a fixed target is produced. RNPs containing the putative nuclease and a suitable guide RNA (single guide or dual-RNA guide targeting the fixed target) are incubated with each of the fragments in the library and assessed for binding using a shift in electrophoretic mobility of the DNA fragment, size exclusion liquid chromatography, or co-precipitation using a solid support with affinity for either component.

7.5 PAM determination using a parallel plasmid DNA library

A plasmid library is produced containing a target sequence, ACTACAACAGCCACAACGTCTATATCATGG (set forth as SEQ ID NO: 313), preceded by an 8 bp degenerate sequence (NNNNNNNN set forth as SEQ ID NO: 176). Upon transformation into competent cells and plating onto agar plates with selective media, each colony resulting from the transformation of this

reaction corresponds in principle to a clonal plasmid DNA sequence and thus preparations of plasmid DNA from cultures deriving from single colonies are unique plasmid preparations sampled from the original library. Plasmid preparations are obtained from a sampling of 96 colonies. These preparations are individually subjected to Sanger sequencing to verify their PAM sequence.

Purified APG05405 is incubated with single guide RNA (sgRNA) Gsg.2 (set forth as SEQ ID NO: 194) in 1X Cutsmart buffer (New England Biolabs B7204S) at a final concentration of 200 nM nuclease and 400 nM sgRNA for 20 minutes at room temperature.

These RNP solutions were added at a final concentration of 100 nM to solutions of plasmid DNA and ssDNA reporter strand comprising a fluorophore and a quencher at 50 nM in 1.5X Cutsmart buffer (New England Biolabs B7204S). To monitor fluorescence intensity as a function of time, 10 µl of each reaction is incubated in a Corning low volume 384-well microplate at 37°C in a microplate reader (CLARIOstar Plus). Each well corresponds to an individual digestion reaction with a specific PAM sequence. Upon completion of data collection, the rate of fluorescence increase will be determined. A consensus PAM sequence will be built by analyzing the sequences that correspond to wells with high rates of fluorescence increase. If inconclusive, an additional library can be produced and evaluated.

Example 8: Use of ssDNA cleavage as a diagnostic

Due to the capability of these nucleases to generate an optically detectable signal in the presence of a target DNA sequence, they promise utility for implementation into diagnostic devices for the detection of genetic diseases or agents of infectious disease, such as bacteria, viruses, or fungi.

A diagnostic procedure may include isolation or amplification of nucleic acids from a sample to be tested. It may also be suitable to use some samples without performing any isolation or purification of nucleic acids, as they may be present in the sample at high enough quantities to be detectable without amplification (such as PCR) or free of materials that interfere with detection or signal production.

RNPs formed as described in the other examples could then be exposed to the sample (or processed sample as described in the preceding paragraph) along with a reporter, such as the fluorophore and quencher modified ssDNA oligonucleotides used in previous examples, or some other sort of ssDNA substrate that produces a visible or otherwise easily detectable signal when cleaved. If using fluorophore-quencher conjugated DNA oligonucleotides (as in the previously described examples), these can be detected using a fluorimeter as described in previous examples. To simplify detection, an endpoint assay can be performed instead of the kinetic assays described above, meaning that the assays can be performed for a fixed time and read out at the end of this elapsed time, relative to positive and negative controls.

These reagents may also be integrated into a lateral flow testing device which allows for the detection of a given disease-causing agent or specific nucleic acid sequence (such as a diseased allele in an individual) with very little instrumentation. In this assay, the ssDNA reporter would be conjugated to multiple molecules suitable for antibody or affinity reagent capture, such as fluorescein, biotin, and/or digoxigenin.

Example 8.1 – COVID19 diagnostic assay

Samples are collected from patients using nasopharyngeal swabs in universal transport medium (UTM) according to standard practice, and RNA is extracted. Reverse Transcription Loop-Mediated Isothermal Amplification (RT-LAMP) is used to amplify the genetic material, similar to Broughton *et al* 2020 (*Nat. Biotechnol.* 38: 870-874). In some embodiments, single-stranded DNA (ssDNA) produced by RT-LAMP may be sufficient for PAM-independent detection by an RGN disclosed herein. In some embodiments, the RT-LAMP produces ssDNA using amplification by a phosphorothioate primer only on the target strand, allowing for T7 exonuclease digestion of the non-target strand.

RT-LAMP amplification is performed with appropriate primers to amplify the N gene and the E gene of the SARS-CoV2 genome, as well as human RNase P as a quality control check for sample collection and preparation, similar to Broughton *et al* 2020. One of the two LAMP internal primers (commonly called FIP or BIP) would contain phosphorothioate groups. Upon treatment of the completed PCR reaction with T7 exonuclease, the ssDNA extended from the phosphorothioate primer is the major species present in the solution. Guides against this sequence are evaluated for specific and efficient activation using the fluorescence assays described above. For specificity, the detection scheme can be tested against homologous genes in other coronaviruses, to ensure lack of cross-reactivity, such as HCoV-OC43, HCoV-HKU1, HCoV-229E, HCoV-NL63, MERS-CoV, and/or SARS-CoV. The assay may be converted into a lateral flow assay by utilizing an oligonucleotide containing FAM and biotin.

THAT WHICH IS CLAIMED:

1. A nucleic acid molecule comprising a polynucleotide encoding an RNA-guided nuclease (RGN) polypeptide, wherein said polynucleotide comprises a nucleotide sequence encoding an RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109;

wherein said RGN polypeptide is capable of binding a target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to said target DNA sequence, and

wherein said polynucleotide encoding an RGN polypeptide is operably linked to a promoter heterologous to said polynucleotide.

2. The nucleic acid molecule of claim 1, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

3. The nucleic acid molecule of claim 1, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

4. The nucleic acid molecule of any one of claims 1-3, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

5. The nucleic acid molecule of claim 4, wherein said RGN polypeptide is capable of cleaving said target DNA sequence upon binding.

6. The nucleic acid molecule of any one of claims 1-3, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

7. The nucleic acid molecule of claim 6, wherein said RGN polypeptide is capable of cleaving said target DNA sequence upon binding.

8. The nucleic acid molecule of claim 7, wherein cleavage by said RGN polypeptide generates a double-stranded break.

9. The nucleic acid molecule of claim 7, wherein cleavage by said RGN polypeptide generates a single-stranded break.

10. The nucleic acid molecule of any one of claims 6-9, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

11. The nucleic acid molecule of claim 10, wherein the base-editing polypeptide is a deaminase.

12. The nucleic acid molecule of any one of claims 6-11, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

13. The nucleic acid molecule of any one of claims 1-12, wherein the RGN polypeptide comprises one or more nuclear localization signals.

14. The nucleic acid molecule of any one of claims 1-13, wherein the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

15. A vector comprising the nucleic acid molecule of any one of claims 1-14.

16. The vector of claim 15, further comprising at least one nucleotide sequence encoding said gRNA capable of hybridizing to said target DNA sequence.

17. The vector of claim 16, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said gRNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

18. The vector of claim 16, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said gRNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

19. The vector of claim 16, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said gRNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

20. The vector of claim 16, wherein said gRNA comprises a tracrRNA.

21. The vector of claim 20, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence

identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

22. The vector of claim 20, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

23. The vector of claim 20, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said gRNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

24. The vector of any one of claims 20-23, where said gRNA is a single guide RNA.

25. The vector of any one of claims 20-23, wherein said gRNA is a dual-guide RNA.

26. A cell comprising the nucleic acid molecule of any one of claims 1-14 or the vector of any one of claims 15-25.

27. A method for making an RGN polypeptide comprising culturing the cell of claim 26 under conditions in which the RGN polypeptide is expressed.

28. A method for making an RGN polypeptide comprising introducing into a cell a heterologous nucleic acid molecule comprising a nucleotide sequence encoding an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3 and 5 to 109;

wherein said RGN polypeptide binds a target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to said target DNA sequence;

and culturing said cell under conditions in which the RGN polypeptide is expressed.

29. The method of claim 28, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

30. The method of claim 28, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

31. The method of any one of claims 28-30, further comprising purifying said RGN polypeptide.

32. The method of any one of claims 28-30, wherein said cell further expresses one or more guide RNAs that binds to said RGN polypeptide to form an RGN ribonucleoprotein complex.

33. The method of claim 32, further comprising purifying said RGN ribonucleoprotein complex.

34. An isolated RNA-guided nuclease (RGN) polypeptide, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109; and

wherein said RGN polypeptide is capable of binding a target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA (gRNA) capable of hybridizing to said target DNA sequence.

35. The isolated RGN polypeptide of claim 34, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

36. The isolated RGN polypeptide of claim 34, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

37. The isolated RGN polypeptide of any one of claims 34-36, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

38. The isolated RGN polypeptide of claim 37, wherein said RGN polypeptide is capable of cleaving said target DNA sequence upon binding.

39. The isolated RGN polypeptide of any one of claims 34-36, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

40. The isolated RGN polypeptide of claim 39, wherein said RGN polypeptide is capable of cleaving said target DNA sequence upon binding.

41. The isolated RGN polypeptide of claim 40, wherein cleavage by said RGN polypeptide generates a double-stranded break.

42. The isolated RGN polypeptide of claim 40, wherein cleavage by said RGN polypeptide generates a single-stranded break.

43. The isolated RGN polypeptide of any one of claims 39-42, wherein the RGN polypeptide is operably fused to a base-editing polypeptide.

44. The isolated RGN polypeptide of claim 43, wherein the base-editing polypeptide is a deaminase.

45. The isolated RGN polypeptide of any one of claims 34-44, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

46. The isolated RGN polypeptide of any one of claims 34-45, wherein the RGN polypeptide comprises one or more nuclear localization signals.

47. A system for binding a target DNA sequence of a DNA molecule, said system comprising:

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

b) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109 or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide;

wherein at least one of said nucleotide sequence encoding the one or more guide RNAs and said nucleotide sequence encoding the RGN polypeptide is 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 RGN polypeptide in order to direct said RGN polypeptide to bind to said target DNA sequence of the DNA molecule.

48. A system for binding a target DNA sequence of a DNA molecule, said system comprising:

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

b) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109;

wherein the one or more guide RNAs are capable of hybridizing to the target DNA sequence, and

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

49. The system of claim 48, wherein at least one of said nucleotides sequences encoding the one or more guide RNAs is operably linked to a promoter heterologous to said nucleotide sequence.

50. The system of any one of claims 47-49, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

51. The system of any one of claims 47-49, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

52. The system of any one of claims 47-51, wherein said RGN polypeptide and said one or more guide RNAs are not found complexed to one another in nature.

53. The system of any one of claims 47-51, wherein said target DNA sequence is a eukaryotic target DNA sequence.

54. The system of any one of claims 47-53, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

55. The system of any one of claims 47-53, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

56. The system of any one of claims 47-53, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

57. The system of any one of claims 47-53, wherein said one or more guide RNAs comprise a tracrRNA.

58. The system of claim 57, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

59. The system of claim 57, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least

95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

60. The system of claim 57, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

61. The system of any one of claims 57-60, wherein said one or more guide RNAs are a single guide RNA (sgRNA).

62. The system of any one of claims 57-60, wherein said one or more guide RNAs are a dual-guide RNA.

63. The system of any one of claims 47-62, wherein the target DNA sequence is within a cell.

64. The system of claim 63, wherein the cell is a eukaryotic cell.

65. The system of claim 64, wherein the eukaryotic cell is a plant cell.

66. The system of claim 64, wherein the eukaryotic cell is a mammalian cell.

67. The system of claim 64, wherein the eukaryotic cell is an insect cell.

68. The system of claim 63, wherein the cell is a prokaryotic cell.

69. The system of any one of claims 47-68, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

70. The system of claim 69, wherein when transcribed the one or more guide RNAs is capable of hybridizing to the target DNA sequence and the guide RNA is capable of forming a complex with the RGN polypeptide to direct cleavage of the target DNA sequence.

71. The system of any one of claims 47-68, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

72. The system of claim 71, wherein when transcribed the one or more guide RNAs is capable of hybridizing to the target DNA sequence and the guide RNA is capable of forming a complex with the RGN polypeptide to direct cleavage of the target DNA sequence.

73. The system of claim 72, wherein said RGN polypeptide is capable of generating a double-stranded break.

74. The system of claim 72, wherein said RGN polypeptide is capable of generating a single-stranded break.

75. The system of any one of claims 71-74, wherein the RGN polypeptide is operably linked to a base-editing polypeptide.

76. The system of claim 75, wherein said base-editing polypeptide is a deaminase.

77. The system of any one of claims 71-76, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

78. The system of any one of claims 47-77, wherein the RGN polypeptide comprises one or more nuclear localization signals.

79. The system of any one of claims 47-78, wherein the RGN polypeptide is codon optimized for expression in a eukaryotic cell.

80. The system of any one of claims 47-79, wherein polynucleotides comprising the nucleotide sequences encoding the one or more guide RNAs and the polynucleotide comprising the nucleotide sequence encoding an RGN polypeptide are located on one vector.

81. The system of any one of claims 47-80, wherein said system further comprises one or more donor polynucleotides or one or more polynucleotides comprising one or more nucleotide sequences encoding the one or more donor polynucleotides.

82. A pharmaceutical composition comprising the nucleic acid molecule of any one of claims 1-14, the vector of any one of claims 15-25, the cell of claim 26, the isolated RGN polypeptide of any one of claims 34-46, and the system of any one of claims 47-81, and a pharmaceutically acceptable carrier.

83. A method for binding a target DNA sequence of a DNA molecule comprising delivering a system according to any one of claims 47-81, to said target DNA sequence or a cell comprising the target DNA sequence.

84. The method of claim 83, wherein said RGN polypeptide or said guide RNA further comprises a detectable label, thereby allowing for detection of said target DNA sequence.

85. The method of claim 83, wherein said guide RNA or said RGN polypeptide further comprises an expression modulator, thereby modulating expression of said target DNA sequence or a gene under transcriptional control by said target DNA sequence.

86. A method for cleaving a target DNA sequence of a DNA molecule comprising delivering a system according to any one of claims 47-81, to said target DNA sequence or a cell comprising the target DNA sequence.

87. The method of claim 86, wherein said modified target DNA sequence comprises insertion of heterologous DNA into the target DNA sequence.

88. The method of claim 86, wherein said modified target DNA sequence comprises deletion of at least one nucleotide from the target DNA sequence.

89. The method of claim 86, wherein said modified target DNA sequence comprises mutation of at least one nucleotide in the target DNA sequence.

90. A method for binding a target DNA sequence of a DNA molecule, said method comprising:

a) assembling an RNA-guided nuclease (RGN) ribonucleotide complex *in vitro* by combining:

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

ii) an RGN polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109;

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

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

wherein the one or more guide RNAs hybridize to the target DNA sequence, thereby directing said RGN polypeptide to bind to said target DNA sequence.

91. The method of claim 90, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

92. The method of claim 90, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

93. The method of claim 92, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

94. The method of any one of claims 90-93, wherein said RGN polypeptide or said guide RNA further comprises a detectable label, thereby allowing for detection of said target DNA sequence.

95. The method of any one of claims 90-93, wherein said guide RNA or said RGN polypeptide further comprises an expression modulator, thereby allowing for the modulation of expression of said target DNA sequence.

96. A method for cleaving and/or modifying a target DNA sequence of a DNA molecule, comprising contacting the DNA molecule with:

a) an RNA-guided nuclease (RGN) polypeptide, wherein said RGN comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109; and

b) one or more guide RNAs capable of targeting the RGN of (a) to the target DNA sequence;

wherein the one or more guide RNAs hybridize to the target DNA sequence, thereby directing said RGN polypeptide to bind to said target DNA sequence and cleavage and/or modification of said target DNA sequence occurs.

97. The method of claim 96, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

98. The method of claim 96, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

99. The method of claim 98, wherein cleavage by said RGN polypeptide generates a double-stranded break.

100. The method of claim 98, wherein cleavage by said RGN polypeptide generates a single-stranded break.

101. The method of any one of claims 98-100, wherein said RGN polypeptide is operably linked to a base-editing polypeptide.

102. The method of claim 101, wherein said base-editing polypeptide comprises a deaminase.

103. The method of any one of claims 98-102, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

104. The method of any one of claims 96-103, wherein said modified target DNA sequence comprises insertion of heterologous DNA into the target DNA sequence.

105. The method of any one of claims 96-103, wherein said modified target DNA sequence comprises deletion of at least one nucleotide from the target DNA sequence.

106. The method of any one of claims 96-103, wherein said modified target DNA sequence comprises mutation of at least one nucleotide in the target DNA sequence.

107. The method of any one of claims 90-106, wherein said RGN comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

108. The method of any one of claims 90-106, wherein said RGN comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

109. The method of any one of claims 90-108, wherein said RGN polypeptide and said one or more guide RNAs are not found complexed to one another in nature.

110. The method of any one of claims 90-109, wherein said target DNA sequence is a eukaryotic target DNA sequence.

111. The method of any one of claims 90-110, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

112. The method of any one of claims 90-110, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

113. The method of any one of claims 90-110, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

114. The method of any one of claims 90-110, wherein said one or more guide RNAs comprise a tracrRNA.

115. The method of claim 114, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

116. The method of claim 114, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

117. The method of claim 114, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100%

sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

118. The method of any one of claims 114-117, wherein said one or more guide RNAs are a single guide RNA (sgRNA).

119. The method of any one of claims 114-117, wherein said one or more guide RNAs are a dual-guide RNA.

120. The method of any one of claims 83-119, wherein the target DNA sequence is within a cell.

121. The method of claim 120, wherein the cell is a eukaryotic cell.

122. The method of claim 121, wherein the eukaryotic cell is a plant cell.

123. The method of claim 121, wherein the eukaryotic cell is a mammalian cell.

124. The method of claim 121, wherein the eukaryotic cell is an insect cell.

125. The method of claim 120, wherein the cell is a prokaryotic cell.

126. A cell comprising a modified target DNA sequence according to the method of any one of claims 96-119.

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

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

129. A plant comprising the cell of claim 128.

130. A seed comprising the cell of claim 128.

131. The cell of claim 127, wherein the eukaryotic cell is a mammalian cell.

132. The cell of claim 131, wherein the mammalian cell is a human cell.

133. The cell of claim 132, wherein the human cell is an immune cell.

134. The cell of claim 133, wherein the human cell is a stem cell.

135. The cell of claim 134, wherein the stem cell is an induced pluripotent stem cell.

136. The cell of claim 127, wherein the eukaryotic cell is an insect cell.

137. The cell of claim 126, wherein the cell is a prokaryotic cell.

138. A pharmaceutical composition comprising the cell of any one of embodiments 127 and 131-135, and a pharmaceutically acceptable carrier.

139. A kit for detecting a target DNA sequence of a DNA molecule in a sample, the kit comprising:

a) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109 or a polynucleotide comprising a nucleotide sequence encoding the RGN polypeptide, wherein said RGN polypeptide is capable of binding and cleaving said target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA capable of hybridizing to said target DNA sequence;

b) said guide RNA or a polynucleotide comprising a nucleotide sequence encoding said guide RNA; and

c) a detector single-stranded DNA (ssDNA) that does not hybridize with the guide RNA.

140. The kit of claim 139, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

141. The kit of claim 139, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

142. The kit of any one of claims 139-141, wherein at least one of said nucleotide sequence encoding the guide RNA and said nucleotide sequence encoding the RGN polypeptide is operably linked to a promoter heterologous to said nucleotide sequence.

143. The kit of any one of claims 139-142, wherein said RGN polypeptide and said one or more guide RNAs are not found complexed to one another in nature.

144. The kit of any one of claims 139-142, wherein said target DNA sequence is a eukaryotic target DNA sequence.

145. The kit of any one of claims 139-144, wherein said detector ssDNA comprises a fluorophore/quencher pair.

146. The kit of any one of claims 139-144, wherein said detector ssDNA comprises a fluorescence resonance energy transfer (FRET) pair.

147. The kit of any one of claims 139-146, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

148. The kit of any one of claims 139-146, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

149. The kit of any one of claims 139-146, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said one or more guide RNAs comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

150. The kit of any one of claims 139-146, wherein said one or more guide RNAs comprise a tracrRNA.

151. The kit of claim 150, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

152. The kit of claim 150, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least

95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

153. The kit of claim 150, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said one or more guide RNAs further comprise a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

154. The kit of any one of claims 150-153, wherein said one or more guide RNAs are a single guide RNA (sgRNA).

155. The kit of any one of claims 150-153, wherein said one or more guide RNAs are a dual-guide RNA.

156. The kit of any one of claims 139-155, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

157. The kit of any one of claims 139-155, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

158. The kit of claim 157, wherein cleavage by said RGN polypeptide generates a double-stranded break.

159. The kit of claim 157, wherein cleavage by said RGN polypeptide generates a single-stranded break.

160. The kit of any one of claims 157-159, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

161. A method of detecting a target DNA sequence of a DNA molecule in a sample, the method comprising:

a) contacting the sample with:

i) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109, wherein said RGN polypeptide is capable of binding and cleaving said target DNA sequence of a DNA molecule in an RNA-guided sequence specific manner when bound to a guide RNA capable of hybridizing to said target DNA sequence;

ii) said guide RNA; and

iii) a detector single-stranded DNA (ssDNA) that does not hybridize with the guide RNA; and

b) measuring a detectable signal produced by cleavage of the detector ssDNA by the RGN, thereby detecting the target DNA sequence.

162. The method of claim 161, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

163. The method of claim 161, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

164. The method of any one of claims 161-163, wherein said sample comprises DNA molecules from a cell lysate.

165. The method of any one of claims 161-163, wherein said sample comprises cells.

166. The method of claim 165, wherein said cells are eukaryotic cells.

167. The method of any one of claims 161-163, wherein the DNA molecule comprising the target DNA sequence is produced by reverse-transcription of an RNA template molecule present in a sample comprising RNA.

168. The method of claim 167, wherein the RNA template molecule is an RNA virus.

169. The method of claim 168, wherein the RNA virus is a coronavirus.

170. The method of claim 169, wherein the coronavirus is a bat SARS-like coronavirus, SARS-CoV, or SARS-CoV-2.

171. The method of any one of claims 167-170, wherein the samples comprising RNA is derived from a sample comprising cells.

172. The method of any one of claims 161-171, wherein said detector ssDNA comprises a fluorophore/quencher pair.

173. The method of any one of claims 161-171, wherein said detector ssDNA comprises a fluorescence resonance energy transfer (FRET) pair.

174. The method of any one of claims 161-173, wherein said method further comprises amplifying nucleic acids in the sample prior to or together with the contacting of step a).

175. A method of cleaving single-stranded DNAs (ssDNAs), the method comprising contacting a population of nucleic acids, wherein said population comprises a DNA molecule comprising a target DNA sequence and a plurality of non-target ssDNAs with:

a) an RNA-guided nuclease (RGN) polypeptide comprising an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109, wherein said RGN polypeptide is

capable of binding and cleaving said target DNA sequence in an RNA-guided sequence specific manner when bound to a guide RNA capable of hybridizing to said target DNA sequence; and

b) said guide RNA;

wherein the RGN polypeptide cleaves non-target ssDNAs of said plurality.

176. The method of claim 175, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

177. The method of claim 175, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 2, 4, 1, 3, and 5 to 109.

178. The method of any one of claims 175-177, wherein said population of nucleic acids are within a cell lysate.

179. The method of any one of claims 175-178, wherein the DNA molecule comprising the target DNA sequence is produced by reverse-transcription of an RNA template molecule.

180. The method of any one of claims 161-179, wherein said RGN polypeptide and said guide RNA are not found complexed to one another in nature.

181. The method of any one of claims 161-180, wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 11 and said guide RNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 116.

182. The method of any one of claims 161-180, wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 11 and said guide RNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 116.

183. The method of any one of claims 161-180, wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 11 and said guide RNA comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 116.

184. The method of any one of claims 161-180, wherein said guide RNA comprises a tracrRNA.

185. The method of claim 184, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 121, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 123, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 120, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 122, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 124, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 125, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 126, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 127, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 90% sequence identity to SEQ ID NO: 128, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 90% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 16.

186. The method of claim 184, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 121, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 123, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 120, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 122, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 124, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 125, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 126, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 127, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having at least 95% sequence identity to SEQ ID NO: 128, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having at least 95% sequence identity to SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO: 16.

187. The method of claim 184, wherein said tracrRNA is selected from the group consisting of:

a) a tracrRNA having 100% sequence identity to SEQ ID NO: 121, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 111, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 2;

b) a tracrRNA having 100% sequence identity to SEQ ID NO: 123, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 113, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 4;

c) a tracrRNA having 100% sequence identity to SEQ ID NO: 120, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 110, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 1;

d) a tracrRNA having 100% sequence identity to SEQ ID NO: 122, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 112, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 3;

e) a tracrRNA having 100% sequence identity to SEQ ID NO: 124, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 114, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 5;

f) a tracrRNA having 100% sequence identity to SEQ ID NO: 125, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 115, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 6;

g) a tracrRNA having 100% sequence identity to SEQ ID NO: 126, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 117, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 12;

h) a tracrRNA having 100% sequence identity to SEQ ID NO: 127, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to SEQ ID NO: 118, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 13; and

i) a tracrRNA having 100% sequence identity to SEQ ID NO: 128, wherein said guide RNA further comprises a CRISPR RNA comprising a CRISPR repeat sequence having 100% sequence identity to

SEQ ID NO: 119, and wherein said RGN polypeptide comprises an amino acid sequence having 100% sequence identity to SEQ ID NO: 16.

188. The method of any one of claims 184-187, wherein said guide RNA is a single guide RNA (sgRNA).

189. The method of any one of claims 184-187, wherein said guide RNA is a dual-guide RNA.

190. The method of any one of claims 161-189, wherein said target DNA sequence is within a region of said DNA molecule that is single-stranded.

191. The method of any one of claims 161-189, wherein said target DNA sequence is within a region of said DNA molecule that is double-stranded.

192. The method of claim 191, wherein cleavage of said target DNA sequence by said RGN polypeptide generates a double-stranded break.

193. The method of claim 191, wherein cleavage of said target DNA sequence by said RGN polypeptide generates a single-stranded break.

194. The method of any one of claims 191-193, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

195. A method of treating a disease, said method comprising administering to a subject in need thereof an effective amount of a pharmaceutical composition of embodiment 82 or 138.

196. The method of embodiment 195, wherein said disease is associated with a causal mutation and said effective amount of said pharmaceutical composition corrects said causal mutation.

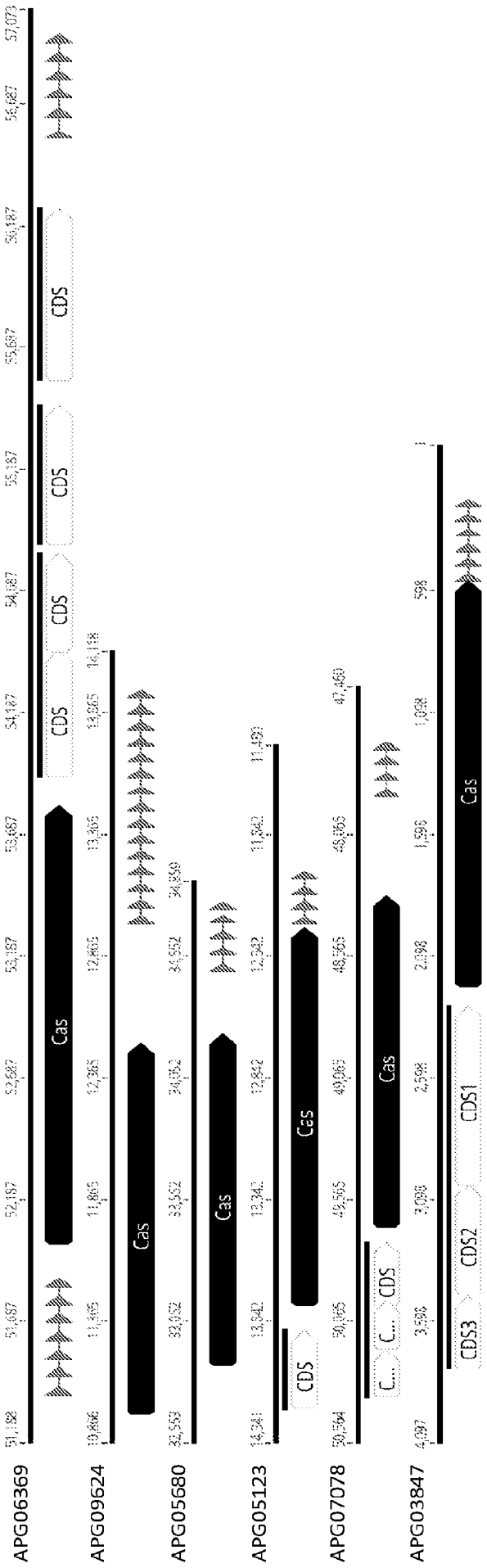

APG06369
APG09624
APG05680
APG05123
APG07078
APG03847
Cas
CDS
CDS1
CDS2
CDS3

FIGURE 1

SEQUENCE LISTING

<110> LIFEEDIT THERAPEUTICS, INC.

<120> RNA-GUIDED NUCLEASES AND ACTIVE FRAGMENTS AND VARIANTS THEREOF AND METHODS OF USE

<130> L103438 1180WO (0077.8)

<140>
<141>

<150> 63/058,169
<151> 2020-07-29

<150> 62/955,014
<151> 2019-12-30

<160> 314

<170> PatentIn version 3.5

<210> 1
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 1

```
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1               5                  10                  15


Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
            20                  25                  30


Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
        35                  40                  45


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


Thr Glu Lys Leu Lys His Ser Asp Ala Glu Tyr Gln Lys His Leu Thr
65                  70                  75                  80


Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Ala
                85                  90                  95


Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
            100                 105                 110
```

```
Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
        115                 120                 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
    130                 135                 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

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

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320
```

```
Ile Asp Asp Phe Leu Arg Val Arg Thr Gln Leu Gln Ser Gln Arg Lys
                325                 330                 335

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445

Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460

Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495

Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505
```

<210> 2
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 2

```
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1               5                  10                  15

Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
            20                  25                  30

Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
        35                  40                  45

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

Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
65                  70                  75                  80

Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
                85                  90                  95

Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
            100                 105                 110

Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
        115                 120                 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
    130                 135                 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

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

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220
```

```
Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Asn Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Asn Val Arg Thr Gln Leu Gln Asn Gln Arg Lys
                325                 330                 335

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430
```

```
Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445


Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460


Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480


Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495


Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505
```

<210> 3
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 3

```
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1               5                   10                  15


Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
            20                  25                  30


Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
        35                  40                  45


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


Thr Glu Lys Leu Lys His Ser Asp Ala Glu Tyr Gln Lys His Leu Thr
65                  70                  75                  80


Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Ala
                85                  90                  95


Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
            100                 105                 110
```

```
Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
        115                 120                 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
    130                 135                 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

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

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Asn Val Arg Thr Gln Leu Gln Asn Gln Arg Lys
                325                 330                 335
```

```
Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445

Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460

Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Asn Glu Met
                485                 490                 495

Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505
```

<210> 4
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 4

```
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1               5                   10                  15
```

```
Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
            20                  25                  30

Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
        35                  40                  45

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

Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
65                  70                  75                  80

Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
                85                  90                  95

Lys Glu Lys Val Asn Asp Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
            100                 105                 110

Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
        115                 120                 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
    130                 135                 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

Ile Ser His Phe Asn Asn Asp Met Lys Thr Gly Val Leu Arg Gly Asp
            180                 185                 190

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220
```

```
Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Lys Val Arg Thr Gln Leu Gln Ser Gln Arg Lys
                325                 330                 335

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Asn
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445
```

```
Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460

Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495

Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505
```

<210> 5
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 5

```
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1               5                   10                  15

Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
            20                  25                  30

Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
        35                  40                  45

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

Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
65                  70                  75                  80

Lys Tyr Arg Glu Val Ser Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
                85                  90                  95

Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
            100                 105                 110

Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
        115                 120                 125
```

```
Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
    130                 135                 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

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

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Asn Val Arg Thr Gln Leu Gln Asn Gln Arg Lys
                325                 330                 335
```

```
Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445

Tyr Glu Lys Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460

Lys Glu Cys Lys Gly Tyr Ala His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495

Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505
```

<210> 6
<211> 450
<212> PRT
<213> Bacillus sp.

<400> 6

```
Met Ser Ile Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1               5                   10                  15
```

```
Asn Val Glu Trp Lys Ala Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr
            20                  25                  30

Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe
        35                  40                  45

Asp Asn Leu Ser Leu Asn His Phe Lys Glu Thr Gly Glu Tyr Pro Ser
    50                  55                  60

Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65                  70                  75                  80

Tyr Asn Gln Leu Arg Glu Glu Tyr Gln Asp Ile Asn Lys Ala Asn Met
                85                  90                  95

Ser Thr Thr Ile Gln Lys Ser Leu Lys Asn Trp Asn Ser Arg Lys Lys
            100                 105                 110

Glu Ile Trp Arg Gly Glu Met Ser Ile Pro Ser Phe Arg Ser Asp Leu
        115                 120                 125

Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Ile Lys Glu Lys Ser
    130                 135                 140

Gly Asp Tyr Phe Ala Asn Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
145                 150                 155                 160

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

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr
            180                 185                 190

Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Asn Lys Trp Tyr
        195                 200                 205

Leu Ser Ile Thr Tyr Lys Ala Thr Lys Lys Glu Gly His Lys Phe Asp
    210                 215                 220

Glu Glu Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
225                 230                 235                 240
```

```
Phe Ala Tyr Asn Lys Gly Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
               245                 250                 255

Ile Glu Ala Phe Arg Lys Lys Ile Glu Tyr Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Glu Asn Arg Ile Gly Lys Gly Arg Lys
        275                 280                 285

Lys Arg Ile Lys Pro Ile Glu Val Leu Asn Asp Lys Ile Thr Lys Phe
    290                 295                 300

Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320

Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Gln Gly Ile
               325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350

Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Ile Lys Val Val Lys
        355                 360                 365

Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
    370                 375                 380

His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln His Cys
385                 390                 395                 400

Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Ile
               405                 410                 415

Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
            420                 425                 430

Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Glu Asn
        435                 440                 445
```

Ile Asn
450

<210> 7
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 7
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1 5 10 15

Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
20 25 30

Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
35 40 45

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

Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
65 70 75 80

Lys Tyr Arg Glu Val Ser Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
85 90 95

Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
100 105 110

Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
115 120 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
130 135 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145 150 155 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
165 170 175

```
Leu Ser His Phe Asn Asn Asp Met Lys Thr Gly Val Leu Arg Gly Asp
            180                 185                 190

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Asn Val Arg Thr Gln Leu Gln Asn Gln Arg Lys
                325                 330                 335

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400
```

Lys Leu Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
405 410 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
420 425 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
435 440 445

Tyr Glu Lys Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
450 455 460

Lys Glu Cys Lys Gly Tyr Ala His Lys Val Asn Ala Asp Phe Asn Ala
465 470 475 480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
485 490 495

Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
500 505

<210> 8
<211> 450
<212> PRT
<213> Bacillus sp.

<400> 8
Met Ser Leu Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1 5 10 15

Asn Val Glu Trp Lys Val Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr
20 25 30

Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe
35 40 45

Asp Asn Leu Ser Leu Lys His Phe Lys Glu Thr Gly Glu Tyr Pro Ser
50 55 60

Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65 70 75 80

```
Tyr Asp Gln Leu Lys Glu Glu Tyr Lys Asp Ile Asn Lys Ala Asn Met
                85                  90                  95

Ser Thr Thr Ile Gln Lys Thr Leu Lys Asn Trp Asn Ser Arg Lys Lys
            100                 105                 110

Glu Ile Trp Arg Gly Glu Met Ser Ile Pro Ser Phe Arg Ser Asp Leu
        115                 120                 125

Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Ile Lys Glu Lys Gly
    130                 135                 140

Gly Asp Tyr Ile Ala Ser Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
145                 150                 155                 160

Glu Asn Asp Leu Pro Asn Gly Lys Ile Leu Met Lys Leu Ser Thr Arg
                165                 170                 175

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr
            180                 185                 190

Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Lys Lys Trp Tyr
        195                 200                 205

Leu Ser Ile Thr Tyr Lys Ser Asn Ile Lys Glu Glu Leu Lys Phe Asp
    210                 215                 220

Glu Asp Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
225                 230                 235                 240

Phe Ala Phe Asn Lys Ser Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
                245                 250                 255

Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Gly Asn Arg Ile Gly Lys Gly Arg Glu
        275                 280                 285
```

```
Lys Arg Ile Lys Pro Ile Asp Val Leu Asn Asp Lys Val Ala Lys Phe
    290                 295                 300

Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320

Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Lys Gly Ile
                325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350

Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Val Lys Val Val Lys
        355                 360                 365

Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
    370                 375                 380

His Lys Asp Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Leu Gln Cys
385                 390                 395                 400

Ser Phe Lys Ala His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Val
                405                 410                 415

Tyr Thr Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
            420                 425                 430

Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Ile Lys Asn
        435                 440                 445

Ile Asn
    450


<210> 9
<211> 486
<212> PRT
<213> Bacillus sp.

<400> 9
Met Ile Thr Ser Arg Lys Ile Arg Leu Ser Ile Val Ser Asp Asn Ala
1               5                   10                  15
```

```
Thr Glu Ala Tyr Asn Phe Ile Arg Lys Glu Met Arg Glu Gln Asn Lys
            20                  25                  30

Ala Leu Asn Val Ala Leu Asn His Leu Tyr Phe Asn Ser Ile Ala Arg
        35                  40                  45

Gln Lys Ile Leu Leu Ala Asp Glu Ala Tyr Gln Gln Lys Leu Lys Asp
    50                  55                  60

Ala Ile Asn Ser Gln Glu Lys Ser Phe Asn Thr Leu Lys Glu Leu Glu
65                  70                  75                  80

Arg Lys Gln Cys Ile Glu Glu Asp Phe Glu Lys Lys Gln Thr Leu Lys
                85                  90                  95

Glu Arg Leu Ile Lys Ala Lys Asn Ala Tyr Glu Lys Ala Lys Glu Lys
            100                 105                 110

Val Ser His Leu Arg Lys Ser Arg Ser Lys Asp Ser Phe Gln Glu Tyr
        115                 120                 125

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

Ser Ser Gln Phe Asn Leu His Ser Asp Thr Lys Asp Arg Leu Thr Met
145                 150                 155                 160

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

Asn Cys Ser Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu Tyr Ile Arg
            180                 185                 190

Gly Arg Asn Thr Val Leu Tyr Lys Glu Gly Asn Glu Phe Phe Ile Lys
        195                 200                 205

Trp Ile Lys Gly Ile Val Phe Lys Cys Ile Leu Gly Val Lys Asn Gln
    210                 215                 220

Asn Lys Thr Glu Leu Tyr Lys Thr Leu Glu Cys Val Leu Ala Gly Asn
225                 230                 235                 240
```

```
Tyr Lys Ile Cys Asp Ser Ser Met Asn Phe Asn Gln Gln Asn Lys Leu
                245                 250                 255

Ile Met Asn Leu Thr Leu Asn Met Pro Asp Lys Asp Glu Asn Arg Lys
            260                 265                 270

Val Pro Gly Arg Ile Ala Gly Ile Asp Leu Gly Leu Lys Ile Pro Ala
        275                 280                 285

Tyr Phe Ala Val Asn Asp Ala Pro Tyr Ile Arg Lys Ala Leu Gly Lys
    290                 295                 300

Ile Glu Asp Phe Leu Lys Val Arg Thr Ser Ile Gln Ser Gln Lys Arg
305                 310                 315                 320

Ser Leu Glu Arg Ala Leu Gln Ser Ser Lys Gly Gly Lys Gly Arg Lys
                325                 330                 335

Lys Lys Leu Arg Ala Leu Asp Gln Phe Lys Gly Lys Glu Lys Arg Tyr
            340                 345                 350

Val Thr Thr Tyr Asn His Phe Ile Ser Lys Lys Ile Ile Ser Leu Ala
        355                 360                 365

Ile Gln Tyr Gly Val Glu Gln Ile Asn Leu Glu Leu Leu Thr Leu Lys
    370                 375                 380

Glu Thr Gln Lys Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
385                 390                 395                 400

Gln Gln Phe Ile Glu Tyr Lys Ala Lys Arg Glu Gly Ile Leu Ile Lys
                405                 410                 415

Tyr Val Asp Pro Phe Asn Thr Ser Gln Thr Cys Ser Lys Cys Asp His
            420                 425                 430

Tyr Glu Gly Gly Gln Arg Glu Lys Gln Ala Asn Phe Leu Cys Lys Ser
        435                 440                 445
```

```
Cys Gly Phe Glu Glu Asn Ala Asp Phe Asn Ala Ala Arg Asn Ile Ala
    450                 455                 460

Lys Ser Lys Asn Tyr Ile Thr Arg Lys Glu Glu Ser Glu Tyr Tyr Lys
465                 470                 475                 480

Arg Asn Asn Glu Ile Ala
                485
```

<210> 10
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 10

```
Met Ser Val Met Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Thr
1               5                   10                  15

Asn Val Asp Trp Ser Thr Phe Glu Lys Asn Leu Arg Asp Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Tyr Asp Tyr Phe Lys Glu Thr Gly Thr Ser Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Lys Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Val Leu Gln Ser Lys Tyr Pro Asp Val His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Ser
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Gln Leu
        115                 120                 125

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

```
Gly Asp Tyr Ile Ala Gly Leu Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Lys Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Ser Gly
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ser Ile Thr Glu Asn Lys Phe
    210                 215                 220

Asp Glu Asn Leu Ile Met Gly Ile Asp Met Gly Gly Val Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Val Arg Ser Asn Ile Arg Ser Asp
                245                 250                 255

Glu Ile Arg Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Asn Gln Ser Lys Tyr Cys Ser Asp Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Ala Lys Arg Leu Gln Pro Ile Asp Val Ile Ser Asn Lys Ile Ala Gln
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Phe Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Ile Glu His Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asp Asp Lys Val Leu Lys Asp Trp Thr Tyr Tyr Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Lys Gln Ala Gln Ala Tyr Gly Ile Glu Val Ile
        355                 360                 365
```

```
Thr Ile Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Ser Asn Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Glu Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Arg Leu His Ser Lys Lys Tyr Met Glu Asp His Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Ser Gly
    450
```

<210> 11
<211> 599
<212> PRT
<213> Gordonia sp.

<400> 11

```
Met Ala Ile Thr Val His Thr Met Gly Ala His Tyr Arg Trp Glu Ile
1               5                   10                  15

Pro Glu Gln Leu Arg Ala Gln Leu Trp Leu Ala His Asn Leu Arg Glu
            20                  25                  30

Asp Leu Val Thr Leu Gln His Glu Tyr Glu Ala Arg Thr Lys Glu Val
        35                  40                  45

Trp Ser Ser Phe Pro Asp Val Ala Ala Thr Glu Asp Arg Leu Thr Ala
    50                  55                  60

Ala Glu Leu Glu Ala Glu Ala Leu Ala Glu Lys Val Lys Ser Glu Arg
65                  70                  75                  80

Ile Arg Gln Arg Thr Lys Arg Val Thr Gly Pro Val Ala Asp Gln Leu
                85                  90                  95
```

```
Ala Ala Ala Arg Lys Thr Ala Lys Glu Ala Arg Ile Ala Arg Arg Ala
            100                 105                 110

Ala Ile Ala Ala Val Arg Asp Thr Ala Lys Ala Gln Leu Asn Glu Leu
        115                 120                 125

Ser Asp Glu Leu Lys Ala Ala His Lys Arg Leu Tyr Ala Glu Tyr Cys
    130                 135                 140

Gln Ser Gly His Leu Tyr Trp Ala Thr Phe Asn Ala Thr Leu Asp His
145                 150                 155                 160

His Lys Thr Ala Val Lys Arg Met Gln Gln Leu Arg Ala Ala Gly Arg
                165                 170                 175

Pro Ala Gln Met Arg His His Arg Phe Asp Gly Thr Gly Thr Ile Ala
            180                 185                 190

Val Gln Leu Gln Arg Gln Ser Gly Gln Pro Gln Arg Thr Pro Ala Leu
        195                 200                 205

Ile Gly Asp Pro Glu Gly Lys Tyr Arg Asn Val Phe Tyr Ser Pro Trp
    210                 215                 220

Val Asp Pro Asp Gln Trp Asp Thr Met Thr Arg Ser Glu Gln Arg Lys
225                 230                 235                 240

Ala Gly Arg Val Thr Val Thr Met Arg Cys Gly Ser Thr Asp Asp Gly
                245                 250                 255

Pro Ala Trp Ile Ala Val Pro Val Gln Gln His Arg Met Leu Pro Ala
            260                 265                 270

Asp Ala Asp Ile Thr Gly Ala Gln Leu Thr Val Thr Arg Lys Gly Ser
        275                 280                 285

Ala Thr His Val Lys Ile Ala Val Thr Ala Arg Val Gly Glu Pro Glu
    290                 295                 300
```

```
Pro Ile Thr Asp Gly Pro Thr Val Ala Val His Leu Gly Trp Arg Asp
305                 310                 315                 320

Thr Asp Thr Gly Thr Val Val Ala His Trp Arg Ala Ser Glu Pro Leu
                325                 330                 335

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

Arg Thr Gly Thr Val Val Ile Pro Thr Arg Ile Val Asp Arg Ile Glu
        355                 360                 365

Ser Ala Ala Gly Ile Ala Ser Gln Arg Gly Asp Ala Gln Asn Glu Met
    370                 375                 380

Lys Arg Gln Leu Val Glu Trp Leu Thr Glu Val Gly Pro Gln Pro His
385                 390                 395                 400

Pro Thr Arg Asp Gly Glu Glu Ile Ser Ala Ala Asp Ala Ala Arg Trp
                405                 410                 415

Arg Asn Pro Gly Arg Phe Ala Ala Leu Ala Met Ala Trp Arg Asp Asn
            420                 425                 430

Pro Pro Asn Arg Gly Glu Gln Ile Ala Tyr Val Leu Glu Ser Trp Arg
        435                 440                 445

Ala Thr Asp Lys Gly Leu Trp Asn Arg Gln Glu Gly Gly Arg Arg Lys
    450                 455                 460

Ala Leu Gly His Arg Thr Asp Ile Tyr Arg Gln Val Ala Ala Leu Ile
465                 470                 475                 480

Ala Asp Gln Ala Gly Val Val Val Val Asp Asp Thr Ser Ile Ala Asp
                485                 490                 495

Ile Ala Ala Arg Pro Thr Glu Leu Pro Asn Glu Val Glu Ala Arg Ile
            500                 505                 510

Ala Arg Arg Arg Gly Ser Ala Ala Pro Gly Glu Leu Arg Glu Ala Val
        515                 520                 525
```

Arg Ser Ala Ala Val Arg Asp Gly Val Gly Val Glu Val Val Ala His
530 535 540

Lys Gly Leu Ser Arg Thr His Ala Ala Cys Gly His Glu Asn Pro Ala
545 550 555 560

Asp Asp Arg Tyr Met Thr Val Gly Val Val Cys Asp Gly Cys Gly Arg
565 570 575

Thr Tyr Asp Gln Asp Leu Ser Ala Thr Leu Leu Met Leu Gln Arg Ala
580 585 590

Thr Ser Thr Ala Ala Ala Thr
595

<210> 12
<211> 553
<212> PRT
<213> Micrococcus sp.

<400> 12
Met Thr Thr Thr Thr Gly Thr Asp Leu Val Pro Arg Leu Arg Ala Phe
1 5 10 15

Lys His Arg Leu Asp Pro Asn Pro Ala Gln Ala Thr Leu Leu Ala Gln
20 25 30

Tyr Ala Gly Ala Ala Arg Val Ala Tyr Asn Met Leu Thr Ala His Asn
35 40 45

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

Thr Gly Leu Ala Gly Pro Glu Leu Ala Ala Arg Met Lys Ala Glu Arg
65 70 75 80

Ala Ala Asp Pro Thr Leu Arg Val Ala Ser Tyr Gln Ser Tyr Ser Thr
85 90 95

Thr His Leu Thr Pro Leu Ile Arg Arg His Arg Glu Ala Ala Ala Ala
100 105 110

```
Ile Ala Ala Gly Ala Asp Pro Ala Glu Ala Trp Thr Asp Glu Arg Tyr
        115                 120                 125

Ala Glu Pro Trp Met His Thr Val Pro Arg Arg Val Leu Val Ser Gly
    130                 135                 140

Leu Gln Asn Ala Ala Lys Ala Thr Glu Asn Trp Met Ala Ser Ala Ser
145                 150                 155                 160

Gly Thr Arg Ala Gly Ala Arg Val Gly Leu Pro Arg Phe Lys Lys Lys
                165                 170                 175

Gly Arg Ser Arg Asp Ser Phe Thr Ile Pro Ala Pro Glu Val Ile Gly
            180                 185                 190

Ala Ala Gly Thr Pro Tyr Lys Arg Gly Glu Pro Arg Arg Gly Val Ile
        195                 200                 205

Thr Asp His Arg His Leu Arg Leu Ala Ser Leu Gly Thr Ile Arg Thr
    210                 215                 220

Tyr Asp Lys Thr Ser Arg Leu Val Arg Ala Cys Arg Arg Gly Ala Gln
225                 230                 235                 240

Ile Arg Ser Met Thr Ile Ser Gln Ala Gly Gly Arg Trp Tyr Ala Ser
                245                 250                 255

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

Gln Arg Ala Asn Asp Ala Val Gly Val Asp Leu Gly Val Lys His Leu
        275                 280                 285

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

Arg Gln Ala Ala Arg Leu Thr Arg Leu Gln Arg Ala Tyr Ala Arg Thr
305                 310                 315                 320
```

```
Gln Pro Gly Ser Asn Arg Arg Glu Arg Val Arg Arg Gln Ile Ala Ala
                325                 330                 335

Leu His His Gly Ile Ala Leu Arg Arg Ala Gly Leu Leu His Gln Val
            340                 345                 350

Ser Thr Arg Leu Ala Met Asp Phe Ala Val Val Ala Leu Glu Asp Leu
        355                 360                 365

Asn Val Ala Gly Met Thr Arg Ser Ala Arg Gly Thr Leu Glu Ala Pro
    370                 375                 380

Gly Arg Asn Val Ala Ala Lys Ser Gly Leu Asn Arg Ala Ile Leu Asp
385                 390                 395                 400

Ala Gly Leu Gly Met Leu Arg Arg Gln Leu Asp Tyr Lys Thr Ser Trp
                405                 410                 415

Ala Gly Ser Gln Val Lys Met Ile Asp Arg Phe Ala Pro Ser Ser Lys
            420                 425                 430

Ala Cys Ser Arg Cys Gly Thr Val Lys Ser Thr Leu Ser Leu Ala Glu
        435                 440                 445

Arg Thr Phe Glu Cys Glu Ala Cys His Leu Val Ile Asp Arg Asp Val
    450                 455                 460

Asn Ala Ala Ile Asn Ile Arg Ala Trp Ala Val Gln Glu Glu Arg Gly
465                 470                 475                 480

Ala Gly Val Gly Leu Ala Arg Gly Arg Arg Glu Ser Arg Asn Gly Arg
                485                 490                 495

Gly Ala Ala Val Ser Gly Pro Pro Ser Gly Gly Ala Ala Gly Gln Gly
            500                 505                 510

Arg Gly Ser Val Lys Pro Ala Pro Gln Gly Val Gly Met Ser Ser Arg
        515                 520                 525

Ala Thr Gly Trp Ser Ser Gln Pro Pro Ser Thr Glu Gly Glu Ser Ala
    530                 535                 540
```

Glu Arg Gly Ala Ser Ala Leu Ala Arg
545 550

<210> 13
<211> 553
<212> PRT
<213> Micrococcus sp.

<400> 13
Met Thr Thr Thr Thr Gly Thr Asp Leu Ala Pro Arg Leu Arg Ala Phe
1 5 10 15

Lys His Arg Leu Asp Pro Asn Pro Ala Gln Ala Thr Leu Leu Ala Gln
20 25 30

Tyr Ala Gly Ala Ala Arg Val Ala Tyr Asn Met Leu Thr Ala His Asn
35 40 45

Arg Ala Ala Leu Ala Ala Gly Ala Ala Arg Arg Thr Glu Leu Ala Glu
50 55 60

Thr Gly Leu Ala Gly Pro Glu Leu Ala Ala Arg Met Lys Ala Glu Arg
65 70 75 80

Ala Ala Asp Pro Thr Leu Arg Val Ala Ser Tyr Gln Ser Tyr Ser Thr
85 90 95

Ala His Leu Thr Pro Leu Ile Arg Arg His Arg Glu Ala Ala Ala Ala
100 105 110

Ile Ala Ala Gly Ala Asp Pro Ala Glu Ala Trp Thr Asp Glu Arg Tyr
115 120 125

Ala Glu Pro Trp Met His Thr Val Pro Arg Arg Val Leu Val Ser Gly
130 135 140

Leu Gln Asn Ala Ala Lys Ala Thr Glu Asn Trp Met Ala Ser Ala Ser
145 150 155 160

Gly Thr Arg Ala Gly Ala Arg Val Gly Leu Pro Arg Phe Lys Lys Lys
165 170 175

```
Gly Arg Ser Arg Asp Ser Phe Thr Ile Pro Ala Pro Glu Val Ile Gly
            180                 185                 190

Ala Ala Gly Thr Pro Tyr Lys Arg Gly Glu Pro Arg Arg Gly Val Ile
        195                 200                 205

Thr Asp His Arg His Leu Arg Leu Ala Ser Leu Gly Thr Ile Arg Thr
    210                 215                 220

Tyr Asp Lys Thr Ser Arg Leu Val Arg Ala Cys Arg Arg Gly Ala Gln
225                 230                 235                 240

Ile Arg Ser Met Thr Ile Ser Gln Ala Gly Gly Arg Trp Tyr Ala Ser
                245                 250                 255

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

Gln Arg Ala Asn Gly Ala Val Gly Val Asp Leu Gly Val Lys His Leu
        275                 280                 285

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

Arg Gln Ala Ala Arg Leu Thr Arg Leu Gln Arg Ala His Ala Arg Thr
305                 310                 315                 320

Gln Pro Gly Ser Asn Arg Arg Glu Arg Val Arg Arg Gln Ile Ala Ala
                325                 330                 335

Leu Gln His Gly Ile Ala Leu Arg Arg Ala Gly Leu Leu His Gln Val
            340                 345                 350

Ser Thr Arg Leu Ala Thr Asp Phe Ala Val Val Ala Leu Glu Asp Leu
        355                 360                 365

Asn Val Ala Gly Met Thr Arg Ser Ala Arg Gly Thr Leu Glu Ala Pro
    370                 375                 380
```

```
Gly Arg Asn Val Ala Ala Lys Ser Gly Leu Asn Arg Ala Ile Leu Asp
385                 390                 395                 400

Ala Gly Leu Gly Met Leu Arg Arg Gln Leu Asp Tyr Lys Thr Ser Trp
                405                 410                 415

Ala Gly Ser Gln Val Lys Met Ile Asp Arg Phe Ala Pro Ser Ser Lys
            420                 425                 430

Ala Cys Ser Arg Cys Gly Thr Val Lys Ser Thr Leu Ser Leu Ala Glu
        435                 440                 445

Arg Thr Phe Glu Cys Glu Ala Cys His Leu Val Ile Asp Arg Asp Val
    450                 455                 460

Asn Ala Ala Ile Asn Ile Arg Ala Trp Ala Val Gln Glu Glu Arg Gly
465                 470                 475                 480

Ala Gly Val Glu Leu Ala Arg Gly Arg Arg Glu Ser Arg Asn Gly Arg
                485                 490                 495

Gly Ala Ala Val Ser Gly Pro Pro Ser Gly Gly Ala Ala Arg Gln Gly
            500                 505                 510

Arg Gly Ser Val Lys Pro Ala Pro Gln Gly Val Gly Met Ser Ser Arg
        515                 520                 525

Ala Thr Gly Trp Ser Ser Gln Pro Pro Ser Thr Glu Gly Glu Ser Ala
    530                 535                 540

Glu Arg Gly Ala Ser Ala Leu Ala Arg
545                 550
```

<210> 14
<211> 505
<212> PRT
<213> Paeniglutamicibacter sp.

<400> 14

```
Met Ser Val Gln Ala Ala Met Ala Thr Thr Val Ile His Arg Ala Tyr
1               5                   10                  15
```

```
Arg Leu Thr Leu Asp Pro Thr Pro Gln Gln Ala Gln Lys Leu Ser Gln
            20                  25                  30

Trp Ala Gly Ala Ala Arg Ala Met Tyr Asn His Ala Ile Ala Ala Lys
        35                  40                  45

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

Glu Gln Glu Leu Thr Glu Glu Gln Ala Arg Lys Ser Ile Lys Val Pro
65                  70                  75                  80

Ile Pro Thr Ala Tyr Gly Phe Asn Ala Trp Leu Thr Glu Thr Arg Asn
                85                  90                  95

Thr His His Asp Ala Ala Glu Lys Gly Leu Leu Leu Pro Gly Arg Asp
            100                 105                 110

Gly Arg Glu His Glu Pro Trp Leu His Ala Val Asn Arg Ser Ala Leu
        115                 120                 125

Met Gly Ala Met Arg His Ala Asp Asp Ala Trp Thr Asn Trp Ile Asp
    130                 135                 140

Ser Leu Thr Gly Thr Arg Ala Gly Arg Lys Ile Gly Tyr Pro Arg Phe
145                 150                 155                 160

Lys Lys Arg Gly Val Ala Arg Asp Ser Phe Thr Ile Thr His Asp Arg
                165                 170                 175

Lys Ser Pro Gly Ile Arg Leu Ala Thr Thr Arg Arg Leu Arg Ile Pro
            180                 185                 190

Thr Phe Gly Glu Ile Arg Ile His Asp His Ala Lys Arg Leu His Arg
        195                 200                 205

Lys Leu His Thr Gly Thr Val Glu Val Thr Ser Val Thr Val Ser Arg
    210                 215                 220

His Gly Pro Arg Trp Tyr Ala Ser Leu Thr Val Glu Glu Thr Ile Pro
225                 230                 235                 240
```

```
Thr Pro Arg Leu Ser Lys Arg Lys Arg Ala Ala Gly Ile Ile Gly Val
                245                 250                 255

Asp Leu Gly Val Lys Ile Thr Ala Ala Leu Ser Asn Gly Asp Leu Ile
            260                 265                 270

Pro Asn Pro Arg Val Lys Ala Ser His Ala Lys Lys Leu Ala Arg Leu
        275                 280                 285

Gln Lys Ala Leu Ala Lys Ser Gln Lys Gly Ser Arg Asn Arg Ala Gln
    290                 295                 300

Leu Val Gln Lys Ile Gly Cys Leu Thr His Leu Glu Ala Arg Gln Arg
305                 310                 315                 320

Glu Gly His Ala His Asn Leu Ala Asn Arg Leu Val His Thr Trp Ala
                325                 330                 335

Ile Ile Gly Ile Glu Asp Leu Asn Val Ala Gly Met Thr Arg Ser Ser
            340                 345                 350

Arg Gly Thr Ile Glu Lys Pro Gly Lys Asn Val Arg Ala Lys Ala Gly
        355                 360                 365

Leu Asn Arg Ser Ile Leu Asp Val Ala Pro Ala Gln Ile Arg His Leu
    370                 375                 380

Leu Asp Tyr Lys Thr Ala Trp Ser Gly Thr Gln Leu Val Val Ile Asp
385                 390                 395                 400

Arg Trp Ala Pro Thr Ser Lys Lys Cys Ser Thr Cys Gly Ala Val Lys
                405                 410                 415

Ala Lys Leu Thr Leu Ala Glu Arg Thr Phe Glu Cys Glu Ala Cys Gly
            420                 425                 430

Leu Val Leu Asp Arg Asp Ile Asn Ala Ala Arg Asn Ile Ala Ala Leu
        435                 440                 445
```

```
Ala Ala Val Ala Pro Ser Thr Glu Glu Thr Gln Asn Ala Arg Arg Ala
    450                 455                 460

Ala Pro Arg Lys Pro Val Pro Ser Thr Val Lys Gln Arg Ala Ala Met
465                 470                 475                 480

Lys Arg Glu Asp Pro Pro Gly Ser Ser Pro Pro Ser Asn Gly Arg Thr
                485                 490                 495

Phe His Thr Val Leu Val Gln Val Ser
            500                 505
```

<210> 15
<211> 515
<212> PRT
<213> Streptomyces sp.

<400> 15

```
Met Glu Arg Glu Val Leu Arg Ala Phe Lys Phe Ala Leu Asp Pro Thr
1               5                   10                  15

Pro Ala Gln Ala Glu Ala Leu Ala Arg His Ala Gly Ala Ala Arg Trp
            20                  25                  30

Ala Phe Asn Tyr Ala Leu Ala Val Lys Val Ser Ala His Gln Arg Trp
        35                  40                  45

Arg Ala Glu Val Ala Gly Leu Val Ala Gln Gly Ile Glu Glu Ala Glu
    50                  55                  60

Ala Arg Arg Arg Val Lys Val Pro Val Pro Ser Lys Pro Gln Ile Gln
65                  70                  75                  80

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

Pro Glu Gly Thr Phe Gly Pro Glu Arg Pro Cys Pro Trp Trp Tyr Glu
            100                 105                 110

Val Asn Thr Tyr Ala Phe Gln Ser Ala Phe Ile Asp Ala Asp Arg Ala
        115                 120                 125
```

```
Trp Lys Asn Trp Leu Asp Ser Leu Arg Gly Val Arg Ala Gly Arg Lys
    130                 135                 140

Val Gly Tyr Pro Arg Phe Lys Lys Lys Gly Arg Ser Arg Asp Ser Phe
145                 150                 155                 160

Arg Leu His His Asp Val Lys Arg Pro Gly Ile Arg Leu Ala Thr Tyr
                165                 170                 175

Arg Arg Leu Arg Leu Pro Thr Ile Gly Glu Val Arg Leu His Asp Ser
            180                 185                 190

Gly Lys Arg Leu Gly Arg Leu Val Asp Arg Ser Leu Ala Val Val Gln
        195                 200                 205

Ser Val Thr Val Ser Arg Ala Gly His Arg Trp Tyr Ala Ser Val Leu
    210                 215                 220

Cys Lys Val Thr Met Thr Val Pro Asp Gln Pro Ser Arg Arg Gln Arg
225                 230                 235                 240

Glu Arg Gly Thr Ile Gly Val Asp Leu Gly Val Lys Thr Leu Ala Ala
                245                 250                 255

Leu Ser Gln Pro Leu Asp Pro Thr Asp Pro Asp Ser Asp Leu Ile Ala
            260                 265                 270

Asn Pro Arg His Leu Ala Arg Ala Gln Gln Arg Leu Leu Lys Ala Gln
        275                 280                 285

Arg Ala Leu Ser Arg Thr Glu Lys Gly Ser Arg Arg Arg Asp Arg Ala
    290                 295                 300

Arg Arg Lys Val Ala Arg Leu His His Glu Val Ala Leu Arg Arg Glu
305                 310                 315                 320

Ser Ala Leu His Ala Val Thr Lys Arg Leu Ala Thr Ala Phe Ala Val
                325                 330                 335

Val Ala Val Glu Asp Leu His Val Ala Gly Met Thr Ala Ser Ala Arg
            340                 345                 350
```

```
Gly Thr Leu Glu Lys Pro Gly Arg Arg Val Arg Gln Lys Ala Gly Leu
        355                 360                 365

Asn Arg Ala Val Leu Asp Ala Ala Pro Gly Glu Phe Arg Arg Gln Leu
    370                 375                 380

Thr Tyr Lys Thr Ser Trp Tyr Gly Ser Lys Leu Ala Val Cys Asp Arg
385                 390                 395                 400

Trp Phe Pro Ser Ser Lys Thr Cys Ser Ala Cys Gly Trp Gln Asn Pro
                405                 410                 415

His Leu Thr Leu Thr Asp Arg Thr Phe His Cys Pro Asp Cys Gly Leu
            420                 425                 430

Thr Ile Asp Arg Asp Leu Asn Ala Ala Arg Asn Ile Ala Arg His Ala
        435                 440                 445

Thr Val Ala Asp Ala Pro Pro Val Ala Pro Gly Arg Gly Glu Thr Gln
    450                 455                 460

Asn Ala Arg Arg Ala Pro Val Arg Pro Asp Asp Arg Lys Ala Ala Arg
465                 470                 475                 480

His Gly Ala Met Lys Arg Glu Asp Thr Arg Pro Leu Gly Gln Val Pro
                485                 490                 495

Pro Gln Arg Ser Asn Pro Leu Ala Ser Pro Pro Thr Gln Lys Gln Ala
            500                 505                 510

Thr Leu Phe
        515
```

<210> 16
<211> 553
<212> PRT
<213> Micrococcus sp.

<400> 16

```
Met Thr Thr Thr Thr Gly Thr Asp Leu Ala Pro Arg Leu Arg Ala Phe
1               5                   10                  15
```

```
Lys His Arg Leu Asp Pro Asn Pro Ala Gln Ala Thr Leu Leu Ala Gln
            20                  25                  30

Tyr Ala Gly Ala Ala Arg Val Ala Tyr Asn Met Leu Ile Ala His Asn
        35                  40                  45

Arg Ala Ala Leu Ala Ala Gly Ala Ala Arg Arg Thr Glu Leu Ala Glu
    50                  55                  60

Ser Gly Leu Ala Gly Pro Glu Leu Ala Ala Arg Met Lys Ala Glu Arg
65                  70                  75                  80

Ala Ala Asp Pro Thr Leu Arg Val Ala Ser Tyr Gln Ser Tyr Ser Thr
                85                  90                  95

Ala His Leu Thr Pro Leu Ile Arg Arg His Arg Glu Ala Ala Ala Ala
            100                 105                 110

Ile Ala Ala Gly Ala Asp Pro Ala Glu Ala Trp Thr Asp Glu Arg Tyr
        115                 120                 125

Ala Glu Pro Trp Met His Thr Val Pro Arg Arg Val Leu Val Ser Gly
    130                 135                 140

Leu Gln Asn Ala Ala Lys Ala Thr Glu Asn Trp Met Ala Ser Ala Ser
145                 150                 155                 160

Gly Thr Arg Ala Gly Ala Arg Val Gly Leu Pro Arg Phe Lys Lys Lys
                165                 170                 175

Gly Arg Ser Arg Asp Ser Phe Thr Ile Pro Ala Pro Glu Val Ile Gly
            180                 185                 190

Ala Ala Gly Thr Pro Tyr Lys Arg Gly Glu Pro Arg Arg Gly Val Ile
        195                 200                 205

Thr Asp His Arg His Leu Arg Leu Ala Ser Leu Gly Thr Ile Arg Thr
    210                 215                 220
```

```
Tyr Asp Lys Thr Ser Arg Leu Val Arg Ala Cys Arg Arg Gly Ala Gln
225                 230                 235                 240

Ile Arg Ser Met Thr Ile Ser Gln Ala Gly Gly Arg Trp Tyr Ala Ser
                245                 250                 255

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

Gln Arg Ala Asn Gly Ala Val Gly Val Asp Leu Gly Val Lys His Leu
        275                 280                 285

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

Arg Gln Ala Ala Arg Leu Ala Arg Leu Gln Arg Ala Tyr Ala Arg Thr
305                 310                 315                 320

Gln Pro Gly Ser Asn Arg Arg Glu Arg Val Arg Arg Gln Ile Ala Ala
                325                 330                 335

Leu His His Gly Ile Ala Leu Arg Arg Ala Gly Leu Leu His Gln Val
            340                 345                 350

Ser Thr Arg Leu Ala Thr Asp Phe Ala Val Val Ala Leu Glu Asp Leu
        355                 360                 365

Asn Val Ala Gly Met Thr Arg Ser Ala Arg Gly Thr Leu Glu Ala Pro
    370                 375                 380

Gly Arg Asn Val Ala Ala Lys Ser Gly Leu Asn Arg Ala Ile Leu Asp
385                 390                 395                 400

Ala Gly Leu Gly Met Leu Arg Arg Gln Leu Asp Tyr Lys Thr Ser Trp
                405                 410                 415

Ala Gly Ser Gln Val Lys Met Ile Asp Arg Phe Ala Pro Ser Ser Lys
            420                 425                 430

Ala Cys Ser Arg Cys Gly Thr Val Lys Ser Thr Leu Ser Leu Ala Glu
        435                 440                 445
```

```
Arg Thr Phe Glu Cys Glu Ala Cys His Leu Val Ile Asp Arg Asp Val
    450                 455                 460

Asn Ala Ala Ile Asn Ile Arg Ala Trp Ala Val Gln Glu Glu Arg Gly
465                 470                 475                 480

Ala Gly Val Glu Leu Ala Arg Gly Arg Arg Glu Ser Arg Asn Gly Arg
                485                 490                 495

Gly Ala Ala Val Ser Gly Pro Pro Ser Gly Gly Ala Ala Gly Gln Gly
            500                 505                 510

Arg Gly Ser Val Lys Pro Ala Pro Gln Gly Val Gly Met Ser Ser Arg
        515                 520                 525

Ala Thr Gly Trp Ser Ser Gln Pro Pro Ser Thr Glu Gly Glu Ser Ala
    530                 535                 540

Glu Arg Gly Ala Ser Ala Leu Ala Arg
545                 550
```

<210> 17
<211> 557
<212> PRT
<213> Bacillus sp.

<400> 17

```
Met Thr Thr Arg Glu Ile Leu Arg Ala Tyr Arg Val Pro Leu Asp Pro
1               5                   10                  15

Thr Asp Ala Gln Thr Ala Ala Leu Ala Ser His Ala Gly Ala Ser Arg
            20                  25                  30

Ala Ala Phe Asn Trp Ala Leu Gly Ala Lys Val His Ala His Arg Met
        35                  40                  45

Trp Ser Ala Cys Val Ala Asp Leu Thr Tyr Thr Arg Tyr Gly His Leu
    50                  55                  60

Asp Ala Asp Gln Ala Leu Ala Ala Ala Lys Lys Asp Ala Ser Arg Tyr
65                  70                  75                  80
```

```
Tyr Arg Ile Pro Thr Ser Gln Thr Asn Glu Lys Ala Phe Asp Arg Asp
                85                  90                  95

Pro Asp Tyr Ala Trp Arg Thr Glu Val Asn Arg Arg Ser Trp Val Ser
            100                 105                 110

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

Thr Gly Arg Arg Ala Gly Arg Arg Val Gly Tyr Pro Arg Phe Lys Ser
    130                 135                 140

Lys Gly Arg Cys Arg Asp Ser Phe Thr Leu Ala His Asp Val Lys Arg
145                 150                 155                 160

Pro Ser Ile Arg Pro Asp Gly Tyr Arg Arg Leu Thr Leu Pro Lys Lys
                165                 170                 175

Ile Ser Val Thr Gly Ser Ile Arg Leu Lys Gly Asn Ile Arg His Leu
            180                 185                 190

Ala Arg Arg Ile Arg Arg Gly Val Ala Arg Ile Gln Ser Ala Thr Ile
        195                 200                 205

Ser Arg Ala Gly Asn Gly Trp Ser Val Ser Ile Leu Ala Leu Glu Thr
    210                 215                 220

Leu Asp Ile Pro Asp His Pro Thr Pro Arg Gln Gln Ala Ala Gly Ala
225                 230                 235                 240

Val Gly Val Asp Val Gly Val His His Leu Met Ala Phe Ser Asp Ala
                245                 250                 255

Thr Ile Ile Asp Asn Pro Arg His Leu Arg Ala Ala Gln Lys Arg Leu
            260                 265                 270

Thr Arg Ala Gln Arg Ala Leu Ser Arg Ser Lys Trp Arg Leu Pro Asn
        275                 280                 285
```

```
Gly Asp Leu Ile Asp Thr Pro Lys Arg Gly Gln Arg Val Thr Pro Thr
    290                 295                 300

Thr Gly Arg Val Lys Ala Arg Ala Arg Leu Ala Arg Glu His Ala Ala
305                 310                 315                 320

Val Ala Gln His Arg Ala Ser Thr Leu His Ala Ile Thr Lys Gln Leu
                325                 330                 335

Ala Thr Ser His Ala Val Val Ala Val Asp Asp Leu Asn Val Ala Val
            340                 345                 350

Met Thr Arg Ser Ala Arg Gly Ser Ile Asp Lys Pro Gly Arg Asn Val
        355                 360                 365

Ala Ala Lys Ala Gly Leu Asn Arg Ser Ile Leu Asp Ala Ser Phe Ala
    370                 375                 380

Glu Met Arg Arg Gln Leu Thr Tyr Lys Thr Ser Trp Tyr Arg Ser Gln
385                 390                 395                 400

Leu Leu Pro Ser Gly Cys Phe Val Pro Thr Ser Arg Thr Cys Ser Thr
                405                 410                 415

Cys Gly Ala Glu Lys Ala Asn Leu Pro Arg Ser Glu Arg Val Tyr His
            420                 425                 430

Cys Glu Asn Cys Ala Thr Val Leu Asp Arg Asp Val Asn Ala Ala Lys
        435                 440                 445

Asn Val Leu Arg Thr Ala Leu Ala Ser His Asp Ala Pro Gly Met Glu
    450                 455                 460

Glu Ser Gln Asn Ala Arg Gly Gly Arg Gly Glu Thr Ser Val Ala Arg
465                 470                 475                 480

Arg Ser Ala Lys Arg Glu Asp Pro Pro Arg Gly Gly Pro Pro Arg Pro
                485                 490                 495

Cys Asn Arg Val Arg Ser Ser Pro Pro Asp Gly Glu Ala Lys Val Arg
            500                 505                 510
```

```
Gly Lys Leu Pro Pro Pro Thr Ala Lys Arg Gly Asn Thr Ala Pro Arg
        515                 520                 525

Lys Gly Arg Pro Arg Ala Arg Gly Asp Glu Pro Asn Lys Pro Asn Arg
    530                 535                 540

Gly Ala Leu Leu Lys Met Ser Ala Pro Arg Thr Arg Arg
545                 550                 555
```

<210> 18
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 18

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Ser Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

```
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Met Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Leu Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

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

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350
```

Gln Glu Lys Ile Lys Asn Gln Ala Glu Phe His Gly Ile Glu Val Ile
355 360 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
370 375 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385 390 395 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
405 410 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
420 425 430

Lys Leu His Ser Lys Lys Cys Ile Glu Glu Tyr Ile Glu Glu Leu Gly
435 440 445

Tyr Leu Asp
450

<210> 19
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 19
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

```
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300
```

```
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320


Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335


Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350


Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365


Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380


Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400


Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415


Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450


<210> 20
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 20
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30
```

```
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Ala Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240
```

```
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Thr Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Ile Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Ala Glu Tyr Met Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 21
<211> 377
<212> PRT
<213> Bacillus sp.

<400> 21
Met Leu Val Asn Lys Ala Tyr Lys Phe Arg Ile Tyr Pro Asn Lys Glu
1   5   10   15

Gln Glu Ile Leu Ile Ala Lys Thr Ile Gly Cys Ser Arg Phe Val Phe
            20   25   30

Asn His Phe Leu Gly Met Trp Asn Asp Thr Tyr Lys Glu Thr Gly Lys
        35   40   45

Gly Leu Thr Tyr Asn Ser Cys Ser Ala Gln Leu Pro Gln Leu Lys Ile
    50   55   60

Glu Leu Glu Trp Leu Lys Glu Val Asp Ser Ile Ala Ile Gln Ser Ala
65   70   75   80

Leu Lys Asn Leu Val Asp Ala Tyr Asn Arg Phe Phe Lys Lys Gln Asn
                85   90   95

Asp Lys Pro Arg Phe Lys Ser Lys Lys Asn Asp Val Gln Ser Tyr Lys
            100   105   110

Thr Lys His Thr Asn Gly Asn Ile Ala Ile Val Asn Asn Lys Ile Lys
        115   120   125

Leu Pro Lys Leu Gly Phe Val Thr Phe Ala Lys Ser Arg Glu Val Asp
    130   135   140

Gly Arg Ile Met Asn Ala Thr Val Arg Arg Asn Ser Ser Gly Lys Tyr
145   150   155   160

Phe Val Ala Ile Leu Thr Glu Val Glu Ile Gln Pro Leu Lys Lys Ala
                165   170   175

Asp Ser Ala Ile Gly Ile Asp Leu Gly Ile Thr Asp Phe Ala Ile Leu
            180   185   190

```
Ser Asp Gly His Lys Ile Asp Asn Asn Lys Phe Thr Ser Lys Met Glu
        195                 200                 205

Lys Lys Leu Lys Arg Glu Gln Arg Lys Leu Ser Lys Arg Ala Leu Leu
    210                 215                 220

Ala Lys Asn Lys Gly Ile His Leu Leu Asp Ala Gln Asn Tyr Gln Lys
225                 230                 235                 240

Gln Lys Cys Lys Val Ala Arg Leu His Glu Arg Val Ile Asn Gln Arg
                245                 250                 255

Asp Asp Phe Leu Asn Lys Leu Ser Thr Glu Ile Ile Lys Asn His Asp
            260                 265                 270

Ile Ile Cys Ile Glu Asp Leu Asn Thr Lys Gly Met Leu Arg Asn His
        275                 280                 285

Lys Leu Ala Lys Ser Ile Ser Asp Val Ser Trp Ser Ala Phe Val Ser
    290                 295                 300

Lys Leu Glu Tyr Lys Ala Thr Trp Tyr Gly Lys Thr Ile Val Lys Val
305                 310                 315                 320

Ser Arg Trp Phe Pro Ser Ser Gln Ile Cys Ser Asp Cys Gly His His
                325                 330                 335

Asp Gly Lys Lys Ser Leu Glu Ile Arg Gly Trp Thr Cys Pro Ile Cys
            340                 345                 350

His Ala Asn His Asp Arg Asp Phe Asn Ala Ser Lys Asn Ile Leu Ala
        355                 360                 365

Glu Gly Leu Arg Thr Leu Ala Leu Val
    370                 375
```

<210> 22
<211> 445
<212> PRT
<213> Bacillus sp.

<400> 22
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1   5   10   15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20   25   30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35   40   45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50   55   60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65   70   75   80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85   90   95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100   105   110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115   120   125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145   150   155   160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180   185   190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195   200   205

```
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr
            260                 265                 270

Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro
        275                 280                 285

Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn
    290                 295                 300

His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys
305                 310                 315                 320

Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg
                325                 330                 335

Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn
            340                 345                 350

Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr
        355                 360                 365

Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
    370                 375                 380

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
385                 390                 395                 400

Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
                405                 410                 415

Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
            420                 425                 430
```

```
Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445


<210> 23
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 23
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Ala Ile Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110


Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125


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


Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160


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

```
Lys Lys Gln Lys Ser Met Lys Ile Ile Leu Asp Arg Leu Met Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Asn Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Lys Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Asn Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Ala Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Ile Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Val His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Ser Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380
```

```
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
385                 390                 395                 400


Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415


Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450
```

<210> 24
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 24

```
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1               5                   10                  15


Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
            20                  25                  30


Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
        35                  40                  45


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


Thr Glu Lys Leu Lys His Ser Asp Ala Glu Tyr Gln Lys His Leu Thr
65                  70                  75                  80


Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Ala
                85                  90                  95


Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
            100                 105                 110
```

```
Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
        115                 120                 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
    130                 135                 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

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

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Arg Val Arg Thr Gln Leu Gln Ser Gln Arg Lys
                325                 330                 335
```

```
Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445

Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460

Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495

Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505
```

<210> 25
<211> 377
<212> PRT
<213> Bacillus sp.

<400> 25

```
Met Leu Val Asn Lys Ala Tyr Lys Phe Arg Leu Tyr Pro Thr Lys Glu
1               5                   10                  15
```

```
Gln Lys Thr Leu Ile Ala Lys Thr Ile Gly Cys Ser Arg Phe Val Phe
            20                  25                  30

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

Gly Leu Thr Tyr Asn Ser Ser Ser Ala Glu Leu Thr Lys Leu Lys Lys
    50                  55                  60

Glu Leu Val Arg Leu Lys Glu Val Asp Ser Ile Ala Leu Gln Ser Ser
65                  70                  75                  80

Leu Lys Asn Leu Ala Asp Ser Tyr Ser Arg Phe Phe Lys Lys Gln Asn
                85                  90                  95

Asn Ala Pro Arg Phe Lys Ser Lys Arg Asn Arg Val Gln Ser Tyr Thr
            100                 105                 110

Thr Lys Glu Thr Asn Gly Asn Ile Ala Val Val Gly Asn Lys Met Lys
        115                 120                 125

Leu Pro Lys Leu Gly Leu Val Arg Phe Ala Lys Ser Arg Glu Val His
    130                 135                 140

Gly Arg Val Leu Asn Ala Thr Val Arg Arg Thr Pro Ser Gly Lys Tyr
145                 150                 155                 160

Phe Val Ser Ile Leu Ala Glu Val Asp Val Leu Pro Met Glu Lys Ala
                165                 170                 175

Glu Ser Ser Ile Gly Ile Asp Leu Gly Ile Thr Asp Phe Ala Ile Phe
            180                 185                 190

Ser Asp Gly Arg Met Ile Asp Asn Asn Lys Phe Thr Ala Lys Met Glu
        195                 200                 205

Lys Lys Leu Lys Arg Glu Gln Arg Lys Leu Ser Arg Arg Ala Leu His
    210                 215                 220
```

```
Ala Lys Gln Asn Gly Ile Asn Leu Leu Asp Ala Lys Asn Tyr Gln Lys
225                 230                 235                 240


Gln Lys Arg Lys Val Ala Arg Leu His Glu Arg Val Leu Asn Gln Arg
                245                 250                 255


Asp Asp Phe Leu Asn Lys Leu Ser Thr Glu Ile Ile Lys Asn His Asp
            260                 265                 270


Leu Ile Cys Ile Glu Asn Leu Asn Thr Lys Gly Met Leu Arg Asn His
        275                 280                 285


Lys Leu Ala Lys Ser Ile Ser Asp Val Ser Trp Ser Ala Phe Val Ser
    290                 295                 300


Lys Leu Glu Tyr Lys Ala Thr Trp Tyr Gly Lys Ala Ile Met Lys Val
305                 310                 315                 320


Ser Lys Trp Phe Pro Ser Ser Gln Ile Cys Ser Asp Cys Gly His Gln
                325                 330                 335


Asp Cys Lys Lys Ser Leu Ala Ile Arg Glu Trp Thr Cys Pro Ile Cys
            340                 345                 350


His Gln His His Asp Arg Asp Ile Asn Ala Ser Lys Asn Ile Leu Ala
        355                 360                 365


Glu Gly Leu Arg Thr Leu Ala Leu Ala
    370                 375


<210> 26
<211> 445
<212> PRT
<213> Bacillus sp.

<400> 26
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30
```

```
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255
```

```
Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr
            260                 265                 270

Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro
        275                 280                 285

Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn
    290                 295                 300

His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys
305                 310                 315                 320

Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg
                325                 330                 335

Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn
            340                 345                 350

Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr
        355                 360                 365

Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
    370                 375                 380

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
385                 390                 395                 400

Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
                405                 410                 415

Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
            420                 425                 430

Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445
```

<210> 27
<211> 448
<212> PRT
<213> Bacillus sp.

<400> 27
Met Thr Phe Phe Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu
1   5   10   15

Cys Pro Leu Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn
            20   25   30

Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu
        35   40   45

Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr
    50   55   60

Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp
65   70   75   80

Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys
                85   90   95

Gly Asn Met Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser
            100   105   110

Arg Arg Asn Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg
        115   120   125

Asn Arg Ile Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys
    130   135   140

Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp
145   150   155   160

Phe His Lys Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys
                165   170   175

Leu Ala Thr Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu
            180   185   190

Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys
        195   200   205

```
Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu
    210                 215                 220

Asn Lys Phe Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile
225                 230                 235                 240

Asn Thr Val Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile
                245                 250                 255

Lys Ser Asp Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln
            260                 265                 270

Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg
        275                 280                 285

Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn
    290                 295                 300

Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys
305                 310                 315                 320

His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys
                325                 330                 335

Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys
            340                 345                 350

Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala
        355                 360                 365

Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys
    370                 375                 380

Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr
385                 390                 395                 400

Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp
                405                 410                 415

Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His
            420                 425                 430
```

```
Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445


<210> 28
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 28
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110


Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125


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


Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160


Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr
                165                 170                 175
```

```
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380
```

```
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln
385                 390                 395                 400


Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415


Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450
```

<210> 29
<211> 471
<212> PRT
<213> Bacillus sp.

<400> 29

```
Met Ile Ile Ala Arg Lys Ile Lys Leu Ile Ile Ile Gly Glu Asp Arg
1               5                   10                  15


Asp Thr Gln Tyr Lys Phe Ile Arg Glu Glu Arg Tyr Lys Gln Asn Lys
            20                  25                  30


Ala Leu Asn Val Ala Met Asn His Leu Tyr Phe Leu His Val Ala Lys
        35                  40                  45


Glu Lys Ile Arg Leu Leu Asp Asn Lys Phe Leu Gln Asp Glu Lys Lys
    50                  55                  60


Leu Gln Glu Gly Ile Lys Lys Leu Tyr Ala Glu Lys Lys Val Ile Lys
65                  70                  75                  80


Asp Gly Lys Lys Arg Asn Glu Leu Glu Lys Lys Ile Glu Lys Gln Thr
                85                  90                  95


Asn Glu Leu Lys Lys Leu Arg Ser Lys Gly Asn Lys Glu Ala Asp Lys
            100                 105                 110
```

```
Ile Leu Gln Glu Ala Ile Lys Ile Asn Leu Ser Ser Thr Thr Arg Glu
        115                 120                 125

Val Ile Ser Lys Gln Phe Asp Leu Ile Ser Asp Thr Lys Asp Arg Ile
    130                 135                 140

Thr Gln Lys Val Tyr Gln Asp Phe Lys Ser Asp Leu Lys Asn Gly Leu
145                 150                 155                 160

Leu Ser Gly Glu Arg Val Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu
                165                 170                 175

Leu Ile Arg Gly Arg Ala Leu Asn Phe Tyr Arg Glu Gly Lys Asp Val
            180                 185                 190

Met Ile Lys Trp Phe Gly Gly Ile Ile Phe Lys Cys Met Leu Gly Gln
        195                 200                 205

His Lys Asn Asn Ala Gln Glu Leu Lys Ala Thr Leu Asn Lys Val Leu
    210                 215                 220

Glu Gly Ser Tyr Lys Val Cys Asp Ser Ser Ile Ser Val Gly Lys Glu
225                 230                 235                 240

Leu Ile Leu Asn Ile Ser Leu Asp Ile Gly Glu Val Asn Ser Asn Val
                245                 250                 255

Ser Cys Lys Lys Gly Arg Val Leu Gly Val Asp Leu Gly Met Lys Val
            260                 265                 270

Pro Ala Tyr Met Ser Ile Asn Asp Lys Pro Tyr Ile Arg Lys Ser Leu
        275                 280                 285

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

Arg Arg Asn Leu His Lys Thr Leu Val Ser Val Lys Gly Gly Lys Gly
305                 310                 315                 320

Arg Glu Lys Lys Leu Gln Ala Leu Asp Arg Leu Lys Glu Lys Asn Phe
                325                 330                 335
```

Ala Thr Thr Tyr Asn His Phe Leu Ser Tyr Asn Ile Val Lys Phe Ala
340 345 350

Lys Asp Asn Leu Ala Glu Gln Ile Asn Met Glu Phe Leu Ala Leu Ala
355 360 365

Gly Glu Asp Lys Asn Ile Ile Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
370 375 380

Gln Gln Phe Val Glu Asp Lys Ala Lys Arg Glu Gly Ile Asp Val Lys
385 390 395 400

Tyr Val Asp Pro Tyr Arg Thr Ser Gln Met Cys Ser Lys Cys Arg Asn
405 410 415

Tyr Glu Pro Gly Gln Arg Glu Ser Gln Glu Lys Phe Ile Cys Lys Ser
420 425 430

Cys His Leu Glu Ile Asn Ala Asp Tyr Asn Ala Ser Gln Asn Ile Ala
435 440 445

His Ser Thr Lys Tyr Ile Thr Asn Lys Asn Gln Ser Glu Tyr Phe Lys
450 455 460

Lys Leu Gln His Thr Thr Glu
465 470

<210> 30
<211> 451
<212> PRT
<213> Bacillus sp.

<220>
<221> MOD_RES
<222> (160)..(160)
<223> Any amino acid

<400> 30
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

```
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Xaa
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240
```

```
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445
```

Tyr Leu Asp
450

<210> 31
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 31
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
85 90 95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
100 105 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
115 120 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145 150 155 160

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

```
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400
```

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
405 410 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
420 425 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
435 440 445

Tyr Leu Asp
450

<210> 32
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 32
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
85 90 95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
100 105 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
115 120 125

```
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn
    130                 135                 140

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335
```

```
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 33
<211> 471
<212> PRT
<213> Bacillus sp.

<400> 33

```
Met Ile Ile Ala Arg Lys Ile Lys Leu Ile Ile Ile Gly Glu Asp Arg
1               5                   10                  15

Asp Thr Gln Tyr Lys Phe Ile Arg Glu Glu Arg Tyr Lys Gln Asn Lys
            20                  25                  30

Ala Leu Asn Val Ala Met Asn His Leu Tyr Phe Leu His Val Ala Lys
        35                  40                  45

Glu Lys Ile Arg Leu Leu Asp Asn Lys Phe Leu Gln Asp Glu Lys Lys
    50                  55                  60
```

```
Leu Gln Glu Gly Ile Lys Lys Leu Tyr Ala Glu Lys Lys Val Ile Lys
65                  70                  75                  80

Asp Gly Lys Lys Arg Asn Glu Leu Glu Lys Lys Ile Glu Lys Gln Thr
                85                  90                  95

Asn Glu Leu Lys Lys Leu Arg Ser Lys Gly Asn Lys Glu Ala Asp Lys
            100                 105                 110

Ile Leu Gln Glu Ala Ile Lys Ile Asn Leu Ser Ser Thr Thr Arg Glu
        115                 120                 125

Val Ile Ser Lys Gln Phe Asp Leu Ile Ser Asp Thr Lys Asp Arg Ile
    130                 135                 140

Thr Gln Lys Val Tyr Gln Asp Phe Lys Ser Asp Leu Lys Asn Gly Leu
145                 150                 155                 160

Leu Ser Gly Glu Arg Val Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu
                165                 170                 175

Leu Ile Arg Gly Arg Ala Leu Asn Phe Tyr Arg Glu Gly Lys Asp Val
            180                 185                 190

Met Ile Lys Trp Phe Gly Gly Ile Ile Phe Lys Cys Met Leu Gly Gln
        195                 200                 205

His Lys Asn Asn Ala Gln Glu Leu Lys Ala Thr Leu Asn Lys Val Leu
    210                 215                 220

Glu Gly Ser Tyr Lys Val Cys Asp Ser Ser Ile Ser Val Gly Lys Glu
225                 230                 235                 240

Leu Ile Leu Asn Ile Ser Leu Asp Ile Gly Glu Val Asn Ser Asn Val
                245                 250                 255

Ser Cys Lys Lys Gly Arg Val Leu Gly Val Asp Leu Gly Met Lys Val
            260                 265                 270

Pro Ala Tyr Met Ser Ile Asn Asp Lys Pro Tyr Ile Arg Lys Ser Leu
        275                 280                 285
```

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

Arg Arg Asn Leu His Lys Thr Leu Val Ser Val Lys Gly Gly Lys Gly
305                 310                 315                 320

Arg Glu Lys Lys Leu Gln Ala Leu Asp Arg Leu Lys Glu Lys Asn Phe
                325                 330                 335

Ala Thr Thr Tyr Asn His Phe Leu Ser Tyr Asn Ile Val Lys Phe Ala
            340                 345                 350

Lys Asp Asn Leu Ala Glu Gln Ile Asn Met Glu Phe Leu Ala Leu Ala
        355                 360                 365

Gly Glu Asp Lys Asn Ile Ile Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
    370                 375                 380

Gln Gln Phe Val Glu Asp Lys Ala Lys Arg Glu Gly Ile Asp Val Lys
385                 390                 395                 400

Tyr Val Asp Pro Tyr Arg Thr Ser Gln Met Cys Ser Lys Cys Arg Asn
                405                 410                 415

Tyr Glu Pro Gly Gln Arg Glu Ser Gln Glu Lys Phe Ile Cys Lys Ser
            420                 425                 430

Cys His Leu Glu Ile Asn Ala Asp Tyr Asn Ala Ser Gln Asn Ile Ala
        435                 440                 445

His Ser Thr Lys Tyr Ile Thr Asn Lys Asn Gln Ser Glu Tyr Phe Lys
    450                 455                 460

Lys Leu Gln His Thr Thr Glu
465                 470
```

<210> 34
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 34
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1   5   10   15

Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr
            20   25   30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35   40   45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50   55   60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65   70   75   80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85   90   95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100   105   110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Thr Arg Ile
        115   120   125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145   150   155   160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180   185   190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195   200   205

```
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile His Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430
```

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
435 440 445

Tyr Leu Asp
450

<210> 35
<211> 196
<212> PRT
<213> Bacillus sp.

<400> 35
Met Asp Val Gly Leu Lys Glu Leu Phe Val Ala Ser Asn Gly Met Lys
1 5 10 15

Glu Arg Asn Ile Asn Lys Asp Ala Lys Val Lys Lys Leu Leu Lys Arg
20 25 30

Lys Lys Ser Ala Gln Arg Asp Met Ser Arg Arg Phe Lys Asn Gly Val
35 40 45

Lys Ile Gln Ser Ala Gly Tyr Glu Lys Ala Lys Ala Glu His Leu Arg
50 55 60

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

Thr Ala Thr Leu Val Lys Thr Lys Pro Met Arg Ile Val Val Glu Asp
85 90 95

Leu Ser Ile Ser Asn Leu Leu Lys Asn Lys Lys Leu Ser Lys Ala Leu
100 105 110

Ser Phe Gln Lys Leu Asn Leu Phe Phe Gln Cys Leu Ser Tyr Lys Cys
115 120 125

Glu Lys Tyr Gly Ile Glu Tyr Val Lys Ala Asp Lys Trp Val Ala Ser
130 135 140

Ser Lys Ile Cys Ser Cys Cys Gly Val Lys Tyr Asp His Ser Val Gln
145 150 155 160

```
Ser Glu Gly Gln Trp Ser Leu Lys Ile Arg Glu Trp Arg Cys Val Arg
                165                 170                 175


Cys Asn Ser His His Asp Arg Asp Val Asn Ala Ala Ile Asn Leu Ser
            180                 185                 190


Arg Trp Val Lys
        195


<210> 36
<211> 445
<212> PRT
<213> Bacillus sp.

<400> 36
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110


Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125


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

```
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr
            260                 265                 270

Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro
        275                 280                 285

Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn
    290                 295                 300

His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys
305                 310                 315                 320

Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg
                325                 330                 335

Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn
            340                 345                 350
```

```
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr
        355                 360                 365

Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
    370                 375                 380

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
385                 390                 395                 400

Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
                405                 410                 415

Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
            420                 425                 430

Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445
```

<210> 37
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 37

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95
```

```
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320
```

```
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 38
<211> 362
<212> PRT
<213> Bacillus sp.

<400> 38

```
Met Leu Lys Ala Phe Lys Tyr Arg Ile Tyr Pro Asn Asn Glu Gln Arg
1               5                   10                  15

Ile Phe Phe Ala Lys Thr Phe Gly Cys Val Arg Phe Val Tyr Asn Lys
            20                  25                  30

Met Leu Ala Asp Arg Ile Glu Ser Tyr Gln Glu Ser Gln Asp Lys Pro
        35                  40                  45
```

```
Asp Lys Ser Ile Lys Tyr Pro Thr Pro Ala Gln Tyr Lys Val Glu Phe
    50                  55                  60

Pro Phe Leu Lys Glu Val Asp Ser Leu Ala Leu Ala Asn Ala Gln Met
65                  70                  75                  80

Asn Leu Asn Lys Ala Tyr Ala Asn Phe Phe Arg Asp Lys Ser Val Gly
                85                  90                  95

Phe Pro Lys Phe Lys Ser Lys Lys Asp Arg His Arg Ser Tyr Thr Thr
            100                 105                 110

Asn Asn Gln Lys Gly Thr Val Cys Ile Glu Lys Gly Tyr Ile Lys Leu
        115                 120                 125

Pro Lys Leu Lys Thr Leu Val Lys Ile Lys Gln His Arg Gln Phe Phe
    130                 135                 140

Gly Leu Met Lys Ser Val Thr Ile Ser Gln Thr Pro Thr Gly Lys Tyr
145                 150                 155                 160

Phe Val Ser Val Leu Val Glu Glu Lys Glu Gln Leu Ser Pro Lys Thr
                165                 170                 175

Asp Glu Lys Val Gly Val Asp Leu Gly Leu Lys Asp Phe Ala Ile Leu
            180                 185                 190

Ser Asn Gly Thr Lys Tyr Glu Asn Pro Lys Trp Leu Arg Arg Leu Glu
        195                 200                 205

Lys Arg Leu Ala Phe Leu Gln Arg Ser Leu Ser Arg Lys Lys Lys Gly
    210                 215                 220

Ser Asn Asn Leu Asn Lys Ala Arg Leu Gln Val Ala Arg Leu His Glu
225                 230                 235                 240

Lys Ile Ala Asn Gln Arg Asn Asp Phe Leu His Lys Ile Ser Asn Glu
                245                 250                 255
```

Ile Thr Asn Glu Asn Gln Val Ile Val Ile Glu Asp Leu Lys Val Lys
260 265 270

Asn Met Gln Lys Asn His Lys Leu Thr Arg Ala Ile Ser Glu Val Ser
275 280 285

Trp Ser Lys Phe Arg Glu Tyr Leu Ala Tyr Lys Thr Ala Trp Lys Gly
290 295 300

Arg Asp Leu Ile Val Ala Pro Lys Asn Tyr Ala Ser Ser Gln Leu Cys
305 310 315 320

Ser Cys Cys Gly Tyr Lys Asn Lys Glu Val Lys Asn Leu Asn Leu Arg
325 330 335

Glu Trp Thr Cys Pro Glu Cys Asn Ser His His Asn Arg Asp Ile Asn
340 345 350

Ala Ser Ile Asn Leu Leu Lys Leu Ala Met
355 360

<210> 39
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 39
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Val Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

```
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Ile Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Ile Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Asn Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Asn Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300
```

```
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320


Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335


Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350


Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365


Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380


Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Glu
385                 390                 395                 400


Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415


Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450


<210> 40
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 40
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30
```

```
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240
```

```
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 41
<211> 450
<212> PRT
<213> Bacillus sp.

<400> 41
Met Ser Ile Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1   5   10   15

Asn Ile Glu Trp Lys Thr Phe Glu Ile Tyr Leu Arg Thr Leu Ser Tyr
            20   25   30

His Phe Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Leu Trp Glu Tyr
        35   40   45

Asp Asn Gln Ser Leu Lys His Phe Lys Asp Thr Gly Gln Tyr Pro Ser
    50   55   60

Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65   70   75   80

Tyr Asp Gln Leu Lys Glu Glu Tyr Gln Asp Ile Asn Lys Ala Asn Met
                85   90   95

Ser Thr Thr Leu Gln Lys Thr Ile Arg Thr Trp Asn Ser Arg Lys Lys
            100   105   110

Glu Ile Trp Ser Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Asn Leu
        115   120   125

Pro Ile Asp Ile His Gly Asn Ser Ile Gln Ile Ile Lys Glu Lys Ser
    130   135   140

Gly Asp Tyr Ile Ala Ser Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
145   150   155   160

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

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Ile Ile Asp Ser Thr
            180   185   190

```
Tyr Ala Lys Gly Ala Cys Met Leu His Lys His Lys Lys Lys Trp Tyr
        195                 200                 205

Leu Ser Ile Thr Tyr Lys Ser Asn Ile Lys Glu Glu Leu Lys Phe Asp
    210                 215                 220

Glu Asp Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
225                 230                 235                 240

Phe Ala Phe Asn Lys Gly Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
                245                 250                 255

Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Gly Asn Arg Ile Gly Lys Gly Arg Glu
        275                 280                 285

Lys Arg Ile Lys Pro Ile Asp Val Leu Asn Asp Lys Val Ala Lys Phe
    290                 295                 300

Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320

Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Lys Gly Ile
                325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350

Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Val Lys Val Val Lys
        355                 360                 365

Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
    370                 375                 380

His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln Gln Cys
385                 390                 395                 400
```

```
Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Leu
                405                 410                 415


Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
            420                 425                 430


Leu Asn Gln Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Lys Asn
        435                 440                 445


Ile Asn
    450
```

<210> 42
<211> 450
<212> PRT
<213> Bacillus sp.

<400> 42

```
Met Ser Ile Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1               5                   10                  15


Asn Ile Glu Trp Lys Thr Phe Glu Ile Tyr Leu Arg Thr Leu Ser Tyr
            20                  25                  30


His Phe Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Leu Trp Glu Tyr
        35                  40                  45


Asp Asn Gln Ser Leu Lys His Phe Lys Asp Thr Gly Gln Tyr Pro Ser
    50                  55                  60


Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65                  70                  75                  80


Tyr Asp Gln Leu Lys Glu Glu Tyr Gln Asp Ile Asn Lys Ala Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Thr Ile Arg Thr Trp Asn Ser Arg Lys Lys
            100                 105                 110


Glu Ile Trp Ser Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Asn Leu
        115                 120                 125
```

```
Pro Ile Asp Ile His Gly Asn Ser Ile Gln Ile Ile Lys Glu Lys Ser
    130                 135                 140

Gly Asp Tyr Ile Ala Ser Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
145                 150                 155                 160

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

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Ile Ile Asp Ser Thr
            180                 185                 190

Tyr Ala Lys Gly Ala Cys Met Leu His Lys His Lys Lys Lys Trp Tyr
        195                 200                 205

Leu Ser Ile Thr Tyr Lys Ser Asn Ile Lys Glu Glu Leu Lys Phe Asp
    210                 215                 220

Glu Asp Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
225                 230                 235                 240

Phe Ala Phe Asn Lys Gly Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
                245                 250                 255

Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Gly Asn Arg Ile Gly Lys Gly Arg Glu
        275                 280                 285

Lys Arg Ile Lys Pro Ile Asp Val Leu Asn Asp Lys Val Ala Lys Phe
    290                 295                 300

Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320

Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Lys Gly Ile
                325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350
```

```
Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Val Lys Val Val Lys
        355                 360                 365


Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
    370                 375                 380


His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln Gln Cys
385                 390                 395                 400


Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Leu
                405                 410                 415


Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
            420                 425                 430


Leu Asn Gln Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Lys Asn
        435                 440                 445


Ile Asn
    450
```

<210> 43
<211> 444
<212> PRT
<213> Bacillus sp.

<400> 43

```
Met Lys Tyr Gln Ile Leu Cys Pro Met Asn Val Asp Trp Thr Ile Phe
1               5                   10                  15


Glu Lys His Leu Arg Asn Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn
            20                  25                  30


Arg Thr Ile Gln Gln Leu Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr
        35                  40                  45


Phe Lys Glu Arg Gly Thr Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys
    50                  55                  60


Thr Gln Lys Lys Ile Asp Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys
65                  70                  75                  80
```

```
Tyr Pro Asp Ile His Lys Gly Asn Met Ser Thr Thr Leu Gln Lys Ile
                85                  90                  95

Ile Lys Thr Trp Lys Ser Arg Arg Asn Glu Ile Arg Lys Gly Glu Met
            100                 105                 110

Ser Ile Pro Ser Phe Arg Asn Arg Ile Pro Ile Asp Leu His Asn Asn
        115                 120                 125

Ser Val Asp Ile Ile Lys Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile
    130                 135                 140

Ser Leu Phe Ser Arg Asp Phe His Lys Glu Asn Gly Asp Val Pro Lys
145                 150                 155                 160

Gly Lys Ile Phe Val Lys Leu Gly Thr Gln Lys Gln Lys Ser Met Lys
                165                 170                 175

Val Ile Leu Asp Arg Leu Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys
            180                 185                 190

Met Ile His Lys Tyr Lys Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys
        195                 200                 205

Phe Asn Ala Ile Lys Glu Asn Lys Phe Asp Lys Glu Leu Ile Met Gly
    210                 215                 220

Ile Asp Met Gly Gly Ile Asn Thr Val Tyr Phe Ala Phe Asn Glu Gly
225                 230                 235                 240

Phe Ile Arg Ser Asn Ile Lys Ser Asp Glu Ile Lys Met Phe Asn Glu
                245                 250                 255

Arg Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr Cys
            260                 265                 270

Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro Ile
        275                 280                 285
```

```
Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn His
    290                 295                 300

Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys Gly
305                 310                 315                 320

Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Lys Val
                325                 330                 335

Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln
            340                 345                 350

Ala Glu Ile Tyr Gly Ile Glu Val Ile Lys Val Val Pro Ala Tyr Thr
        355                 360                 365

Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys
    370                 375                 380

Thr Gln Ala Met Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His Ala
385                 390                 395                 400

Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile
                405                 410                 415

Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys Cys
            420                 425                 430

Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440
```

<210> 44
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 44

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg His Leu Thr Tyr
            20                  25                  30
```

```
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Thr Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Val Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255
```

```
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 45
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 45
Met Arg Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1   5   10   15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20   25   30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35   40   45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Lys Gly Thr Tyr Pro Thr
    50   55   60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65   70   75   80

Tyr His Thr Leu Gln Ser Lys Tyr Leu Asp Ile His Lys Gly Asn Met
                85   90   95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100   105   110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115   120   125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145   150   155   160

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

Gln Arg Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180   185   190

```
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Glu Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Gln Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415
```

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
420 425 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
435 440 445

Tyr Leu Asp
450

<210> 46
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 46
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1 5 10 15

Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
20 25 30

Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
35 40 45

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

Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
65 70 75 80

Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
85 90 95

Lys Glu Lys Val Asn Asp Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
100 105 110

Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
115 120 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
130 135 140

```
Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

Ile Ser His Phe Asn Asn Asp Met Lys Thr Gly Val Leu Arg Gly Asp
            180                 185                 190

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Lys Val Arg Thr Gln Leu Gln Ser Gln Arg Lys
                325                 330                 335

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Asn
            340                 345                 350
```

```
Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445

Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460

Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495

Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505
```

<210> 47
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 47

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

Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
            20                  25                  30
```

```
Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
        35                  40                  45

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

Thr Glu Lys Leu Lys His Ser Asp Ala Glu Tyr Gln Lys His Leu Thr
65                  70                  75                  80

Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Ala
                85                  90                  95

Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Glu Lys Arg
            100                 105                 110

Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
        115                 120                 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Lys Glu Thr Tyr Gln Lys
    130                 135                 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

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

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255
```

```
Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Lys Val Arg Thr Gln Leu Gln Asn Gln Arg Lys
                325                 330                 335

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Lys Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Asp His
        435                 440                 445

Tyr Glu Lys Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460
```

```
Lys Glu Cys Lys Gly Tyr Asp His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480


Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495


Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505


<210> 48
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 48
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110


Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125


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

```
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365
```

```
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 49
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 49

```
Met Arg Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Lys Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Leu Asp Ile His Lys Gly Asn Met
                85                  90                  95
```

```
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Arg Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Glu Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300
```

```
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320


Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335


Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350


Gln Glu Lys Ile Gln Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365


Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380


Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
385                 390                 395                 400


Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415


Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450


<210> 50
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 50
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1               5                   10                  15


Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
            20                  25                  30
```

```
Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
        35                  40                  45

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

Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
65                  70                  75                  80

Lys Tyr Arg Glu Val Ser Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
                85                  90                  95

Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
            100                 105                 110

Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
        115                 120                 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
    130                 135                 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

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

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255
```

```
Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Asn Val Arg Thr Gln Leu Gln Asn Gln Arg Lys
                325                 330                 335

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445

Tyr Glu Lys Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460
```

```
Lys Glu Cys Lys Gly Tyr Ala His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480


Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495


Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505
```

<210> 51
<211> 405
<212> PRT
<213> Bacillus sp.

<400> 51

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


Asn Lys Val Leu Arg His Tyr Gly Leu Lys Leu Arg Ala Leu Pro Thr
            20                  25                  30


Pro Phe Gln Glu Glu Lys Ile Ala Lys Thr Ile Gly Ser Ala Arg Phe
        35                  40                  45


Ser Tyr Asn Phe Tyr Leu Asn Glu Lys Ile Glu Ile Tyr Lys Leu Thr
    50                  55                  60


Asn Leu Thr Leu Asp Tyr Ser Ser Phe Lys Lys Ser Phe Asn Gly Leu
65                  70                  75                  80


Lys Gln His Pro Ala Phe Asp Trp Leu Lys Glu Val Asp Lys Phe Ser
                85                  90                  95


Leu Glu Ser Ala Leu Glu Gln Val Asp Asp Ala Phe Glu Arg Phe Phe
            100                 105                 110


Lys Gly Gln Thr Lys Phe Pro Lys Phe Lys Ser Lys His Lys Thr Lys
        115                 120                 125


Gln Ser Tyr Thr Thr Lys Glu Thr Asn Gly Asn Ile Ala Phe Asp Leu
    130                 135                 140
```

```
Glu Asn Gln Val Val Lys Leu Pro Lys Met Gly Lys Val Pro Val Lys
145                 150                 155                 160

Leu Ser Lys Lys His Leu Thr Met Phe Gln Lys Asn Glu Phe Thr Gly
                165                 170                 175

Lys Ile Lys Ser Ala Thr Val Thr Arg His Ser Ser Gly Gln Tyr Tyr
            180                 185                 190

Ile Ser Leu Lys Cys Glu Glu Ile Ile Leu Leu Glu Glu Lys Ile Asp
        195                 200                 205

Val Thr Thr Ile Pro Thr Asp Gly Ile Ile Gly Cys Asp Leu Gly Leu
    210                 215                 220

Thr Tyr Phe Leu Ile Asp Ser Asn Gly Gln Lys Ile Glu Asn Pro Arg
225                 230                 235                 240

Tyr Leu Lys Glu Asn Leu Lys Lys Leu Ala Lys Leu Gln Arg Gly Leu
                245                 250                 255

Lys His Lys Lys Ile Gly Ser Ser Asn Phe Gln Lys Leu Lys Gln Lys
            260                 265                 270

Ile Ala Lys Leu His Leu His Ile Ser Asn Met Arg Lys Asp Phe Leu
        275                 280                 285

His Lys Val Ser Arg Lys Leu Val Asn Glu Asn Gln Val Ile Ile Leu
    290                 295                 300

Glu Asp Leu Asn Val Lys Gly Met Ile Lys Asn Pro Lys Leu Ala Arg
305                 310                 315                 320

Ser Ile Ala Asp Val Gly Trp Gly Met Phe Lys Thr Phe Val Ser Tyr
                325                 330                 335

Lys Ala Asn Trp Ala Asn Lys Ile Leu Ile Leu Ile Asn Arg Phe Phe
            340                 345                 350

Pro Ser Ser Lys Gln Cys Asn Gly Cys Lys Glu Lys Asn Thr Leu Leu
        355                 360                 365
```

```
Ser Leu Ser Asp Arg Leu Trp Met Cys Pro Ser Cys Gly Thr His His
    370                 375                 380

Asp Arg Asp Asp Asn Ala Ala Leu Asn Ile Lys Glu Glu Gly Ile Arg
385                 390                 395                 400

Leu Leu Leu Asn Ala
                405
```

<210> 52
<211> 252
<212> PRT
<213> Bacillus sp.

<400> 52

```
Met Ile His Lys Tyr Lys Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys
1               5                   10                  15

Phe Asn Ala Ile Lys Glu Asn Lys Phe Asp Lys Glu Leu Ile Met Gly
            20                  25                  30

Ile Asp Met Gly Gly Ile Asn Thr Val Tyr Phe Ala Phe Asn Glu Gly
        35                  40                  45

Phe Ile Arg Ser Asn Ile Lys Ser Asp Glu Ile Lys Ala Phe Asn Glu
    50                  55                  60

Arg Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr Cys
65                  70                  75                  80

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

Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn His
            100                 105                 110

Lys Tyr Ala Asn Tyr Ile Ile Lys Gln Cys Leu Lys His Asn Cys Gly
        115                 120                 125

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

```
Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Gln Asn Gln
145                 150                 155                 160


Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr Thr
                165                 170                 175


Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys
            180                 185                 190


Thr Gln Ala Val Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His Ala
        195                 200                 205


Asp Tyr Asn Ala Ala Lys Asn Ile Ala Thr Tyr Asp Ile Glu Asn Ile
    210                 215                 220


Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys Cys
225                 230                 235                 240


Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
                245                 250


<210> 53
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 53
Met Arg Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Lys Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80
```

```
Tyr His Thr Leu Gln Ser Lys Tyr Leu Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Arg Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285
```

```
Glu Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Gln Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 54
<211> 445
<212> PRT
<213> Bacillus sp.

<400> 54

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1               5                   10                  15
```

```
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240
```

```
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr
            260                 265                 270

Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro
        275                 280                 285

Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn
    290                 295                 300

His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys
305                 310                 315                 320

Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg
                325                 330                 335

Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn
            340                 345                 350

Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr
        355                 360                 365

Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
    370                 375                 380

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
385                 390                 395                 400

Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
                405                 410                 415

Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
            420                 425                 430

Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445
```

<210> 55
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 55
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1   5   10   15

Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr
            20   25   30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35   40   45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50   55   60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65   70   75   80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85   90   95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100   105   110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115   120   125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145   150   155   160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180   185   190

```
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415
```

```
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450


<210> 56
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 56
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Ala Arg Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Arg Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Ile Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110


Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125


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

```
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Arg Leu Ala Thr
                165                 170                 175

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Met Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Leu Ile Thr Tyr Lys Phe Asn Ala Val Lys Glu Asn Ile Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Leu Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350
```

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
355 360 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
370 375 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385 390 395 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
405 410 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
420 425 430

Lys Leu His Ser Lys Lys Tyr Met Glu Glu Tyr Ile Glu Glu Leu Gly
435 440 445

Tyr Leu Asp
450

<210> 57
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 57
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

```
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300
```

```
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320


Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335


Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350


Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365


Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380


Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400


Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415


Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450


<210> 58
<211> 445
<212> PRT
<213> Bacillus sp.

<400> 58
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30
```

```
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240
```

```
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr
            260                 265                 270

Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro
        275                 280                 285

Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn
    290                 295                 300

His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys
305                 310                 315                 320

Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg
                325                 330                 335

Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn
            340                 345                 350

Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr
        355                 360                 365

Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
    370                 375                 380

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
385                 390                 395                 400

Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
                405                 410                 415

Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
            420                 425                 430

Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445
```

<210> 59
<211> 445

<212> PRT
<213> Bacillus sp.

<400> 59
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1   5   10   15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20   25   30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35   40   45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50   55   60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65   70   75   80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85   90   95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100   105   110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115   120   125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145   150   155   160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180   185   190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195   200   205

```
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr
            260                 265                 270

Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro
        275                 280                 285

Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn
    290                 295                 300

His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys
305                 310                 315                 320

Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg
                325                 330                 335

Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn
            340                 345                 350

Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr
        355                 360                 365

Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
    370                 375                 380

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
385                 390                 395                 400

Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
                405                 410                 415
```

```
Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
            420                 425                 430


Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445


<210> 60
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 60
Met Arg Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Lys Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110


Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125


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


Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160
```

```
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr
                165                 170                 175

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Asp Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Glu Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Gln Asn Gln Thr Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380
```

```
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
385                 390                 395                 400


Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415


Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450


<210> 61
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 61
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Ala Ile Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110
```

```
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Lys Lys Gln Lys Ser Met Lys Ile Ile Leu Asp Arg Leu Met Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Asn Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Lys Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Asn Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Ala Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320
```

```
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Ile Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Val His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Ser Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 62
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 62

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45
```

```
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270
```

```
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 63
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 63
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1   5   10   15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20   25   30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35   40   45

Asp Val Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50   55   60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65   70   75   80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85   90   95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100   105   110

Glu Ile Arg Lys Gly Glu Ile Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115   120   125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145   150   155   160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180   185   190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195   200   205

```
Tyr Leu Ser Ile Ile Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Asn Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Asn Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Glu
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430
```

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
435 440 445

Tyr Leu Asp
450

<210> 64
<211> 445
<212> PRT
<213> Bacillus sp.

<400> 64
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
85 90 95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
100 105 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
115 120 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145 150 155 160

```
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr
                165                 170                 175

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr
            260                 265                 270

Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro
        275                 280                 285

Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn
    290                 295                 300

His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys
305                 310                 315                 320

Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg
                325                 330                 335

Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn
            340                 345                 350

Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr
        355                 360                 365
```

```
Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
    370                 375                 380

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
385                 390                 395                 400

Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
                405                 410                 415

Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
            420                 425                 430

Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445
```

<210> 65
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 65

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110
```

```
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335
```

```
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 66
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 66

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Ala Ile Gly Thr Tyr Pro Thr
    50                  55                  60
```

```
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Lys Lys Gln Lys Ser Met Lys Ile Ile Leu Asp Arg Leu Met Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Asn Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Lys Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270
```

```
Leu Lys Gln Ser Asn Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Ala Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Ile Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Val His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Ser Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 67
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 67

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Val Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Ile Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Ile Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220
```

```
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Ile
225                 230                 235                 240

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

Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Asn Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Ala Gln Ser
            420                 425                 430
```

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Arg
435 440 445

Tyr Leu Asp
450

<210> 68
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 68
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Ser Pro Thr
50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
85 90 95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
100 105 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
115 120 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145 150 155 160

```
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr
                165                 170                 175

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Met Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Leu Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

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

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Phe His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380
```

```
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400


Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415


Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Ile Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450


<210> 69
<211> 496
<212> PRT
<213> Bacillus sp.

<400> 69
Met Ile Leu Thr Arg Lys Val Lys Leu Val Ile Val Ser Asp Asn Arg
1               5                   10                  15


Asp Glu Gly Tyr Lys Leu Ile Arg Asn Glu Ile Arg Glu Gln His Lys
            20                  25                  30


Ala Leu Asn Leu Ala Tyr Asn His Leu Tyr Phe Glu His Asn Ala Ile
        35                  40                  45


Gln Ile Leu Lys Gln Asn Asp Glu Asp Tyr Lys Gln Lys Arg Asn Lys
    50                  55                  60


Leu Gln Glu Leu Ile Asn Lys Lys Tyr Glu Glu His Gln Lys Ala Lys
65                  70                  75                  80


Asn Leu Glu Arg Lys Glu Ala Leu Arg Glu Ala Tyr Asn Asn Lys Lys
                85                  90                  95


Gln Glu Leu Tyr Lys Phe Glu Arg Glu Cys Asn Glu Glu Ala Arg Lys
            100                 105                 110
```

```
Ala Tyr Gln Gln Val Val Gly Phe Thr Gln Gln Thr Arg Val Arg Asn
        115                 120                 125

Leu Ile Asn Arg Glu Cys Asn Leu Met Ser Asp Thr Lys Asp Gly Ile
    130                 135                 140

Thr Ser Lys Val Thr Gln Asp Tyr Lys Asn Asp Cys Lys Ala Gly Leu
145                 150                 155                 160

Leu Ile Gly Lys Arg Ser Leu Arg Asn Tyr Lys Lys Asp Asn Pro Leu
                165                 170                 175

Leu Val Arg Gly Arg Ser Leu Lys Phe Tyr Lys Glu Asp Gly Asp Tyr
            180                 185                 190

Phe Ile Lys Trp Asn Lys Gly Thr Val Phe Lys Cys Ile Leu His Ile
        195                 200                 205

Arg Lys Lys Asn Val Ala Glu Leu Gln Ser Val Leu Glu Asn Val Leu
    210                 215                 220

Leu Gly Ala Tyr Lys Ile Cys Asp Ser Ser Ile Gly Phe Asn Asn Lys
225                 230                 235                 240

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

Asp Tyr Ile Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile
            260                 265                 270

Pro Ala Tyr Val Ser Leu Ser Asp Lys Val Tyr Val Arg Lys Gly Ile
        275                 280                 285

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

Arg Arg Gln Leu Gln Glu Ser Leu Ala Ala Val Lys Gly Gly Lys Gly
305                 310                 315                 320
```

```
Arg Glu Lys Lys Leu Lys Ala Leu Asp His Leu Lys Gly Lys Glu Ala
                325                 330                 335

Asn Phe Ala Lys Thr Tyr Asn His Phe Leu Ser Thr Gln Ile Val Thr
            340                 345                 350

Phe Ala Val Lys Asn Gln Ala Gly Gln Ile Asn Met Glu Phe Leu Glu
        355                 360                 365

Phe Asp Lys Met Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr
    370                 375                 380

Gln Leu Gln Met Met Val Glu Tyr Lys Ala Lys Arg Glu Gly Ile Ile
385                 390                 395                 400

Ile Lys Tyr Val Asp Ala Tyr Leu Thr Ser Gln Thr Cys Ser Lys Cys
                405                 410                 415

Asp Tyr Tyr Glu Glu Gly Gln Arg Glu Lys Gln Glu Lys Phe Ile Cys
            420                 425                 430

Lys Ser Cys Ala Phe Glu Val Asn Ala Asp Tyr Asn Ala Ser Gln Asn
        435                 440                 445

Ile Ala Lys Ser Ala Arg Tyr Ile Ser Asp Ser Thr Glu Arg Glu Tyr
    450                 455                 460

His Lys Lys Lys Gln Glu Asp Leu Lys Glu Ile Leu Gly Glu Asn Asp
465                 470                 475                 480

Ile Ile Asn Glu Gln Leu Ser Leu Phe Asp Asn His Asp Asp Ile Ala
                485                 490                 495
```

<210> 70
<211> 496
<212> PRT
<213> Bacillus sp.

<400> 70

```
Met Ile Leu Thr Arg Lys Val Lys Leu Val Ile Val Ser Asp Asn Arg
1               5                   10                  15
```

```
Asp Glu Gly Tyr Lys Leu Ile Arg Asn Glu Ile Arg Glu Gln His Lys
            20                  25                  30

Ala Leu Asn Leu Ala Tyr Asn His Leu Tyr Phe Glu His Asn Ala Ile
        35                  40                  45

Gln Ile Leu Lys Gln Asn Asp Glu Asp Tyr Lys Gln Lys Arg Asn Lys
    50                  55                  60

Leu Gln Glu Leu Ile Asn Lys Lys Tyr Glu Glu His Gln Lys Ala Lys
65                  70                  75                  80

Asn Leu Glu Arg Lys Glu Ala Leu Arg Glu Ala Tyr Asn Asn Lys Lys
                85                  90                  95

Gln Glu Leu Tyr Lys Phe Glu Arg Glu Cys Asn Glu Glu Ala Arg Lys
            100                 105                 110

Ala Tyr Gln Gln Val Val Gly Phe Thr Gln Gln Thr Arg Val Arg Asn
        115                 120                 125

Leu Ile Asn Arg Glu Cys Asn Leu Met Ser Asp Thr Lys Asp Gly Ile
    130                 135                 140

Thr Ser Lys Val Thr Gln Asp Tyr Lys Asn Asp Cys Lys Ala Gly Leu
145                 150                 155                 160

Leu Ile Gly Lys Arg Ser Leu Arg Asn Tyr Lys Lys Asp Asn Pro Leu
                165                 170                 175

Leu Val Arg Gly Arg Ser Leu Lys Phe Tyr Lys Glu Asp Gly Asp Tyr
            180                 185                 190

Phe Ile Lys Trp Asn Lys Gly Thr Val Phe Lys Cys Ile Leu His Ile
        195                 200                 205

Arg Lys Lys Asn Val Ala Glu Leu Gln Ser Val Leu Glu Asn Val Leu
    210                 215                 220

Leu Gly Ala Tyr Lys Ile Cys Asp Ser Ser Ile Gly Phe Asn Asn Lys
225                 230                 235                 240
```

```
Asp Met Ile Leu Asn Leu Ser Leu Asn Ile Pro Asp Lys Glu Thr Tyr
                245                 250                 255

Asp Tyr Ile Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile
            260                 265                 270

Pro Ala Tyr Val Ser Leu Ser Asp Lys Val Tyr Val Arg Lys Gly Ile
        275                 280                 285

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

Arg Arg Gln Leu Gln Glu Ser Leu Ala Ala Val Lys Gly Gly Lys Gly
305                 310                 315                 320

Arg Glu Lys Lys Leu Lys Ala Leu Asp His Leu Lys Gly Lys Glu Ala
                325                 330                 335

Asn Phe Ala Lys Thr Tyr Asn His Phe Leu Ser Thr Gln Ile Val Thr
            340                 345                 350

Phe Ala Val Lys Asn Gln Ala Gly Gln Ile Asn Met Glu Phe Leu Glu
        355                 360                 365

Phe Asp Lys Met Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr
    370                 375                 380

Gln Leu Gln Met Met Val Glu Tyr Lys Ala Lys Arg Glu Gly Ile Ile
385                 390                 395                 400

Ile Lys Tyr Val Asp Ala Tyr Leu Thr Ser Gln Thr Cys Ser Lys Cys
                405                 410                 415

Asp Tyr Tyr Glu Glu Gly Gln Arg Glu Lys Gln Glu Lys Phe Ile Cys
            420                 425                 430

Lys Ser Cys Ala Phe Glu Val Asn Ala Asp Tyr Asn Ala Ser Gln Asn
        435                 440                 445
```

```
Ile Ala Lys Ser Ala Arg Tyr Ile Ser Asp Ser Thr Glu Arg Glu Tyr
    450                 455                 460

His Lys Lys Lys Gln Glu Asp Leu Lys Glu Ile Leu Gly Glu Asn Asp
465                 470                 475                 480

Ile Ile Asn Glu Gln Leu Ser Leu Phe Asp Asn His Asp Asp Ile Ala
                485                 490                 495
```

<210> 71
<211> 496
<212> PRT
<213> Bacillus sp.

<400> 71

```
Met Ile Leu Thr Arg Lys Val Lys Leu Val Ile Val Ser Asp Asn Arg
1               5                   10                  15

Asp Glu Gly Tyr Lys Leu Ile Arg Asn Glu Ile Arg Glu Gln His Lys
            20                  25                  30

Ala Leu Asn Leu Ala Tyr Asn His Leu Tyr Phe Glu His Asn Ala Ile
        35                  40                  45

Gln Ile Leu Lys Gln Asn Asp Glu Asp Tyr Lys Gln Lys Arg Asn Lys
    50                  55                  60

Leu Gln Glu Leu Ile Asn Lys Lys Tyr Glu Glu His Gln Lys Ala Lys
65                  70                  75                  80

Asn Leu Glu Arg Lys Glu Ala Leu Arg Glu Ala Tyr Asn Asn Lys Lys
                85                  90                  95

Gln Glu Leu Tyr Lys Phe Glu Arg Glu Cys Asn Glu Glu Ala Arg Lys
            100                 105                 110

Ala Tyr Gln Gln Val Val Gly Phe Thr Gln Gln Thr Arg Val Arg Asn
        115                 120                 125

Leu Ile Asn Arg Glu Cys Asn Leu Met Ser Asp Thr Lys Asp Gly Ile
    130                 135                 140
```

```
Thr Ser Lys Val Thr Gln Asp Tyr Lys Asn Asp Cys Lys Ala Gly Leu
145                 150                 155                 160

Leu Ile Gly Lys Arg Ser Leu Arg Asn Tyr Lys Lys Asp Asn Pro Leu
                165                 170                 175

Leu Val Arg Gly Arg Ser Leu Lys Phe Tyr Lys Glu Asp Gly Asp Tyr
            180                 185                 190

Phe Ile Lys Trp Asn Lys Gly Thr Val Phe Lys Cys Ile Leu His Ile
        195                 200                 205

Arg Lys Lys Asn Val Ala Glu Leu Gln Ser Val Leu Glu Asn Val Leu
    210                 215                 220

Leu Gly Ala Tyr Lys Ile Cys Asp Ser Ser Ile Gly Phe Asn Asn Lys
225                 230                 235                 240

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

Asp Tyr Ile Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile
            260                 265                 270

Pro Ala Tyr Val Ser Leu Ser Asp Lys Val Tyr Val Arg Lys Gly Ile
        275                 280                 285

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

Arg Arg Gln Leu Gln Glu Ser Leu Ala Ala Val Lys Gly Gly Lys Gly
305                 310                 315                 320

Arg Glu Lys Lys Leu Lys Ala Leu Asp His Leu Lys Gly Lys Glu Ala
                325                 330                 335

Asn Phe Ala Lys Thr Tyr Asn His Phe Leu Ser Thr Gln Ile Val Thr
            340                 345                 350

Phe Ala Val Lys Asn Gln Ala Gly Gln Ile Asn Met Glu Phe Leu Glu
        355                 360                 365
```

```
Phe Asp Lys Met Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr
    370                 375                 380

Gln Leu Gln Met Met Val Glu Tyr Lys Ala Lys Arg Glu Gly Ile Ile
385                 390                 395                 400

Ile Lys Tyr Val Asp Ala Tyr Leu Thr Ser Gln Thr Cys Ser Lys Cys
                405                 410                 415

Asp Tyr Tyr Glu Glu Gly Gln Arg Glu Lys Gln Glu Lys Phe Ile Cys
            420                 425                 430

Lys Ser Cys Ala Phe Glu Val Asn Ala Asp Tyr Asn Ala Ser Gln Asn
        435                 440                 445

Ile Ala Lys Ser Ala Arg Tyr Ile Ser Asp Ser Thr Glu Arg Glu Tyr
    450                 455                 460

His Lys Lys Lys Gln Glu Asp Leu Lys Glu Ile Leu Gly Glu Asn Asp
465                 470                 475                 480

Ile Ile Asn Glu Gln Leu Ser Leu Phe Asp Asn His Asp Asp Ile Ala
                485                 490                 495


<210> 72
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 72
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60
```

```
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile His Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270
```

```
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 73
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 73

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Val Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Ile Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Ile Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220
```

```
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Asn Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Asn Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Glu
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430
```

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
435 440 445

Tyr Leu Asp
450

<210> 74
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 74
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
85 90 95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
100 105 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
115 120 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145 150 155 160

```
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr
                165                 170                 175

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380
```

```
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln
385                 390                 395                 400


Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415


Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450
```

<210> 75
<211> 473
<212> PRT
<213> Bacillus sp.

<400> 75

```
Met Ile Ile Ala Arg Lys Ile Lys Leu Ile Ile Ile Gly Glu Asp Arg
1               5                   10                  15


Asp Thr Gln Tyr Lys Phe Ile Arg Glu Glu Arg Tyr Lys Gln Asn Lys
            20                  25                  30


Ala Leu Asn Val Ala Met Asn His Leu Tyr Phe Leu His Val Ala Lys
        35                  40                  45


Glu Lys Ile Arg Leu Leu Asp Asn Lys Phe Leu Gln Asp Glu Lys Lys
    50                  55                  60


Leu Gln Glu Gly Ile Lys Lys Leu Tyr Ala Glu Lys Lys Val Ile Lys
65                  70                  75                  80


Asp Gly Lys Lys Arg Asn Glu Leu Glu Lys Lys Ile Glu Lys Gln Thr
                85                  90                  95


Asn Glu Leu Lys Lys Leu Arg Ser Lys Gly Asn Lys Glu Ala Asp Lys
            100                 105                 110
```

```
Ile Leu Gln Glu Ala Ile Lys Ile Asn Leu Ser Ser Thr Thr Arg Glu
        115                 120                 125

Val Ile Ser Lys Gln Phe Asp Leu Ile Ser Asp Thr Lys Asp Arg Ile
    130                 135                 140

Thr Gln Lys Val Tyr Gln Asp Phe Lys Ser Asp Leu Lys Asn Gly Leu
145                 150                 155                 160

Leu Ser Gly Glu Arg Val Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu
                165                 170                 175

Leu Ile Arg Gly Arg Ala Leu Asn Phe Tyr Arg Glu Gly Lys Asp Val
            180                 185                 190

Met Ile Lys Trp Phe Gly Gly Ile Ile Phe Lys Cys Met Leu Gly Gln
        195                 200                 205

His Lys Asn Asn Ala Gln Glu Leu Lys Ala Thr Leu Asn Lys Val Leu
    210                 215                 220

Glu Gly Ser Tyr Lys Val Cys Asp Ser Ser Ile Ser Val Gly Lys Glu
225                 230                 235                 240

Leu Ile Leu Asn Ile Ser Leu Asp Ile Gly Glu Val Asn Ser Asn Val
                245                 250                 255

Ser Cys Lys Lys Gly Arg Val Leu Gly Val Asp Leu Gly Met Lys Val
            260                 265                 270

Pro Ala Tyr Met Ser Ile Asn Asp Lys Pro Tyr Ile Arg Lys Ser Leu
        275                 280                 285

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

Arg Arg Asn Leu His Lys Thr Leu Val Ser Val Lys Gly Gly Lys Gly
305                 310                 315                 320
```

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

Asn Phe Ala Thr Thr Tyr Asn His Phe Leu Ser Tyr Asn Ile Val Lys
340 345 350

Phe Ala Lys Asp Asn Leu Ala Glu Gln Ile Asn Met Glu Phe Leu Ala
355 360 365

Leu Ala Gly Glu Asp Lys Asn Ile Ile Leu Arg Asn Trp Ser Tyr Tyr
370 375 380

Gln Leu Gln Gln Phe Val Glu Asp Lys Ala Lys Arg Glu Gly Ile Asp
385 390 395 400

Val Lys Tyr Val Asp Pro Tyr Arg Thr Ser Gln Met Cys Ser Lys Cys
405 410 415

Arg Asn Tyr Glu Pro Gly Gln Arg Glu Ser Gln Glu Lys Phe Ile Cys
420 425 430

Lys Ser Cys His Leu Glu Ile Asn Ala Asp Tyr Asn Ala Ser Gln Asn
435 440 445

Ile Ala His Ser Thr Lys Tyr Ile Thr Asn Lys Asn Gln Ser Glu Tyr
450 455 460

Phe Lys Lys Leu Gln His Thr Thr Glu
465 470

<210> 76
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 76
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
20 25 30

```
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255
```

```
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 77
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 77
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1   5   10   15

Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr
            20   25   30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35   40   45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50   55   60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65   70   75   80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85   90   95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100   105   110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115   120   125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145   150   155   160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180   185   190

```
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415
```

```
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450


<210> 78
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 78
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110


Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125


Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn
    130                 135                 140
```

```
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350
```

```
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 79
<211> 449
<212> PRT
<213> Bacillus sp.

<400> 79

```
Met Asp Lys Lys Ala Ser Lys Val Met Lys Tyr Glu Ile Ile Lys Pro
1               5                   10                  15

Val Asp Val Asp Trp Asp Val Phe Gly Ser Val Leu Arg Glu Leu Gln
            20                  25                  30

Phe Glu Ser Lys Asn Leu Met Asn Lys Thr Ile Gln Leu Cys Trp Glu
        35                  40                  45

Trp Gln Gly Phe Ser Ser Asp Tyr Lys Lys Glu Asn Gly Glu Tyr Pro
    50                  55                  60

Lys Leu Glu Asn His Thr Lys Tyr Lys Ala Ile Thr Gly Phe Val Tyr
65                  70                  75                  80
```

```
Asp Arg Met Lys Thr Gln Tyr Asn Lys Phe Asn Thr Gly Asn Met Ser
                85                  90                  95

Val Ser Ile Lys Ser Ala Thr Asp Lys Trp Lys Ser Asp Leu Lys Asp
            100                 105                 110

Ile Leu Arg Gly Glu Lys Ser Ile Pro Ser Tyr Lys Lys Asn Val Pro
        115                 120                 125

Ile Asp Ile His Gly Asn Ser Ile Lys Ile Lys Glu Val Asn Lys Asn
    130                 135                 140

Gly Ala Ile Leu Gln Leu Ser Leu Ile Ser Ser Thr Tyr Lys Lys Glu
145                 150                 155                 160

Leu Gly Ile Lys Asn Gly Phe Phe Asn Val Leu Ile Lys Ile Gly Asp
                165                 170                 175

Asn Ser Gln His Glu Ile Val Lys Arg Leu Phe Glu Gly Asp Tyr Lys
            180                 185                 190

Ile Ser Ala Ser Lys Ile Val Lys His Lys Tyr Lys Asn Lys Trp Phe
        195                 200                 205

Leu Asn Leu Thr Tyr Ser Phe Phe Leu Glu Lys Arg Glu Leu Asn Pro
    210                 215                 220

Asp Asn Ile Met Gly Ile Asp Val Gly Val Val Asn Ala Leu Tyr Met
225                 230                 235                 240

Ala Phe Asn Glu Ser Leu Ser Arg Tyr Ser Ile Glu Gly Gly Glu Ile
                245                 250                 255

Thr Lys Phe Arg Lys Gly Val Glu Ala Arg Arg Lys Ser Leu Leu Arg
            260                 265                 270

Gln Gly Lys Tyr Cys Gly Glu Gly Arg Lys Gly Arg Gly Arg Ala Thr
        275                 280                 285

Arg Ile Lys Pro Ile Glu Lys Leu Ser Gln Arg Val Asp Asn Phe Lys
    290                 295                 300
```

```
Asp Ser Cys Asn His Lys Tyr Ser Lys Tyr Val Ile Asp Met Ala Leu
305                 310                 315                 320


Lys His Asn Cys Gly Thr Ile Gln Met Glu Asp Leu Thr Gly Ile Ala
                325                 330                 335


Glu Gly Glu Lys Lys Ser Ser Phe Leu Gly Asn Trp Thr Tyr Phe Asp
            340                 345                 350


Leu Gln Glu Lys Ile Thr Tyr Lys Ala Lys Glu His Gly Ile Lys Val
        355                 360                 365


Val Lys Ile Lys Pro Lys Tyr Thr Ser Gln Arg Cys Ser Lys Cys Gly
    370                 375                 380


Phe Ile Ser Lys Asp Ser Arg Pro Asp Gln Ala Thr Phe Glu Cys Ile
385                 390                 395                 400


Lys Cys Asn Phe Lys Thr Ser Ala Asp Tyr Asn Ala Ala Arg Asn Ile
                405                 410                 415


Ala Met Lys Asp Ile Glu Lys Ile Ile Ala Glu Gln Leu Lys Val Gln
            420                 425                 430


Glu Lys Ala Lys Lys Leu Thr Lys Lys Gln Leu Ala Asn Ile Leu Asp
        435                 440                 445


Asp
```

<210> 80
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 80

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30
```

```
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240
```

```
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Glu Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 81
<211> 444
<212> PRT
<213> Bacillus sp.

<400> 81
Met Lys Tyr Gln Ile Leu Cys Pro Met Asn Val Asp Trp Thr Ile Phe
1   5   10   15

Glu Lys His Leu Arg Asn Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn
            20   25   30

Arg Thr Ile Gln Gln Leu Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr
        35   40   45

Phe Lys Glu Arg Gly Thr Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys
    50   55   60

Thr Gln Lys Lys Ile Asp Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys
65   70   75   80

Tyr Pro Asp Ile His Lys Gly Asn Met Ser Thr Thr Leu Gln Lys Ile
                85   90   95

Ile Lys Thr Trp Lys Ser Arg Arg Asn Glu Ile Arg Lys Gly Glu Met
            100   105   110

Ser Ile Pro Ser Phe Arg Asn Arg Ile Pro Ile Asp Leu His Asn Asn
        115   120   125

Ser Val Asp Ile Ile Lys Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile
    130   135   140

Ser Leu Phe Ser Arg Asp Phe His Lys Glu Asn Gly Asp Val Pro Lys
145   150   155   160

Gly Lys Ile Phe Val Lys Leu Gly Thr Gln Lys Gln Lys Ser Met Lys
                165   170   175

Val Ile Leu Asp Arg Leu Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys
            180   185   190

```
Met Ile His Lys Tyr Lys Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys
        195                 200                 205

Phe Asn Ala Ile Lys Glu Asn Lys Phe Asp Lys Glu Leu Ile Met Gly
    210                 215                 220

Ile Asp Met Gly Gly Ile Asn Thr Val Tyr Phe Ala Phe Asn Glu Gly
225                 230                 235                 240

Phe Ile Arg Ser Asn Ile Lys Ser Asp Glu Ile Lys Met Phe Asn Glu
                245                 250                 255

Arg Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr Cys
            260                 265                 270

Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro Ile
        275                 280                 285

Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn His
    290                 295                 300

Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys Gly
305                 310                 315                 320

Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Lys Val
                325                 330                 335

Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln
            340                 345                 350

Ala Glu Ile Tyr Gly Ile Glu Val Ile Lys Val Val Pro Ala Tyr Thr
        355                 360                 365

Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys
    370                 375                 380

Thr Gln Ala Met Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His Ala
385                 390                 395                 400
```

Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile
405 410 415

Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys Cys
420 425 430

Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
435 440

<210> 82
<211> 448
<212> PRT
<213> Bacillus sp.

<400> 82
Met Thr Phe Leu Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu
1 5 10 15

Cys Pro Leu Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn
20 25 30

Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu
35 40 45

Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr
50 55 60

Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp
65 70 75 80

Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys
85 90 95

Gly Asn Met Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser
100 105 110

Arg Arg Asn Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg
115 120 125

Asn Arg Ile Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys
130 135 140

```
Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp
145                 150                 155                 160


Phe His Lys Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys
                165                 170                 175


Leu Ala Thr Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu
            180                 185                 190


Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys
        195                 200                 205


Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu
    210                 215                 220


Asn Lys Phe Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile
225                 230                 235                 240


Asn Thr Val Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile
                245                 250                 255


Lys Ser Asp Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln
            260                 265                 270


Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg
        275                 280                 285


Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn
    290                 295                 300


Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys
305                 310                 315                 320


His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys
                325                 330                 335


Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys
            340                 345                 350


Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala
        355                 360                 365
```

Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys
370 375 380

Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr
385 390 395 400

Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp
405 410 415

Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His
420 425 430

Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
435 440 445

<210> 83
<211> 444
<212> PRT
<213> Bacillus sp.

<400> 83
Met Lys Tyr Gln Ile Leu Cys Pro Met Asn Val Asp Trp Thr Ile Phe
1 5 10 15

Glu Lys His Leu Arg Asn Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn
20 25 30

Arg Thr Ile Gln Gln Leu Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr
35 40 45

Phe Lys Glu Arg Gly Thr Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys
50 55 60

Thr Gln Lys Lys Ile Asp Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys
65 70 75 80

Tyr Pro Asp Ile His Lys Gly Asn Met Ser Thr Thr Leu Gln Lys Ile
85 90 95

Ile Lys Thr Trp Lys Ser Arg Arg Asn Glu Ile Arg Lys Gly Glu Met
100 105 110

```
Ser Ile Pro Ser Phe Arg Asn Arg Ile Pro Ile Asp Leu His Asn Asn
        115                 120                 125

Ser Val Asp Ile Thr Lys Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile
    130                 135                 140

Ser Leu Phe Ser Arg Asp Phe His Lys Glu Asn Asp Asp Val Pro Lys
145                 150                 155                 160

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

Val Ile Leu Asp Arg Leu Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys
            180                 185                 190

Met Ile His Lys Tyr Lys Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys
        195                 200                 205

Phe Asn Ala Ile Lys Glu Asn Lys Phe Asp Lys Glu Leu Ile Met Gly
    210                 215                 220

Ile Asp Leu Gly Gly Ile Asn Thr Val Tyr Phe Ala Phe Asn Glu Gly
225                 230                 235                 240

Phe Ile Arg Ser Asn Ile Lys Ser Asp Glu Ile Lys Met Phe Asn Glu
                245                 250                 255

Arg Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr Cys
            260                 265                 270

Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro Ile
        275                 280                 285

Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn His
    290                 295                 300

Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys Gly
305                 310                 315                 320
```

```
Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg Ile
                325                 330                 335

Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln
            340                 345                 350

Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr Thr
        355                 360                 365

Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys
    370                 375                 380

Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His Ala
385                 390                 395                 400

Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile
                405                 410                 415

Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys Cys
            420                 425                 430

Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440
```

<210> 84
<211> 448
<212> PRT
<213> Bacillus sp.

<400> 84

```
Met Thr Phe Leu Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu
1               5                   10                  15

Cys Pro Leu Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn
            20                  25                  30

Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu
        35                  40                  45

Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr
    50                  55                  60
```

```
Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp
65                  70                  75                  80

Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys
                85                  90                  95

Gly Asn Met Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser
            100                 105                 110

Arg Arg Asn Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg
        115                 120                 125

Asn Arg Ile Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys
    130                 135                 140

Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp
145                 150                 155                 160

Phe His Lys Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys
                165                 170                 175

Leu Ala Thr Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu
            180                 185                 190

Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys
        195                 200                 205

Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu
    210                 215                 220

Asn Lys Phe Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile
225                 230                 235                 240

Asn Thr Val Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile
                245                 250                 255

Lys Ser Asp Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln
            260                 265                 270

Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg
        275                 280                 285
```

```
Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn
    290                 295                 300

Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys
305                 310                 315                 320

His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys
                325                 330                 335

Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys
            340                 345                 350

Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala
        355                 360                 365

Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys
    370                 375                 380

Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr
385                 390                 395                 400

Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp
                405                 410                 415

Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His
            420                 425                 430

Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445
```

<210> 85
<211> 486
<212> PRT
<213> Bacillus sp.

<400> 85

```
Met Ile Thr Ser Arg Lys Ile Arg Leu Ser Ile Val Ser Asp Asn Ser
1               5                   10                  15

Thr Glu Ala Tyr Asn Phe Ile Arg Lys Glu Met Arg Glu Gln Asn Lys
            20                  25                  30
```

```
Ala Leu Asn Val Ser Met Asn His Leu Tyr Phe Asn Ser Ile Ala Arg
        35                  40                  45

Gln Lys Ile Leu Leu Ala Asp Glu Ala Tyr Gln Gln Lys Leu Asn Asp
    50                  55                  60

Ala Ile Asn Ser Gln Glu Lys Ser Phe Ala Ala Leu Lys Glu Ile Glu
65                  70                  75                  80

Gln Lys Gln Cys Ile Glu Asp Ser Glu Lys Lys Gln Val Leu Lys Glu
                85                  90                  95

Arg Leu Ile Lys Thr Lys Asn Ala Tyr Glu Lys Ala Lys Glu Lys Val
            100                 105                 110

Asn Asn Leu Arg Lys Ser Arg Ser Lys Asp Ser Phe Gln Glu Tyr Lys
        115                 120                 125

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

Ser Gln Phe Asn Leu His Ser Asp Thr Lys Asp Arg Leu Thr Met Ile
145                 150                 155                 160

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

Arg Ser Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu Tyr Ile Arg Ser
            180                 185                 190

Arg Asn Thr Val Leu Tyr Lys Glu Gly Asn Glu Phe Phe Ile Lys Trp
        195                 200                 205

Ile Lys Gly Ile Val Phe Lys Cys Ile Leu Gly Val Lys Asn Gln Asn
    210                 215                 220

Lys Thr Glu Leu Tyr Lys Thr Leu Glu Cys Val Leu Ala Gly Ile Tyr
225                 230                 235                 240
```

```
Lys Ile Cys Asp Ser Ser Met Asn Phe Asn Gln Gln Asn Lys Leu Ile
                245                 250                 255

Leu Asn Leu Thr Leu Asp Met Pro Asp Lys Ser Glu His Arg Lys Val
            260                 265                 270

Pro Glu Arg Ile Ala Gly Ile Asp Leu Gly Leu Lys Ile Pro Ala Tyr
        275                 280                 285

Phe Ala Val Asn Asp Val Ser Tyr Ile Arg Gln Ala Ile Gly Lys Ile
    290                 295                 300

Glu Asp Phe Leu Lys Val Arg Thr Ser Ile Gln Ser Gln Lys Arg Ser
305                 310                 315                 320

Leu Glu His Ala Leu Gln Ser Ser Lys Gly Gly Lys Gly Arg Lys Lys
                325                 330                 335

Lys Leu Lys Ala Leu Gly Gln Phe Lys Glu Lys Glu Lys Ser Tyr Ile
            340                 345                 350

Thr Thr Tyr Asn His Phe Ile Ser Lys Lys Ile Ile Ser Leu Ala Ile
        355                 360                 365

Gln Tyr Gly Val Val Gln Ile Asn Leu Glu Leu Leu Thr Leu Lys Glu
    370                 375                 380

Thr Gln Lys Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu Gln
385                 390                 395                 400

Gln Phe Ile Glu Tyr Lys Ala Glu Arg Ala Gly Ile Leu Val Lys Tyr
                405                 410                 415

Val Asp Pro Phe His Thr Ser Gln Thr Cys Ser Lys Cys Gly His Tyr
            420                 425                 430

Glu Asp Gly Gln Arg Glu Lys Gln Asp Thr Phe Cys Cys Lys Ser Cys
        435                 440                 445

Gly Phe Thr Glu Asn Ala Asp Tyr Asn Ala Ala Arg Asn Ile Ala Ala
    450                 455                 460
```

```
Ser Thr Arg Tyr Ile Thr Asn Lys Glu Glu Ser Glu Tyr Tyr Arg Lys
465                 470                 475                 480


Asn Asn Asn Glu Ile Ala
                485
```

<210> 86
<211> 264
<212> PRT
<213> Bacillus sp.

<400> 86

```
Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln Thr Tyr Ser Lys Gly
1               5                   10                  15


Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp Tyr Leu Ser Ile Thr
            20                  25                  30


Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe Glu Lys Glu Leu Ile
        35                  40                  45


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


Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp Glu Ile Ile Arg Gln
65                  70                  75                  80


Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg
                85                  90                  95


Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser
            100                 105                 110


Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn
        115                 120                 125


Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met
    130                 135                 140


Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp
145                 150                 155                 160
```

```
Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His
                165                 170                 175

Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys
            180                 185                 190

Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr
        195                 200                 205

Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala
    210                 215                 220

Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln
225                 230                 235                 240

Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr
                245                 250                 255

Ile Glu Glu Leu Gly Tyr Leu Asp
            260
```

<210> 87
<211> 438
<212> PRT
<213> Bacillus sp.

<400> 87

```
Met Lys Tyr Gln Ile Leu Cys Pro Leu Asn Val Asp Trp Thr Ile Phe
1               5                   10                  15

Glu Lys His Leu Arg Asn Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn
            20                  25                  30

Arg Thr Ile Gln Gln Leu Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr
        35                  40                  45

Phe Lys Glu Arg Gly Thr Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys
    50                  55                  60

Thr Gln Lys Lys Ile Asp Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys
65                  70                  75                  80
```

```
Tyr Pro Asp Ile His Lys Gly Asn Met Ser Thr Thr Leu Gln Lys Ile
                85                  90                  95

Ile Lys Thr Trp Lys Ser Arg Arg Asn Glu Ile Arg Lys Gly Glu Met
            100                 105                 110

Ser Ile Pro Ser Phe Arg Asn Arg Ile Pro Ile Asp Leu His Asn Asn
        115                 120                 125

Ser Val Asp Ile Thr Lys Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile
    130                 135                 140

Ser Leu Phe Ser Arg Asp Phe His Lys Glu Asn Asp Asp Val Pro Lys
145                 150                 155                 160

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

Val Ile Leu Asp Arg Leu Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys
            180                 185                 190

Met Ile His Lys Tyr Lys Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys
        195                 200                 205

Phe Asn Ala Ile Lys Glu Asn Lys Phe Asp Lys Glu Leu Ile Met Gly
    210                 215                 220

Ile Asp Leu Gly Gly Ile Asn Thr Val Tyr Ser Ala Phe Asn Glu Gly
225                 230                 235                 240

Phe Ile Arg Ser Asn Ile Lys Ser Asp Glu Ile Ile Arg Gln Arg Arg
                245                 250                 255

Ile Asn Leu Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly
            260                 265                 270

Lys Gly Arg Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys
        275                 280                 285
```

```
Ile Ala Lys Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile
    290                 295                 300

Val Lys Gln Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu
305                 310                 315                 320

Leu Lys Gly Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr
                325                 330                 335

Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile
            340                 345                 350

Glu Val Ile Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln
        355                 360                 365

Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu
    370                 375                 380

Cys Lys Gln Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys
385                 390                 395                 400

Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala
                405                 410                 415

Val Gln Ser Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu
            420                 425                 430

Glu Leu Gly Tyr Leu Asp
        435
```

<210> 88
<211> 450
<212> PRT
<213> Bacillus sp.

<400> 88

```
Met Ser Thr Val Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1               5                   10                  15

Asn Ile Glu Trp Lys Val Phe Glu Thr Tyr Leu Arg Thr Leu Thr Tyr
            20                  25                  30
```

```
Gln Val Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Leu Trp Glu Phe
        35                  40                  45

Asp Asn Gln Ser Leu Asn His Phe Lys Glu Asn Gly Val Tyr Pro Ser
    50                  55                  60

Thr Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65                  70                  75                  80

Tyr Asp Gln Leu Lys Phe Glu Tyr Leu Asp Ile Asn Lys Ala Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Thr Ile Lys Thr Trp Asn Ser Arg Lys Lys
            100                 105                 110

Glu Ile Trp Ser Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Asn Leu
        115                 120                 125

Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Thr Lys Glu Lys Ser
    130                 135                 140

Gly Asp Tyr Ile Ala Asn Leu Ser Leu Phe Ser Ser Ser Phe Ile Arg
145                 150                 155                 160

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

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asp Gly Thr
            180                 185                 190

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

Leu Ser Val Thr Tyr Lys Ala Asn Ala Ile Glu Glu Ser Lys Phe Asp
    210                 215                 220

Glu Asp Phe Ile Met Gly Ile Asp Met Gly Lys Val Asn Val Leu Tyr
225                 230                 235                 240

Phe Ala Phe Asn Lys Gly Phe Ile Arg Gly Ser Ile Ser Gly Glu Glu
                245                 250                 255
```

```
Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Glu Asn Arg Ile Gly Lys Gly Arg Lys
        275                 280                 285

Lys Arg Ile Lys Pro Ile Asp Val Ile Asn Asp Lys Val Ala Arg Phe
    290                 295                 300

Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320

Leu Lys Tyr Lys Cys Gly Cys Ile Gln Leu Glu Asn Leu Gln Gly Ile
                325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350

Glu Lys Ile Glu Asn Gln Ala Asn Gln Tyr Gly Ile Lys Val Val Lys
        355                 360                 365

Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
    370                 375                 380

His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln Gln Cys
385                 390                 395                 400

Asn Leu Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Ile
                405                 410                 415

Tyr Asn Ile Glu Lys Leu Ile Gln Lys Gln Leu Lys Leu Gln Glu Lys
            420                 425                 430

Leu Asn Ser Lys Lys Tyr Thr Glu Gln Tyr Ile Glu Gln Ile Glu Asn
        435                 440                 445

Ile Asn
    450
```

<210> 89
<211> 487
<212> PRT
<213> Bacillus sp.

<400> 89
Met Ile Thr Ser Arg Lys Ile Arg Leu Ser Ile Val Ser Asp Asn Ser
1   5   10   15

Thr Glu Ala Tyr Asn Phe Ile Arg Lys Glu Met Arg Glu Gln Asn Lys
            20   25   30

Ala Leu Asn Val Ala Met Asn His Leu Tyr Phe Asn Tyr Val Ala Arg
        35   40   45

Gln Lys Ile Ser Phe Ala Asp Lys Ala Tyr Gln His Lys Leu Asn Lys
    50   55   60

Ala Ile Glu Ala Gln Gln Lys Ser Phe Ile Leu Leu Lys Glu Leu Glu
65   70   75   80

Gln Lys Gln Ile Val Glu Thr Asp Ser Ser Lys Lys Ser Ile Leu Lys
                85   90   95

Glu Arg Leu Ile Lys Ser Lys Thr Thr Tyr Glu Lys Thr Lys Glu Lys
            100   105   110

Leu Ser Ser Leu Arg Lys Ala Arg Asn Lys Glu Leu Phe Gln Glu Tyr
        115   120   125

Asn Asn Met Ile Gly Gln Leu Glu Asp Thr His Leu Arg Asp Ile Val
    130   135   140

Ser Ser Gln Phe Asn Leu Leu Ser Asp Thr Lys Asp Arg Leu Thr Lys
145   150   155   160

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

Asn Arg Ser Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu Tyr Ile Arg
            180   185   190

```
Gly Arg Thr Leu Asn Leu Phe Lys Glu Ser Glu Glu Phe Phe Ile Lys
        195                 200                 205

Trp Thr Lys Lys Ile Val Phe Lys Cys Val Leu Gly Val Lys Tyr Gln
    210                 215                 220

Asn Lys Thr Glu Leu Tyr Lys Thr Leu Glu Cys Ile Leu Thr Gly Glu
225                 230                 235                 240

Tyr Glu Leu Cys Asp Ser Ser Met Asn Phe Asn Gln Gln Asn Lys Leu
                245                 250                 255

Ile Leu Asn Leu Ala Leu Asn Ile Pro Glu Lys Ser Glu Asn Arg Lys
            260                 265                 270

Val Pro Gly Arg Ile Ala Gly Ile Asp Leu Gly Leu Lys Ile Pro Ala
        275                 280                 285

Tyr Phe Ala Val Asn Asp Val Pro Tyr Ile Arg Lys Ala Leu Gly Lys
    290                 295                 300

Ile Glu Asp Phe Leu Lys Val Arg Thr Asn Ile Gln Ser Gln Lys Arg
305                 310                 315                 320

Ser Leu Gln Arg Ala Leu Gln Ser Ser Lys Gly Gly Lys Gly Arg Lys
                325                 330                 335

Lys Lys Leu Lys Ala Leu Asp Gln Phe Lys Glu Lys Glu Lys Asn Tyr
            340                 345                 350

Ile Thr Thr Tyr Asn His Phe Ile Ser Lys Lys Ile Ile Ser Leu Ala
        355                 360                 365

Val Gln Tyr Gly Val Glu Gln Ile Asn Leu Glu Leu Leu Thr Leu Lys
    370                 375                 380

Glu Thr Gln Lys Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
385                 390                 395                 400

Gln Gln Phe Ile Glu Tyr Lys Ala Asn Arg Glu Glu Ile Leu Ile Lys
                405                 410                 415
```

```
Tyr Val Asp Pro Phe His Thr Ser Gln Thr Cys Ser Lys Cys Gly His
            420                 425                 430

Tyr Glu Asp Gly Gln Arg Glu Lys Gln Asp Thr Phe Cys Cys Lys Ser
        435                 440                 445

Cys Gly Phe Ile Asp Asn Ala Asp Tyr Asn Ala Ala Arg Asn Ile Ala
    450                 455                 460

Ala Ser Thr Arg Tyr Ile Thr Asn Lys Glu Glu Ser Glu Tyr Tyr Arg
465                 470                 475                 480

Lys Asn Asn Asn Glu Ile Ala
                485
```

<210> 90
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 90

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110
```

```
Glu Ile Arg Lys Gly Lys Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Gly Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320
```

```
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 91
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 91

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45
```

```
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Lys Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270
```

```
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Gly Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 92
<211> 496
<212> PRT
<213> Bacillus sp.

<400> 92
Met Ile Leu Thr Arg Lys Ile Gln Leu Val Ile Ile Ser Glu Asn Gln
1   5   10   15

Lys Glu Gly Tyr Ser Leu Ile Arg Asn Glu Ile Gln Glu Gln Tyr Lys
            20   25   30

Ala Leu Asn Leu Ser Tyr Asn His Leu Tyr Phe Glu His Asn Ala Ile
        35   40   45

Gln Lys Leu Lys Trp Asn Asp Glu Asp Tyr Lys Gln Lys Arg Ser Lys
    50   55   60

Leu Gln Glu Leu Val Asn Lys Lys Tyr Glu Glu Tyr Gln Lys Val Lys
65   70   75   80

Asn Leu Glu Lys Lys Glu Ala Leu Arg Glu Ala Tyr Asn Lys Lys Lys
                85   90   95

Gln Glu Leu Tyr Lys Phe Glu Lys Glu Cys Asn Glu Glu Ala Arg Lys
            100   105   110

Val Tyr Gln Gln Val Val Gly Phe Thr Gln Gln Thr Arg Val Arg Asn
        115   120   125

Leu Ile Asn Arg Glu Cys Asn Leu Met Ser Asp Thr Lys Asp Gly Ile
    130   135   140

Thr Ser Lys Val Thr Gln Asp Tyr Lys Asn Asp Cys Lys Ala Gly Leu
145   150   155   160

Leu Ile Gly Lys Arg Ser Leu Arg Asn Tyr Lys Lys Asp Asn Pro Leu
                165   170   175

Leu Val Arg Gly Arg Ser Leu Lys Phe Tyr Lys Glu Asn Gly Glu Tyr
            180   185   190

Phe Ile Lys Trp Asn Lys Gly Thr Ile Phe Lys Cys Ile Leu His Ile
        195   200   205

```
Arg Lys Lys Asn Ile Val Glu Leu Gln Ser Val Leu Glu Asn Val Leu
    210                 215                 220

Gln Gly Ala Tyr Lys Val Cys Asp Ser Ser Ile Ala Phe Asn Asn Arg
225                 230                 235                 240

Asp Met Ile Leu Asn Leu Thr Leu Asn Ile Pro Asn Lys Glu Thr Gln
                245                 250                 255

Asp Tyr Ile Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile
            260                 265                 270

Pro Ala Tyr Val Ser Leu Ser Asp Lys Val Tyr Val Arg Lys Gly Ile
        275                 280                 285

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

Arg Arg Gln Leu Gln Glu Ser Leu Ser Ala Val Lys Gly Gly Lys Gly
305                 310                 315                 320

Arg Lys Lys Lys Leu Lys Ala Leu Glu His Leu Lys Glu Lys Glu Ala
                325                 330                 335

Asn Phe Ala Lys Thr Tyr Asn His Phe Leu Ser Thr Gln Ile Val Thr
            340                 345                 350

Phe Ala Val Lys Asn His Ala Gly Gln Ile Asn Met Glu Phe Leu Glu
        355                 360                 365

Phe Asp Lys Met Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr
    370                 375                 380

Gln Leu Gln Met Met Ile Glu Tyr Lys Ala Lys Arg Glu Gly Ile Ile
385                 390                 395                 400

Ile Lys Tyr Val Asp Ala Tyr Leu Thr Ser Gln Thr Cys Ser Lys Cys
                405                 410                 415

Asp Tyr Tyr Glu Glu Gly Gln Arg Glu Thr Gln Glu Arg Phe Met Cys
            420                 425                 430
```

```
Lys Ser Cys Gly Phe Glu Val Asn Ala Asp Tyr Asn Ala Ser Gln Asn
        435                 440                 445


Ile Ala Lys Ser Thr Arg Tyr Ile Ser Asp Ser Thr Glu Ser Glu Tyr
    450                 455                 460


His Lys Lys Lys Gln Glu Ala Leu Lys Gly Ile Leu Gly Glu Asn Asp
465                 470                 475                 480


Thr Ile Asn Glu Gln Leu Ser Leu Phe Asn Asn Cys Asp Asp Ile Ala
                485                 490                 495


<210> 93
<211> 445
<212> PRT
<213> Bacillus sp.

<400> 93
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110


Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125
```

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr
            260                 265                 270

Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro
        275                 280                 285

Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn
    290                 295                 300

His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys
305                 310                 315                 320

Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg
                325                 330                 335
```

Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn
340 345 350

Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr
355 360 365

Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
370 375 380

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
385 390 395 400

Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
405 410 415

Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
420 425 430

Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
435 440 445

<210> 94
<211> 506
<212> PRT
<213> Bacillus sp.

<400> 94
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1 5 10 15

Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
20 25 30

Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
35 40 45

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

Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
65 70 75 80

```
Lys Tyr Arg Glu Val Ser Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
                85                  90                  95

Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
            100                 105                 110

Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
        115                 120                 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
    130                 135                 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145                 150                 155                 160

Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

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

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300
```

```
Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Asn Val Arg Thr Gln Leu Gln Asn Gln Arg Lys
                325                 330                 335

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445

Tyr Glu Lys Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460

Lys Glu Cys Lys Gly Tyr Ala His Lys Val Asn Ala Asp Phe Asn Ala
465                 470                 475                 480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495

Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505
```

<210> 95
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 95
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1   5   10   15

Asn Val Asp Trp Ile Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20   25   30

Gln Ile Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35   40   45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50   55   60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65   70   75   80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85   90   95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100   105   110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115   120   125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145   150   155   160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180   185   190

```
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Lys Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Glu Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Asn Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
305                 310                 315                 320

Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
                405                 410                 415
```

```
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Ala Gln Ser
            420                 425                 430


Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445


Tyr Leu Asp
    450


<210> 96
<211> 445
<212> PRT
<213> Bacillus sp.

<400> 96
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
1               5                   10                  15


Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30


Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45


Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60


Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80


Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95


Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110


Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125


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

```
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr
            260                 265                 270

Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro
        275                 280                 285

Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn
    290                 295                 300

His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys
305                 310                 315                 320

Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg
                325                 330                 335

Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn
            340                 345                 350
```

```
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr
        355                 360                 365

Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
    370                 375                 380

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
385                 390                 395                 400

Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
                405                 410                 415

Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
            420                 425                 430

Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
        435                 440                 445
```

<210> 97
<211> 438
<212> PRT
<213> Bacillus sp.

<400> 97

```
Met Lys Tyr Gln Ile Leu Cys Pro Leu Asn Val Asp Trp Thr Ile Phe
1               5                   10                  15

Glu Lys His Leu Arg Asn Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn
            20                  25                  30

Arg Thr Ile Gln Gln Leu Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr
        35                  40                  45

Phe Lys Glu Arg Gly Thr Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys
    50                  55                  60

Thr Gln Lys Lys Ile Asp Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys
65                  70                  75                  80

Tyr Pro Asp Ile His Lys Gly Asn Met Ser Thr Thr Leu Gln Lys Ile
                85                  90                  95
```

```
Ile Lys Thr Trp Lys Ser Arg Arg Asn Glu Ile Arg Lys Gly Glu Met
            100                 105                 110

Ser Ile Pro Ser Phe Arg Asn Arg Ile Pro Ile Asp Leu His Asn Asn
        115                 120                 125

Ser Val Asp Ile Thr Lys Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile
    130                 135                 140

Ser Leu Phe Ser Arg Asp Phe His Lys Glu Asn Asp Asp Val Pro Lys
145                 150                 155                 160

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

Val Ile Leu Asp Arg Leu Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys
            180                 185                 190

Met Ile His Lys Tyr Lys Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys
        195                 200                 205

Phe Asn Ala Ile Lys Glu Asn Lys Phe Asp Lys Glu Leu Ile Met Gly
    210                 215                 220

Ile Asp Leu Gly Gly Ile Asn Thr Val Tyr Ser Ala Phe Asn Glu Gly
225                 230                 235                 240

Phe Ile Arg Ser Asn Ile Lys Ser Asp Glu Ile Ile Arg Gln Arg Arg
                245                 250                 255

Ile Asn Leu Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly
            260                 265                 270

Lys Gly Arg Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys
        275                 280                 285

Ile Ala Lys Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile
    290                 295                 300

Val Lys Gln Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu
305                 310                 315                 320
```

```
Leu Lys Gly Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr
                325                 330                 335

Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile
            340                 345                 350

Glu Val Ile Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln
        355                 360                 365

Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu
    370                 375                 380

Cys Lys Gln Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys
385                 390                 395                 400

Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala
                405                 410                 415

Val Gln Ser Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu
            420                 425                 430

Glu Leu Gly Tyr Leu Asp
        435
```

<210> 98
<211> 276
<212> PRT
<213> Bacillus sp.

<400> 98

```
Met Pro Gln Phe Lys Ser Lys Lys Asn Lys Ile Lys Ser Tyr Thr Thr
1               5                   10                  15

Lys Cys Thr Asn His Asn Ile Ala Ile Val Gly His Thr Ile Lys Leu
            20                  25                  30

Pro Lys Leu Gly Phe Val Arg Phe Ala Lys Ser Arg Glu Val Ser Gly
        35                  40                  45

Arg Ile Leu Asn Ala Thr Ile Arg Arg Asn Pro Ser Gly Arg His Phe
    50                  55                  60
```

```
Val Ser Ile Leu Thr Glu Thr Glu Val Gln Pro Val Glu Lys Thr Gly
65                  70                  75                  80

Ser Ser Val Gly Ile Asp Val Gly Leu Lys Asp Tyr Ala Ile Leu Ser
                85                  90                  95

Asp Gly Ile Thr Tyr Lys Asn Pro Lys Phe Phe Arg Thr Leu Glu Glu
            100                 105                 110

Lys Leu Ala Lys Ala Gln Arg Ile Leu Ser Arg Arg Thr Lys Gly Ser
        115                 120                 125

Ser Asn Trp Asn Lys Gln Arg Val Lys Val Ala Arg Ile His Glu His
    130                 135                 140

Ile Ala Asn Ala Arg Ala Asp Tyr Leu His Lys Leu Ser Thr Glu Ile
145                 150                 155                 160

Ile Lys Asn His Asp Val Ile Gly Met Glu Asp Leu Gln Val Arg Asn
                165                 170                 175

Met Leu Lys Asn Arg Lys Leu Ala Lys Ala Ile Gly Glu Val Ser Trp
            180                 185                 190

Ser Gln Phe Arg Thr Met Leu Glu Tyr Lys Ala Lys Trp Tyr Gly Lys
        195                 200                 205

Lys Val Val Met Val Ser Lys Thr Phe Ala Ser Ser Gln Leu Cys Ser
    210                 215                 220

Asn Cys Glu Tyr Lys Asn Lys Asp Val Lys Asn Leu Thr Ile Arg Glu
225                 230                 235                 240

Trp Asp Cys Pro Ser Cys Gly Ala His His Asp Arg Asp Arg Asn Ala
                245                 250                 255

Ala Leu Asn Leu Lys Asn Glu Ala Ile Arg Leu Leu Thr Val Gly Thr
            260                 265                 270
```

Thr Gly Ile Ala
        275

<210> 99
<211> 405
<212> PRT
<213> Bacillus sp.

<400> 99
Met Ala Arg Thr Thr Val Lys Lys Arg Asn Glu Glu Leu Ala Lys Glu
1   5   10   15

Asn Lys Val Leu Arg His Tyr Gly Ile Lys Leu Arg Ala Tyr Pro Thr
            20   25   30

Pro Thr Gln Glu Glu Lys Ile Ala Lys Thr Met Gly Cys Ala Arg Phe
        35   40   45

Ala Phe Asn Phe Tyr Leu His Glu Lys Gln Glu Val Tyr Gln Leu Thr
    50   55   60

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

Lys Gln His Pro Ala Phe Asn Trp Leu Lys Glu Val Asp Lys Phe Ser
                85   90   95

Leu Glu Ser Ala Leu Glu Gln Val Asp Asp Ala Tyr Asp His Phe Phe
            100   105   110

Lys Gly Gln Asn Lys Phe Pro Lys Phe Lys Ser Lys His Thr Ser Lys
        115   120   125

Gln Ser Tyr Thr Thr Lys Glu Thr Asn Gly Asn Ile Ala Leu Asp Val
    130   135   140

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

Leu Ser Lys Lys His Arg Thr Met Phe Gln Lys Asn Gly Phe Thr Ala
                165   170   175

```
Lys Ile Lys Ser Ala Thr Val Thr Arg His Ser Ser Gly Gln Tyr Tyr
            180                 185                 190

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

Val Thr Ala Ile Pro Thr Asn Glu Ile Ile Gly Cys Asp Leu Gly Leu
    210                 215                 220

Thr His Phe Leu Ile Asp Ser Asn Gly Gln Lys Ile Glu Asn Pro Arg
225                 230                 235                 240

Tyr Leu Lys Lys Asn Leu Glu Lys Leu Ala Lys Phe Gln Arg Arg Leu
                245                 250                 255

Lys Tyr Lys Gln Ile Gly Ser Ala Asn Tyr Arg Lys His Ala Gln Lys
            260                 265                 270

Ile Ala Lys Leu His Leu His Ile Ser Asn Leu Arg Lys Asp Phe Leu
        275                 280                 285

His Lys Ile Ser Arg Lys Leu Val Asn Glu Asn Gln Val Ile Ile Leu
    290                 295                 300

Glu Asp Leu Asn Val Lys Gly Met Ile Arg Asn Lys Lys Leu Ala Arg
305                 310                 315                 320

Ser Ile Ala Asp Val Gly Trp Gly Met Phe Lys Thr Phe Val Ser Tyr
                325                 330                 335

Lys Ala Asn Trp Ala Asn Lys Leu Leu Ile Leu Ile Asp Arg Phe Phe
            340                 345                 350

Pro Ser Ser Lys Gln Cys Asn Gly Cys Lys Glu Thr Asn Pro Leu Leu
        355                 360                 365

Ser Leu Ser Asp Arg Val Trp Met Cys Pro Ser Cys Gly Thr His His
    370                 375                 380

Asp Arg Asp Asp Asn Ala Ala Phe Asn Ile Lys Glu Glu Gly Ile Arg
385                 390                 395                 400
```

Leu Leu Leu Asn Ala
405

<210> 100
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 100
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
85 90 95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
100 105 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
115 120 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145 150 155 160

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

```
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380
```

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385 390 395 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
405 410 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
420 425 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
435 440 445

Tyr Leu Asp
450

<210> 101
<211> 487
<212> PRT
<213> Bacillus sp.

<400> 101
Met Ile Thr Ser Arg Lys Ile Arg Leu Ala Ile Ile Ser Asp Asn Ser
1 5 10 15

Thr Glu Thr Tyr Asn Phe Ile Arg Gln Glu Met Arg Glu Gln Asn Lys
20 25 30

Ala Leu Asn Val Ala Met Asn His Leu Tyr Phe Asn Tyr Val Ala Arg
35 40 45

Gln Lys Ile Ser Phe Ala Asp Lys Ala Tyr Gln Gln Lys Leu Asn Lys
50 55 60

Ala Ile Glu Ala Gln Gln Lys Ser Phe Ile Leu Leu Lys Glu Leu Glu
65 70 75 80

Gln Lys Gln Ile Val Glu Thr Asp Ser Ser Lys Lys Ser Leu Leu Lys
85 90 95

Glu Arg Leu Ile Lys Thr Lys Thr Thr Tyr Glu Lys Thr Lys Glu Lys
100 105 110

```
Leu Ser Asn Leu Arg Lys Ala Arg Asn Lys Glu Leu Phe Gln Glu Tyr
        115                 120                 125

Asn Asn Ile Ile Gly Gln Leu Glu Glu Thr His Leu Arg Asp Ile Val
    130                 135                 140

Ser Ser Gln Phe Asn Leu Leu Ser Asp Thr Lys Asp Ser Leu Thr Lys
145                 150                 155                 160

Ile Ala Tyr Gln Asp Phe Lys Asn Asp Ile Val Glu Val Leu Ser Gly
                165                 170                 175

Asp Arg Ser Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu Tyr Ile Arg
            180                 185                 190

Gly Arg Ser Leu Asn Leu Phe Lys Glu Ser Asp Glu Phe Phe Ile Lys
        195                 200                 205

Trp Thr Lys Gly Ile Val Phe Lys Cys Val Leu Gly Val Lys Tyr Gln
    210                 215                 220

Asn Lys Thr Glu Leu Tyr Lys Thr Leu Glu Cys Ile Leu Ala Gly Glu
225                 230                 235                 240

Tyr Lys Ile Cys Asp Ser Ser Met Asn Phe Asn Gln Gln Lys Lys Leu
                245                 250                 255

Ile Leu Asn Leu Val Leu Asn Ile Pro Glu Lys Ser Glu Asn Arg Lys
            260                 265                 270

Val Pro Gly Arg Ile Ala Gly Ile Asp Leu Gly Leu Lys Ile Pro Ala
        275                 280                 285

Tyr Phe Ala Val Asn Asp Val Pro Tyr Ile Arg Lys Ala Ile Gly Asn
    290                 295                 300

Ile Glu Asp Phe Leu Lys Val Arg Thr Asn Ile His Ser Gln Gln Arg
305                 310                 315                 320

Ser Leu Gln Arg Ala Leu Gln Ser Ser Lys Gly Gly Lys Gly Arg Lys
                325                 330                 335
```

```
Lys Lys Leu Lys Ala Leu Asp Gln Phe Lys Glu Lys Glu Lys Asn Tyr
            340                 345                 350

Ile Thr Thr Tyr Asn His Phe Ile Ser Lys Lys Ile Ile Thr Leu Ala
        355                 360                 365

Ile Arg Tyr Gly Val Glu Gln Ile Asn Leu Glu Leu Leu Thr Leu Lys
    370                 375                 380

Asp Thr Gln Lys Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
385                 390                 395                 400

Gln Gln Phe Ile Glu Tyr Lys Ala Asn Arg Glu Gly Ile Leu Ile Lys
                405                 410                 415

Tyr Val Asp Pro Phe His Thr Ser Gln Thr Cys Ser Lys Cys Ser His
            420                 425                 430

Tyr Glu Asp Gly Gln Arg Glu Lys Gln Asp Thr Phe Cys Cys Lys Ser
        435                 440                 445

Cys Gly Phe Lys Asp Asn Ala Asp Tyr Asn Ala Ala Arg Asn Ile Ala
    450                 455                 460

Ala Ser Thr Lys Tyr Ile Ile Lys Lys Glu Glu Ser Glu Tyr Tyr Arg
465                 470                 475                 480

Gln Asn Asn Asn Glu Ile Ala
                485
```

<210> 102
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 102

```
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
1               5                   10                  15

Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
            20                  25                  30
```

```
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
        35                  40                  45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
    50                  55                  60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65                  70                  75                  80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
                85                  90                  95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
            100                 105                 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
        115                 120                 125

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

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
145                 150                 155                 160

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

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
    210                 215                 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225                 230                 235                 240
```

```
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
                245                 250                 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
        355                 360                 365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Leu Asp
    450
```

<210> 103
<211> 377
<212> PRT
<213> Bacillus sp.

<400> 103
Met Leu Val Asn Lys Ala Tyr Lys Phe Arg Ile Tyr Pro Asn Lys Glu
1   5   10   15

Gln Glu Ile Leu Ile Ala Lys Thr Ile Gly Cys Ser Arg Phe Val Phe
            20   25   30

Asn His Phe Leu Gly Met Trp Asn Asp Thr Tyr Lys Glu Thr Gly Lys
        35   40   45

Gly Leu Thr Tyr Asn Ser Cys Ser Ala Gln Leu Pro Gln Leu Lys Ile
    50   55   60

Glu Leu Glu Trp Leu Lys Glu Val Asp Ser Ile Ala Ile Gln Ser Ala
65   70   75   80

Leu Lys Asn Leu Val Asp Ala Tyr Asn Arg Phe Phe Lys Lys Gln Asn
                85   90   95

Asp Lys Pro Arg Phe Lys Ser Lys Lys Asn Asp Val Gln Ser Tyr Lys
            100   105   110

Thr Lys His Thr Asn Gly Asn Ile Ala Ile Val Asn Asn Lys Ile Lys
        115   120   125

Leu Pro Lys Leu Gly Phe Val Thr Phe Ala Lys Ser Arg Glu Val Asp
    130   135   140

Gly Arg Ile Met Asn Ala Thr Val Arg Arg Asn Ser Ser Gly Lys Tyr
145   150   155   160

Phe Val Ala Ile Leu Thr Glu Val Glu Ile Gln Pro Leu Lys Lys Ala
                165   170   175

Asp Ser Ala Ile Gly Ile Asp Leu Gly Ile Thr Asp Phe Ala Ile Leu
            180   185   190

```
Ser Asp Gly His Lys Ile Asp Asn Asn Lys Phe Thr Ser Lys Met Glu
        195                 200                 205

Lys Lys Leu Lys Arg Glu Gln Arg Lys Leu Ser Lys Arg Ala Leu Leu
    210                 215                 220

Ala Lys Asn Lys Gly Ile His Leu Leu Asp Ala Gln Asn Tyr Gln Lys
225                 230                 235                 240

Gln Lys Cys Lys Val Ala Arg Leu His Glu Arg Val Ile Asn Gln Arg
                245                 250                 255

Asp Asp Phe Leu Asn Lys Leu Ser Thr Glu Ile Ile Lys Asn His Asp
            260                 265                 270

Ile Ile Cys Ile Glu Asp Leu Asn Thr Lys Gly Met Leu Arg Asn His
        275                 280                 285

Lys Leu Ala Lys Ser Ile Ser Asp Val Ser Trp Ser Ala Phe Val Ser
    290                 295                 300

Lys Leu Glu Tyr Lys Ala Thr Trp Tyr Gly Lys Thr Ile Val Lys Val
305                 310                 315                 320

Ser Arg Trp Phe Pro Ser Ser Gln Ile Cys Ser Asp Cys Gly His His
                325                 330                 335

Asp Gly Lys Lys Ser Leu Glu Ile Arg Gly Trp Thr Cys Pro Ile Cys
            340                 345                 350

His Ala Asn His Asp Arg Asp Phe Asn Ala Ser Lys Asn Ile Leu Ala
        355                 360                 365

Glu Gly Leu Arg Thr Leu Ala Leu Val
    370                 375
```

<210> 104
<211> 450
<212> PRT
<213> Bacillus sp.

<400> 104
Met Ser Leu Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1   5   10   15

Asn Val Glu Trp Lys Val Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr
            20   25   30

Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe
        35   40   45

Asp Asn Leu Ser Leu Lys His Phe Lys Glu Thr Gly Glu Tyr Pro Ser
    50   55   60

Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65   70   75   80

Tyr Asp Gln Leu Lys Glu Glu Tyr Lys Asp Ile Asn Lys Ala Asn Met
                85   90   95

Ser Thr Thr Ile Gln Lys Thr Leu Lys Asn Trp Asn Ser Arg Lys Lys
            100   105   110

Glu Ile Trp Arg Gly Glu Met Ser Ile Pro Ser Phe Arg Ser Asp Leu
        115   120   125

Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Ile Lys Glu Lys Gly
    130   135   140

Gly Asp Tyr Ile Ala Ser Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
145   150   155   160

Glu Asn Asp Leu Pro Asn Gly Lys Ile Leu Met Lys Leu Ser Thr Arg
                165   170   175

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr
            180   185   190

Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Lys Lys Trp Tyr
        195   200   205

```
Leu Ser Ile Thr Tyr Lys Ser Asn Ile Lys Glu Glu Leu Lys Phe Asp
    210                 215                 220

Glu Asp Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
225                 230                 235                 240

Phe Ala Phe Asn Lys Ser Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
                245                 250                 255

Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Gly Asn Arg Ile Gly Lys Gly Arg Glu
        275                 280                 285

Lys Arg Ile Lys Pro Ile Asp Val Leu Asn Asp Lys Val Ala Lys Phe
    290                 295                 300

Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320

Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Lys Gly Ile
                325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350

Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Val Lys Val Val Lys
        355                 360                 365

Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
    370                 375                 380

His Lys Asp Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Leu Gln Cys
385                 390                 395                 400

Ser Phe Lys Ala His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Val
                405                 410                 415

Tyr Thr Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
            420                 425                 430
```

Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Ile Lys Asn
435 440 445

Ile Asn
450

<210> 105
<211> 451
<212> PRT
<213> Bacillus sp.

<400> 105
Met Ser Val Met Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Thr
1 5 10 15

Asn Val Asp Trp Ser Thr Phe Glu Lys Asn Leu Arg Asp Leu Thr Tyr
20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
35 40 45

Asp Ala Leu Ser Tyr Asp Tyr Phe Lys Glu Thr Gly Thr Ser Pro Thr
50 55 60

Val Gln Asp Leu Tyr Lys Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
65 70 75 80

Tyr His Val Leu Gln Ser Lys Tyr Pro Asp Val His Lys Gly Asn Met
85 90 95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Ser
100 105 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Gln Leu
115 120 125

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

Gly Asp Tyr Ile Ala Gly Leu Ser Leu Phe Ser Arg Asp Phe His Lys
145 150 155 160

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

Lys Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Ser Gly
            180                 185                 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
        195                 200                 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ser Ile Thr Glu Asn Lys Phe
    210                 215                 220

Asp Glu Asn Leu Ile Met Gly Ile Asp Met Gly Gly Val Asn Thr Val
225                 230                 235                 240

Tyr Phe Ala Phe Asn Glu Gly Phe Val Arg Ser Asn Ile Arg Ser Asp
                245                 250                 255

Glu Ile Arg Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
            260                 265                 270

Leu Asn Gln Ser Lys Tyr Cys Ser Asp Ser Arg Thr Gly Lys Gly Arg
        275                 280                 285

Ala Lys Arg Leu Gln Pro Ile Asp Val Ile Ser Asn Lys Ile Ala Gln
    290                 295                 300

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Phe Ile Val Lys Gln
305                 310                 315                 320

Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Ile Glu His Leu Lys Gly
                325                 330                 335

Ile Ser Lys Asp Asp Lys Val Leu Lys Asp Trp Thr Tyr Tyr Asp Leu
            340                 345                 350

Gln Glu Lys Ile Lys Lys Gln Ala Gln Ala Tyr Gly Ile Glu Val Ile
        355                 360                 365
```

```
Thr Ile Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
    370                 375                 380

Ile Ser Asn Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
385                 390                 395                 400

Cys Glu Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
                405                 410                 415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
            420                 425                 430

Arg Leu His Ser Lys Lys Tyr Met Glu Asp His Ile Glu Glu Leu Gly
        435                 440                 445

Tyr Ser Gly
    450
```

<210> 106
<211> 450
<212> PRT
<213> Bacillus sp.

<400> 106

```
Met Ser Leu Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1               5                   10                  15

Asn Val Glu Trp Lys Val Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr
            20                  25                  30

Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe
        35                  40                  45

Asp Asn Leu Ser Leu Lys His Phe Lys Glu Thr Gly Glu Tyr Pro Ser
    50                  55                  60

Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65                  70                  75                  80

Tyr Asp Gln Leu Lys Glu Glu Tyr Lys Asp Ile Asn Lys Ala Asn Met
                85                  90                  95
```

```
Ser Thr Thr Ile Gln Lys Thr Leu Lys Asn Trp Asn Ser Arg Lys Lys
            100                 105                 110

Glu Ile Trp Arg Gly Glu Met Ser Ile Pro Ser Phe Arg Ser Asp Leu
        115                 120                 125

Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Ile Lys Glu Lys Gly
    130                 135                 140

Gly Asp Tyr Ile Ala Ser Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
145                 150                 155                 160

Glu Asn Asp Leu Pro Asn Gly Lys Ile Leu Met Lys Leu Ser Thr Arg
                165                 170                 175

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr
            180                 185                 190

Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Lys Lys Trp Tyr
        195                 200                 205

Leu Ser Ile Thr Tyr Lys Ser Asn Ile Lys Glu Glu Leu Lys Phe Asp
    210                 215                 220

Glu Asp Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
225                 230                 235                 240

Phe Ala Phe Asn Lys Ser Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
                245                 250                 255

Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Gly Asn Arg Ile Gly Lys Gly Arg Glu
        275                 280                 285

Lys Arg Ile Lys Pro Ile Asp Val Leu Asn Asp Lys Val Ala Lys Phe
    290                 295                 300

Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320
```

```
Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Lys Gly Ile
                325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350

Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Val Lys Val Val Lys
        355                 360                 365

Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
    370                 375                 380

His Lys Asp Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Leu Gln Cys
385                 390                 395                 400

Ser Phe Lys Ala His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Val
                405                 410                 415

Tyr Thr Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
            420                 425                 430

Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Ile Lys Asn
        435                 440                 445

Ile Asn
    450
```

<210> 107
<211> 450
<212> PRT
<213> Bacillus sp.

<400> 107

```
Met Ser Leu Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1               5                   10                  15

Asn Val Glu Trp Lys Val Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr
            20                  25                  30

Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe
        35                  40                  45
```

```
Asp Asn Leu Ser Leu Lys His Phe Lys Glu Thr Gly Glu Tyr Pro Ser
    50                  55                  60

Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65                  70                  75                  80

Tyr Asp Gln Leu Lys Glu Glu Tyr Lys Asp Ile Asn Lys Ala Asn Met
                85                  90                  95

Ser Thr Thr Ile Gln Lys Thr Leu Lys Asn Trp Asn Ser Arg Lys Lys
            100                 105                 110

Glu Ile Trp Arg Gly Glu Met Ser Ile Pro Ser Phe Arg Ser Asp Leu
        115                 120                 125

Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Ile Lys Glu Lys Gly
    130                 135                 140

Gly Asp Tyr Ile Ala Ser Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
145                 150                 155                 160

Glu Asn Asp Leu Pro Asn Gly Lys Ile Leu Met Lys Leu Ser Thr Arg
                165                 170                 175

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr
            180                 185                 190

Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Lys Lys Trp Tyr
        195                 200                 205

Leu Ser Ile Thr Tyr Lys Ser Asn Ile Lys Glu Glu Leu Lys Phe Asp
    210                 215                 220

Glu Asp Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
225                 230                 235                 240

Phe Ala Phe Asn Lys Ser Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
                245                 250                 255
```

```
Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Gly Asn Arg Ile Gly Lys Gly Arg Glu
        275                 280                 285

Lys Arg Ile Lys Pro Ile Asp Val Leu Asn Asp Lys Val Ala Lys Phe
    290                 295                 300

Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320

Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Lys Gly Ile
                325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350

Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Val Lys Val Val Lys
        355                 360                 365

Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
    370                 375                 380

His Lys Asp Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Leu Gln Cys
385                 390                 395                 400

Ser Phe Lys Ala His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Val
                405                 410                 415

Tyr Thr Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
            420                 425                 430

Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Ile Lys Asn
        435                 440                 445

Ile Asn
    450
```

<210> 108
<211> 450

<212> PRT
<213> Bacillus sp.

<400> 108
Met Ser Ile Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1   5   10   15

Asn Val Glu Trp Lys Ala Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr
            20   25   30

Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe
        35   40   45

Asp Asn Leu Ser Leu Asn His Phe Lys Glu Thr Gly Glu Tyr Pro Ser
    50   55   60

Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65   70   75   80

Tyr Asn Gln Leu Arg Glu Glu Tyr Gln Asp Ile Asn Lys Ala Asn Met
                85   90   95

Ser Thr Thr Ile Gln Lys Ser Leu Lys Asn Trp Asn Ser Arg Lys Lys
            100   105   110

Glu Ile Trp Arg Gly Glu Met Ser Ile Pro Ser Phe Arg Ser Asp Leu
        115   120   125

Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Ile Lys Glu Lys Ser
    130   135   140

Gly Asp Tyr Phe Ala Asn Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
145   150   155   160

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

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr
            180   185   190

Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Asn Lys Trp Tyr
        195   200   205

```
Leu Ser Ile Thr Tyr Lys Ala Thr Lys Lys Glu Gly His Lys Phe Asp
    210                 215                 220

Glu Glu Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
225                 230                 235                 240

Phe Ala Tyr Asn Lys Gly Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
                245                 250                 255

Ile Glu Ala Phe Arg Lys Lys Ile Glu Tyr Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Glu Asn Arg Ile Gly Lys Gly Arg Lys
        275                 280                 285

Lys Arg Ile Lys Pro Ile Glu Val Leu Asn Asp Lys Ile Thr Lys Phe
    290                 295                 300

Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320

Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Gln Gly Ile
                325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350

Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Ile Lys Val Val Lys
        355                 360                 365

Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
    370                 375                 380

His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln His Cys
385                 390                 395                 400

Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Ile
                405                 410                 415
```

```
Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
            420                 425                 430

Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Glu Asn
        435                 440                 445

Ile Asn
    450
```

<210> 109
<211> 450
<212> PRT
<213> Bacillus sp.

<400> 109

```
Met Ser Ile Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
1               5                   10                  15

Asn Val Glu Trp Lys Ala Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr
            20                  25                  30

Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe
        35                  40                  45

Asp Asn Leu Ser Leu Asn His Phe Lys Glu Thr Gly Glu Tyr Pro Ser
    50                  55                  60

Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
65                  70                  75                  80

Tyr Asn Gln Leu Arg Glu Glu Tyr Gln Asp Ile Asn Lys Ala Asn Met
                85                  90                  95

Ser Thr Thr Ile Gln Lys Ser Leu Lys Asn Trp Asn Ser Arg Lys Lys
            100                 105                 110

Glu Ile Trp Arg Gly Glu Met Ser Ile Pro Ser Phe Arg Ser Asp Leu
        115                 120                 125

Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Ile Lys Glu Lys Ser
    130                 135                 140
```

```
Gly Asp Tyr Phe Ala Asn Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
145                 150                 155                 160

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

Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr
            180                 185                 190

Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Asn Lys Trp Tyr
        195                 200                 205

Leu Ser Ile Thr Tyr Lys Ala Thr Lys Lys Glu Gly His Lys Phe Asp
    210                 215                 220

Glu Glu Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
225                 230                 235                 240

Phe Ala Tyr Asn Lys Gly Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
                245                 250                 255

Ile Glu Ala Phe Arg Lys Lys Ile Glu Tyr Arg Arg Ile Ser Leu Leu
            260                 265                 270

Arg Gln Gly Lys Tyr Cys Ser Glu Asn Arg Ile Gly Lys Gly Arg Lys
        275                 280                 285

Lys Arg Ile Lys Pro Ile Glu Val Leu Asn Asp Lys Ile Thr Lys Phe
    290                 295                 300

Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys
305                 310                 315                 320

Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Gln Gly Ile
                325                 330                 335

Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln
            340                 345                 350

Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Ile Lys Val Val Lys
        355                 360                 365
```

Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile
370 375 380

His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln His Cys
385 390 395 400

Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Ile
405 410 415

Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
420 425 430

Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Glu Asn
435 440 445

Ile Asn
450

<210> 110
<211> 11
<212> RNA
<213> Bacillus sp.

<400> 110
aggauugaaa u 11

<210> 111
<211> 11
<212> RNA
<213> Bacillus sp.

<400> 111
aggauugaaa u 11

<210> 112
<211> 11
<212> RNA
<213> Bacillus sp.

<400> 112
uggauugaaa u 11

<210> 113

<211> 11
<212> RNA
<213> Bacillus sp.

<400> 113
aggauugaaa u 11

<210> 114
<211> 11
<212> RNA
<213> Bacillus sp.

<400> 114
uggauugaaa u 11

<210> 115
<211> 18
<212> RNA
<213> Bacillus sp.

<400> 115
caauagaugu auugaaau 18

<210> 116
<211> 37
<212> RNA
<213> Gordonia sp.

<400> 116
gucgugagua cgacgcccuc acgcuggcgu ugcgacc 37

<210> 117
<211> 13
<212> RNA
<213> Micrococcus sp.

<400> 117
cagggaugag ccc 13

<210> 118
<211> 13
<212> RNA
<213> Micrococcus sp.

<400> 118
cagggaugag ccc 13

<210> 119

<211> 13
<212> RNA
<213> Micrococcus sp.

<400> 119
cagggaugag ccc 13

<210> 120
<211> 201
<212> RNA
<213> Bacillus sp.

<400> 120
aaacagcuag aauguaacuu aaaguagguc aauguuuaaa uucgauguug caauuuguuu 60

ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg 120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc 180

aaagggucaa cucgauacau u 201

<210> 121
<211> 201
<212> RNA
<213> Bacillus sp.

<400> 121
aaacagcuag aauguaacuu aaaguagguc aauguuuaaa uucgauguug caauuuguuu 60

ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg 120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc 180

aaagggucaa cucgauacau u 201

<210> 122
<211> 201
<212> RNA
<213> Bacillus sp.

<400> 122
aaacagcuag aauguaacuu aaaguagguc aauguuuaaa uucgauauug caauuuguuu 60

ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg 120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc 180

aaagggucaa cucgauacau u 201

<210> 123

<211> 201
<212> RNA
<213> Bacillus sp.

<400> 123
aaacagcuag aauguaacuu aaaguagguc aauguuuaaa uucgauguug caauuuguuu 60

ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg 120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc 180

aaagggucaa cucgauacau u 201

<210> 124
<211> 201
<212> RNA
<213> Bacillus sp.

<400> 124
aaacagcuag aauguaacuu aaaguagguc aauguuuaaa uucgauauug caauuuguuu 60

ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg 120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc 180

aaagggucaa cucgauacau u 201

<210> 125
<211> 124
<212> RNA
<213> Bacillus sp.

<400> 125
acggucuuug uguuacaccc uucaaagcau ccaggaugcu uauguaucaa uguuuuuuua 60

acauucgaua cuaccuuaug gagcguuuac acuugguaaa uucaccuugu auuacuucuc 120

cauu 124

<210> 126
<211> 97
<212> RNA
<213> Micrococcus sp.

<400> 126
uucugggcgc caggcagagc gaagcccucc cggcgcggag gggggauggc caggagggcu 60

ucuaagucgc aggcccugcg ggaugggcuc acguucg 97

<210> 127

<211> 96
<212> RNA
<213> Micrococcus sp.

<400> 127
guuuacacag gucgaacgug agcccauccc guagggccug cgacuuagaa gcccuccugg 60

ccaucccccc uccgcgccgg gagggcuucg cucugc 96

<210> 128
<211> 97
<212> RNA
<213> Micrococcus sp.

<400> 128
uucugggcgc caggcagagc gaagcccucc cggcgcggag gggggauggc caggagggcu 60

ucuaagucgc aggcccugcg ggaugggcuc acguucg 97

<210> 129
<211> 221
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 129
aaacagcuag aauguaacuu aaaguagguc aauguuuaaa uucgauguug caauuuguuu 60

ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg 120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc 180

aaagggucaa aaagauccua cauggaugug aggauugaaa u 221

<210> 130
<211> 221
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 130
aaacagcuag aauguaacuu aaaguagguc aauguuuaaa uucgauguug caauuuguuu 60

ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg 120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc 180

aaagggucaa aaagauccua cauggaugug aggauugaaa u 221

<210> 131
<211> 221
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 131
aaacagcuag aauguaacuu aaaguagguc aauguuuaaa uucgauguug caauuuguuu 60

ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg 120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc 180

aaagggucaa aaagauccua cauggaugug aggauugaaa u 221

<210> 132
<211> 216
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 132
aaacagcuag aauguaacuu aaaguagguc aauguuuaaa uucgauauug caauuuguuu 60

ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg 120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc 180

aaagggucaa cucgauacau uaaaguggau ugaaau 216

<210> 133
<211> 114
<212> RNA
<213> Artificial Sequence

<220>
<221> source

<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 133
uucugggcgc caggcagagc gaagcccucc cggcgcggag gggggauggc caggagggcu 60

ucuaagucgc aggcccugcg ggaugggcuc acguucgaaa gcagggauga gccc 114

<210> 134
<211> 113
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 134
guuuacacag gucgaacgug agcccauccc guagggccug cgacuuagaa gcccuccugg 60

ccaucccccc uccgcgccgg gagggcuucg cucugcaaag cagggacgag ccc 113

<210> 135
<211> 114
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 135
uucugggcgc caggcagagc gaagcccucc cggcgcggag gggggauggc caggagggcu 60

ucuaagucgc aggcccugcg ggaugggcuc acguucgaaa gcagggauga gccc 114

<210> 136
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
peptide"

<400> 136
Ala Ala Ala Gly
1

<210> 137
<211> 82
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polypeptide"

<400> 137
Thr Asn Leu Ser Asp His Glu Lys Glu Thr Gly Lys Gln Leu Val Ile
1 5 10 15

Gln Glu Ser Ile Leu Met Leu Pro Glu Glu Val Glu Glu Val Ile Gly
20 25 30

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

Thr Asp Glu Asn Val Met Leu Leu Thr Ser Asp Ala Pro Glu Tyr Lys
50 55 60

Pro Trp Ala Leu Val Ile Gln Asp Ser Asn Gly Glu Asn Lys Ile Lys
65 70 75 80

Met Leu

<210> 138
<211> 1493
<212> PRT
<213> Homo sapiens

<400> 138
Met Pro Asn Glu Gly Ile Pro His Ser Ser Gln Thr Gln Glu Gln Asp
1 5 10 15

Cys Leu Gln Ser Gln Pro Val Ser Asn Asn Glu Glu Met Ala Ile Lys
20 25 30

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

```
Ser Val Gly Asp Gly Leu Ser Thr Ser Ala Val Gly Cys Ala Ser Ala
    50                  55                  60

Ala Pro Arg Arg Gly Pro Ala Leu Leu His Ile Asp Arg His Gln Ile
65                  70                  75                  80

Gln Ala Val Glu Pro Ser Ala Gln Ala Leu Glu Leu Gln Gly Leu Gly
                85                  90                  95

Val Asp Val Tyr Asp Gln Asp Val Leu Glu Gln Gly Val Leu Gln Gln
            100                 105                 110

Val Asp Asn Ala Ile His Glu Ala Ser Arg Ala Ser Gln Leu Val Asp
        115                 120                 125

Val Glu Lys Glu Tyr Arg Ser Val Leu Asp Asp Leu Thr Ser Cys Thr
    130                 135                 140

Thr Ser Leu Arg Gln Ile Asn Lys Ile Ile Glu Gln Leu Ser Pro Gln
145                 150                 155                 160

Ala Ala Thr Ser Arg Asp Ile Asn Arg Lys Leu Asp Ser Val Lys Arg
                165                 170                 175

Gln Lys Tyr Asn Lys Glu Gln Gln Leu Lys Lys Ile Thr Ala Lys Gln
            180                 185                 190

Lys His Leu Gln Ala Ile Leu Gly Gly Ala Glu Val Lys Ile Glu Leu
        195                 200                 205

Asp His Ala Ser Leu Glu Glu Asp Ala Glu Pro Gly Pro Ser Ser Leu
    210                 215                 220

Gly Ser Met Leu Met Pro Val Gln Glu Thr Ala Trp Glu Glu Leu Ile
225                 230                 235                 240

Arg Thr Gly Gln Met Thr Pro Phe Gly Thr Gln Ile Pro Gln Lys Gln
                245                 250                 255

Glu Lys Lys Pro Arg Lys Ile Met Leu Asn Glu Ala Ser Gly Phe Glu
```

```
            260                 265                 270

Lys Tyr Leu Ala Asp Gln Ala Lys Leu Ser Phe Glu Arg Lys Lys Gln
        275                 280                 285

Gly Cys Asn Lys Arg Ala Ala Arg Lys Ala Pro Ala Pro Val Thr Pro
    290                 295                 300

Pro Ala Pro Val Gln Asn Lys Asn Lys Pro Asn Lys Lys Ala Arg Val
305                 310                 315                 320

Leu Ser Lys Lys Glu Glu Arg Leu Lys Lys His Ile Lys Lys Leu Gln
                325                 330                 335

Lys Arg Ala Leu Gln Phe Gln Gly Lys Val Gly Leu Pro Lys Ala Arg
            340                 345                 350

Arg Pro Trp Glu Ser Asp Met Arg Pro Glu Ala Glu Gly Asp Ser Glu
        355                 360                 365

Gly Glu Glu Ser Glu Tyr Phe Pro Thr Glu Glu Glu Glu Glu Glu Glu
    370                 375                 380

Asp Asp Glu Val Glu Gly Ala Glu Ala Asp Leu Ser Gly Asp Gly Thr
385                 390                 395                 400

Asp Tyr Glu Leu Lys Pro Leu Pro Lys Gly Gly Lys Arg Gln Lys Lys
                405                 410                 415

Val Pro Val Gln Glu Ile Asp Asp Asp Phe Phe Pro Ser Ser Gly Glu
            420                 425                 430

Glu Ala Glu Ala Ala Ser Val Gly Glu Gly Gly Gly Gly Gly Arg Lys
        435                 440                 445

Val Gly Arg Tyr Arg Asp Asp Gly Asp Glu Asp Tyr Tyr Lys Gln Arg
    450                 455                 460

Leu Arg Arg Trp Asn Lys Leu Arg Leu Gln Asp Lys Glu Lys Arg Leu
465                 470                 475                 480
```

```
Lys Leu Glu Asp Asp Ser Glu Glu Ser Asp Ala Glu Phe Asp Glu Gly
                485                 490                 495

Phe Lys Val Pro Gly Phe Leu Phe Lys Lys Leu Phe Lys Tyr Gln Gln
            500                 505                 510

Thr Gly Val Arg Trp Leu Trp Glu Leu His Cys Gln Gln Ala Gly Gly
        515                 520                 525

Ile Leu Gly Asp Glu Met Gly Leu Gly Lys Thr Ile Gln Ile Ile Ala
    530                 535                 540

Phe Leu Ala Gly Leu Ser Tyr Ser Lys Ile Arg Thr Arg Gly Ser Asn
545                 550                 555                 560

Tyr Arg Phe Glu Gly Leu Gly Pro Thr Val Ile Val Cys Pro Thr Thr
                565                 570                 575

Val Met His Gln Trp Val Lys Glu Phe His Thr Trp Trp Pro Pro Phe
            580                 585                 590

Arg Val Ala Ile Leu His Glu Thr Gly Ser Tyr Thr His Lys Lys Glu
        595                 600                 605

Lys Leu Ile Arg Asp Val Ala His Cys His Gly Ile Leu Ile Thr Ser
    610                 615                 620

Tyr Ser Tyr Ile Arg Leu Met Gln Asp Asp Ile Ser Arg Tyr Asp Trp
625                 630                 635                 640

His Tyr Val Ile Leu Asp Glu Gly His Lys Ile Arg Asn Pro Asn Ala
                645                 650                 655

Ala Val Thr Leu Ala Cys Lys Gln Phe Arg Thr Pro His Arg Ile Ile
            660                 665                 670

Leu Ser Gly Ser Pro Met Gln Asn Asn Leu Arg Glu Leu Trp Ser Leu
        675                 680                 685

Phe Asp Phe Ile Phe Pro Gly Lys Leu Gly Thr Leu Pro Val Phe Met
```

```
    690                 695                 700

Glu Gln Phe Ser Val Pro Ile Thr Met Gly Gly Tyr Ser Asn Ala Ser
705                 710                 715                 720

Pro Val Gln Val Lys Thr Ala Tyr Lys Cys Ala Cys Val Leu Arg Asp
                725                 730                 735

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

Ser Leu Ser Leu Pro Asp Lys Asn Glu Gln Val Leu Phe Cys Arg Leu
        755                 760                 765

Thr Asp Glu Gln His Lys Val Tyr Gln Asn Phe Val Asp Ser Lys Glu
    770                 775                 780

Val Tyr Arg Ile Leu Asn Gly Glu Met Gln Ile Phe Ser Gly Leu Ile
785                 790                 795                 800

Ala Leu Arg Lys Ile Cys Asn His Pro Asp Leu Phe Ser Gly Gly Pro
                805                 810                 815

Lys Asn Leu Lys Gly Leu Pro Asp Asp Glu Leu Glu Glu Asp Gln Phe
            820                 825                 830

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

Lys Ile Trp His Lys Gln Gly Gln Arg Val Leu Leu Phe Ser Gln Ser
    850                 855                 860

Arg Gln Met Leu Asp Ile Leu Glu Val Phe Leu Arg Ala Gln Lys Tyr
865                 870                 875                 880

Thr Tyr Leu Lys Met Asp Gly Thr Thr Thr Ile Ala Ser Arg Gln Pro
                885                 890                 895

Leu Ile Thr Arg Tyr Asn Glu Asp Thr Ser Ile Phe Val Phe Leu Leu
            900                 905                 910
```

```
Thr Thr Arg Val Gly Gly Leu Gly Val Asn Leu Thr Gly Ala Asn Arg
        915                 920                 925

Val Val Ile Tyr Asp Pro Asp Trp Asn Pro Ser Thr Asp Thr Gln Ala
    930                 935                 940

Arg Glu Arg Ala Trp Arg Ile Gly Gln Lys Lys Gln Val Thr Val Tyr
945                 950                 955                 960

Arg Leu Leu Thr Ala Gly Thr Ile Glu Glu Lys Ile Tyr His Arg Gln
                965                 970                 975

Ile Phe Lys Gln Phe Leu Thr Asn Arg Val Leu Lys Asp Pro Lys Gln
            980                 985                 990

Arg Arg Phe Phe Lys Ser Asn Asp  Leu Tyr Glu Leu Phe  Thr Leu Thr
        995                 1000                1005

Ser Pro  Asp Ala Ser Gln Ser  Thr Glu Thr Ser Ala  Ile Phe Ala
    1010                1015                1020

Gly Thr  Gly Ser Asp Val Gln  Thr Pro Lys Cys His  Leu Lys Arg
    1025                1030                1035

Arg Ile  Gln Pro Ala Phe Gly  Ala Asp His Asp Val  Pro Lys Arg
    1040                1045                1050

Lys Lys  Phe Pro Ala Ser Asn  Ile Ser Val Asn Asp  Ala Thr Ser
    1055                1060                1065

Ser Glu  Glu Lys Ser Glu Ala  Lys Gly Ala Glu Val  Asn Ala Val
    1070                1075                1080

Thr Ser  Asn Arg Ser Asp Pro  Leu Lys Asp Asp Pro  His Met Ser
    1085                1090                1095

Ser Asn  Val Thr Ser Asn Asp  Arg Leu Gly Glu Glu  Thr Asn Ala
    1100                1105                1110

Val Ser  Gly Pro Glu Glu Leu  Ser Val Ile Ser Gly  Asn Gly Glu
```

```
    1115                1120                1125

Cys Ser  Asn Ser Ser Gly Thr  Gly Lys Thr Ser Met  Pro Ser Gly
    1130                1135                1140

Asp Glu  Ser Ile Asp Glu Lys  Leu Gly Leu Ser Tyr  Lys Arg Glu
    1145                1150                1155

Arg Pro  Ser Gln Ala Gln Thr  Glu Ala Phe Trp Glu  Asn Lys Gln
    1160                1165                1170

Met Glu  Asn Asn Phe Tyr Lys  His Lys Ser Lys Thr  Lys His His
    1175                1180                1185

Ser Val  Ala Glu Glu Glu Thr  Leu Glu Lys His Leu  Arg Pro Lys
    1190                1195                1200

Gln Lys  Pro Lys Asn Ser Lys  His Cys Arg Asp Ala  Lys Phe Glu
    1205                1210                1215

Gly Thr  Arg Ile Pro His Leu  Val Lys Lys Arg Arg  Tyr Gln Lys
    1220                1225                1230

Gln Asp  Ser Glu Asn Lys Ser  Glu Ala Lys Glu Gln  Ser Asn Asp
    1235                1240                1245

Asp Tyr  Val Leu Glu Lys Leu  Phe Lys Lys Ser Val  Gly Val His
    1250                1255                1260

Ser Val  Met Lys His Asp Ala  Ile Met Asp Gly Ala  Ser Pro Asp
    1265                1270                1275

Tyr Val  Leu Val Glu Ala Glu  Ala Asn Arg Val Ala  Gln Asp Ala
    1280                1285                1290

Leu Lys  Ala Leu Arg Leu Ser  Arg Gln Arg Cys Leu  Gly Ala Val
    1295                1300                1305

Ser Gly  Val Pro Thr Trp Thr  Gly His Arg Gly Ile  Ser Gly Ala
    1310                1315                1320
```

```
Pro Ala  Gly Lys Lys Ser Arg  Phe Gly Lys Lys Arg  Asn Ser Asn
    1325                1330                1335

Phe Ser  Val Gln His Pro Ser  Ser Thr Ser Pro Thr  Glu Lys Cys
    1340                1345                1350

Gln Asp  Gly Ile Met Lys Lys  Glu Gly Lys Asp Asn  Val Pro Glu
    1355                1360                1365

His Phe  Ser Gly Arg Ala Glu  Asp Ala Asp Ser Ser  Ser Gly Pro
    1370                1375                1380

Leu Ala  Ser Ser Ser Leu Leu  Ala Lys Met Arg Ala  Arg Asn His
    1385                1390                1395

Leu Ile  Leu Pro Glu Arg Leu  Glu Ser Glu Ser Gly  His Leu Gln
    1400                1405                1410

Glu Ala  Ser Ala Leu Leu Pro  Thr Thr Glu His Asp  Asp Leu Leu
    1415                1420                1425

Val Glu  Met Arg Asn Phe Ile  Ala Phe Gln Ala His  Thr Asp Gly
    1430                1435                1440

Gln Ala  Ser Thr Arg Glu Ile  Leu Gln Glu Phe Glu  Ser Lys Leu
    1445                1450                1455

Ser Ala  Ser Gln Ser Cys Val  Phe Arg Glu Leu Leu  Arg Asn Leu
    1460                1465                1470

Cys Thr  Phe His Arg Thr Ser  Gly Gly Glu Gly Ile  Trp Lys Leu
    1475                1480                1485

Lys Pro  Glu Tyr Cys
    1490
```

<210> 139
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 139
gtcacatcct acatggatgt gtggattgaa attggaatgg gaactaaagt aatgg 55

<210> 140
<211> 201
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 140
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatattg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa ctcgatacat t 201

<210> 141
<211> 53
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 141
gtttaaacat aacaatagat gtattgaaat tggaatggga actaaagtaa tgg 53

<210> 142
<211> 124
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 142

acggtctttg tgttacaccc ttcaaagcat ccaggatgct tatgtatcaa tgtttttta 60

acattcgata ctaccttatg gagcgtttac acttggtaaa ttcaccttgt attacttctc 120

catt 124

<210> 143
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 143
gtcacatcct acatggatgt gaggattgaa attggaatgg gaactaaagt aatgg 55

<210> 144
<211> 70
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 144
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg 70

<210> 145
<211> 201
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 145
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa ctcgatacat t 201

<210> 146
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 146
gtcacatcct acatggatgt gtggattgaa attggaatgg gaactaaagt aatgg 55

<210> 147
<211> 201
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 147
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatattg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa ctcgatacat t 201

<210> 148
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 148
cagccagttt acgggcttgg gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag 60

aacctacaaa agcactcacc aaagggtcaa ctcgatacat t 101

<210> 149

<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 149
cctaagaaga aaagaaaggt g 21

<210> 150
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 150
aaaagacctg ccgctacaaa gaaggccggc caggccaaga aaaagaag 48

<210> 151
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 151
gactacaagg accacgacgg cgactacaaa gatcacgata tcgactacaa ggacgacgat 60

gataag 66

<210> 152
<211> 203
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 152

tgatgcggtt ttggcagtac atcaatgggc gtggatagcg gtttgactca cggggatttc 60

caagtctcca ccccattgac gtcaatggga gtttgttttg gcaccaaaat caacgggact 120

ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat gggcggtagg cgtgtacggt 180

gggaggtcta tataagcaga gct 203

<210> 153
<211> 318
<212> DNA
<213> Homo sapiens

<400> 153
tgtacaaaaa agcaggcttt aaaggaacca attcagtcga ctggatccgg taccaaggtc 60

gggcaggaag agggcctatt tcccatgatt ccttcatatt tgcatatacg atacaaggct 120

gttagagaga taattagaat taatttgact gtaaacacaa agatattagt acaaaatacg 180

tgacgtagaa agtaataatt tcttgggtag tttgcagttt taaaattatg ttttaaaatg 240

gactatcata tgcttaccgt aacttgaaag tatttcgatt tcttggcttt atatatcttg 300

tggaaaggac gaaacacc 318

<210> 154
<211> 1644
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 154
gacgcattgg acgattttga tctggatatg ctgggaagtg acgccctcga tgattttgac 60

cttgacatgc ttggttcgga tgcccttgat gactttgacc tcgacatgct cggcagtgac 120

gcccttgatg atttcgacct ggacatgctg attaactcta gaagttccgg atctggtagc 180

cagtacctgc ccgacaccga cgaccggcac cggatcgagg aaaagcggaa gcggacctac 240

gagacattca agagcatcat gaagaagtcc cccttcagcg gccccaccga ccctagacct 300

ccacctagaa gaatcgccgt gcccagcaga tccagcgcca gcgtgccaaa acctgccccc 360

cagccttacc ccttcaccag cagcctgagc accatcaact acgacgagtt ccctaccatg 420

gtgttcccca gcggccagat ctctcaggcc tctgctctgg ctccagcccc tcctcaggtg 480

ctgcctcagg ctcctgctcc tgcaccagct ccagccatgg tgtctgcact ggctcaggca 540

ccagcacccg tgcctgtgct ggctcctgga cctccacagg ctgtggctcc accagcccct 600

aaacctacac aggccggcga gggcacactg tctgaagctc tgctgcagct gcagttcgac 660

gacgaggatc tgggagccct gctgggaaac agcaccgatc ctgccgtgtt caccgacctg 720

gccagcgtgg acaacagcga gttccagcag ctgctgaacc agggcatccc tgtggccct 780

cacaccaccg agcccatgct gatggaatac cccgaggcca tcacccggct cgtgacaggc 840

gctcagaggc ctcctgatcc agctcctgcc cctctgggag caccaggcct gcctaatgga 900

ctgctgtctg gcgacgagga cttcagctct atcgccgata tggatttctc agccttgctg 960

ggctctggca gcggcagccg ggattccagg gaagggatgt ttttgccgaa gcctgaggcc 1020

ggctccgcta ttagtgacgt gtttgagggc cgcgaggtgt gccagccaaa acgaatccgg 1080

ccatttcatc ctccaggaag tccatgggcc aaccgcccac tccccgccag cctcgcacca 1140

acaccaaccg gtccagtaca tgagccagtc gggtcactga ccccggcacc agtccctcag 1200

ccactagatc cagcgcccgc agtgactccc gaggccagtc acctgttgga agatcccgat 1260

gaagagacga gccaggctgt caaagccctt cgggagatgg ccgatactgt gattccccag 1320

aaggaagagg ctgcaatctg tggccaaatg gacctttccc atccgccccc aaggggccat 1380

ctggatgagc tgacaaccac acttgagtcc atgaccgagg atctgaacct ggactcaccc 1440

ctgaccccgg aattgaacga gattctggat accttcctga acgacgagtg cctcttgcat 1500

gccatgcata tcagcacagg actgtccatc ttcgacacat ctctgttttc cggaggatct 1560

agcggaggct cctctggctc tgagacacct ggcacaagcg agagcgcaac acctgaaagc 1620

agcgggggca gcagcggggg gtca 1644

<210> 155
<211> 1518
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 155
atgagcaccc ctctgcagca gcctcaccag aaatccaaga agaccagcca gatgataaca 60

acacgcaagt tcaagctcgc catcgttagc gacaaccgga acgaggccta cagcttcatc 120

cggaacgaga tcagaaacca gaacaaggcc ctgaacgccg cttataacca cctgtacttc 180

gagcacatcg caaccgagaa actgaagcac agcgatgagg aataccagaa gcacctgaca 240

aagtacagag aggtggccac caacaagtac caggactacc tgaaggtgaa agaaaaagtg 300

aacgacagca aggacgacga gaagctgcag aagcgggtgg ataaggctag agaggcctac 360

aacaaggccc aggagaaagt gtacaagatc gagaaggaat tcaacaaaaa gagcatggaa 420

acctaccaga aggtggtcgg cctgagcaaa cagacaagaa tcggaaagct gctcaagagc 480

cagttcaccc tgcactacga caccgaagat agaatcacca gcacagtgat cagccacttt 540

aacaatgata tgaagacagg agtgctgaga ggcgatagaa gcctgagaac atacaagaac 600

agccatcctc tgctggttag agccagatct atgaaggtgt acgaggaaaa cggcgactac 660

ttcatcaagt gggtcaaggg gatcgtgttt aagatcgtga tcagcgccgg cagcaagcag 720

aaagctaata tcggcgagct gaaatctgtg cttatcaaca tcctgaatgg ccactataag 780

gtgtgcgaca gcagcatcag cctgaataaa gacctgatcc tgaacctgtc tctgaacatc 840

cctgtgtcta aggagaacgt gttcgtgccc ggcagagtgg tgggcgtgga cctgggcctg 900

aagatcccag cctatgtgtc cctgaacgac acaccctaca tcaagaaggg catcggcaac 960

atcgacgatt tcctgaaggt gcggacccaa ctgcagagcc agagaaagag actgcaaaag 1020

accctggaat gtacctccgg cggcaagggc agaaacaaga agctgaaggg actggacaga 1080

ctgaaggcca aggaaaaaaa cttcgtgaac acctacaacc acttcctgag caagaaaatc 1140

atccagtttg ccgtgaagaa caacgccggc gtgatccacc tggaagagct gcagttcgac 1200

aagctgaaac ataagtccct gctgcggaac tggtcctact accaactgca gacaatgatc 1260

gagtacaagg ctgaacggga aggcatcgag gtgaagtacg tggacgccag ctacaccagc 1320

cagacctgta gcaagtgcgg ccactacgag gaaggccagc gggtgctgca agacaccttc 1380

acctgcaaga acaaagagtg caagggctac gtgcacaagg ttaacgccga cttcaacgcc 1440

tctcagaata tcgccaagtc taccgacatc atccggtgca ccgagatggc caagaacaac 1500

gatatcgaga agaatgcc 1518

<210> 156
<211> 1518

<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 156
atgagcaccc ctctgcagca gcctcaccag aaatccaaga agaccagcca gatgataaca 60

acacgcaagt tcaagctcgc catcgttagc gacaaccgga acgaggccta cagcttcatc 120

cggaacgaga tcagaaacca gaacaaggcc ctgaacgccg cttataacca cctgtacttc 180

gagcacatcg caaccgagaa actgaagcac agcgatgagg aataccagaa gcacctgaca 240

aagtacagag aggtggccac caacaagtac caggactacc tgaaggtgaa agaaaaagtg 300

aacgacagca aggacgacga gaagctgcag aagcgggtgg ataaggctag agaggcctac 360

aacaaggccc aggagaaagt gtacaagatc gagaaggaat tcaacaaaaa gagcatggaa 420

acctaccaga aggtggtcgg cctgagcaaa cagacaagaa tcggaaagct gctcaagagc 480

cagttcaccc tgcactacga caccgaagat agaatcacca gcacagtgat cagccacttt 540

aacaatgata tgaagacagg agtgctgaga ggcgatagaa gcctgagaac atacaagaac 600

agccatcctc tgctggttag agccagatct atgaaggtgt acgaggaaaa cggcgactac 660

ttcatcaagt gggtcaaggg gatcgtgttt aagatcgtga tcagcgccgg cagcaagcag 720

aaagctaata tcggcgagct gaaatctgtg cttatcaaca tcctgaatgg ccactataag 780

gtgtgcgaca gcagcatcag cctgaataaa gacctgatcc tgaacctgtc tctgaacatc 840

cctgtgtcta aggagaacgt gttcgtgccc ggcagagtgg tgggcgtggc cctgggcctg 900

aagatcccag cctatgtgtc cctgaacgac acaccctaca tcaagaaggg catcggcaac 960

atcgacgatt tcctgaaggt gcggacccaa ctgcagagcc agagaaagag actgcaaaag 1020

accctggaat gtacctccgg cggcaagggc agaaacaaga agctgaaggg actggacaga 1080

ctgaaggcca aggaaaaaaa cttcgtgaac acctacaacc acttcctgag caagaaaatc 1140

atccagtttg ccgtgaagaa caacgccggc gtgatccacc tggccgagct gcagttcgac 1200

aagctgaaac ataagtccct gctgcggaac tggtcctact accaactgca gacaatgatc 1260

gagtacaagg ctgaacggga aggcatcgag gtgaagtacg tggacgccag ctacaccagc 1320

cagacctgta gcaagtgcgg ccactacgag gaaggccagc gggtgctgca agacaccttc 1380

acctgcaaga acaaagagtg caagggctac gtgcacaagg ttaacgccgc cttcaacgcc 1440

tctcagaata tcgccaagtc taccgacatc atccggtgca ccgagatggc caagaacaac 1500

gatatcgaga agaatgcc 1518

<210> 157
<211> 7494
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 157
gcgatcgcgg ctcccgacat cttggaccat tagctccaca ggtatcttct tccctctagt 60

ggtcataaca gcagcttcag ctacctctca attcaaaaaa cccctcaaga cccgtttaga 120

ggccccaagg ggttatgcta tcaatcgttg cgttacacac acaaaaaacc aacacacatc 180

catcttcgat ggatagcgat tttattatct aactgctgat cgagtgtagc cagatctagt 240

aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta 300

cggtaaatgg cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga 360

cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt 420

tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta 480

ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg 540

actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg ctgatgcggt 600

tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc 660

accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat 720

gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct 780

atataagcag agctggttta gtgaaccgtc agatcagatc tttgtcgatc ctaccatcca 840

ctcgacacac ccgccagcgg ccgcgccacc atggccccta agaagaaaag aaaggtggac 900

tacaaggacc acgacggcga ctacaaagat cacgatatcg actacaagga cgacgatgat 960

aagatgagca cccctctgca gcagcctcac cagaaatcca agaagaccag ccagatgata 1020

acaacacgca agttcaagct cgccatcgtt agcgacaacc ggaacgaggc ctacagcttc 1080

atccggaacg agatcagaaa ccagaacaag gccctgaacg ccgcttataa ccacctgtac 1140

ttcgagcaca tcgcaaccga gaaactgaag cacagcgatg aggaatacca gaagcacctg 1200

acaaagtaca gagaggtggc caccaacaag taccaggact acctgaaggt gaaagaaaaa 1260

gtgaacgaca gcaaggacga cgagaagctg cagaagcggg tggataaggc tagagaggcc 1320

tacaacaagg cccaggagaa agtgtacaag atcgagaagg aattcaacaa aaagagcatg 1380

gaaacctacc agaaggtggt cggcctgagc aaacagacaa gaatcggaaa gctgctcaag 1440

agccagttca ccctgcacta cgacaccgaa gatagaatca ccagcacagt gatcagccac 1500

tttaacaatg atatgaagac aggagtgctg agaggcgata gaagcctgag aacatacaag 1560

aacagccatc ctctgctggt tagagccaga tctatgaagg tgtacgagga aaacggcgac 1620

tacttcatca agtgggtcaa ggggatcgtg tttaagatcg tgatcagcgc cggcagcaag 1680

cagaaagcta atatcggcga gctgaaatct gtgcttatca acatcctgaa tggccactat 1740

aaggtgtgcg acagcagcat cagcctgaat aaagacctga tcctgaacct gtctctgaac 1800

atccctgtgt ctaaggagaa cgtgttcgtg cccggcagag tggtgggcgt ggacctgggc 1860

ctgaagatcc cagcctatgt gtccctgaac gacacaccct acatcaagaa gggcatcggc 1920

aacatcgacg atttcctgaa ggtgcggacc caactgcaga gccagagaaa gagactgcaa 1980

aagaccctgg aatgtacctc cggcggcaag ggcagaaaca agaagctgaa gggactggac 2040

agactgaagg ccaaggaaaa aaacttcgtg aacacctaca accacttcct gagcaagaaa 2100

atcatccagt ttgccgtgaa gaacaacgcc ggcgtgatcc acctggaaga gctgcagttc 2160

gacaagctga aacataagtc cctgctgcgg aactggtcct actaccaact gcagacaatg 2220

atcgagtaca aggctgaacg ggaaggcatc gaggtgaagt acgtggacgc cagctacacc 2280

agccagacct gtagcaagtg cggccactac gaggaaggcc agcgggtgct gcaagacacc 2340

ttcacctgca agaacaaaga gtgcaagggc tacgtgcaca aggttaacgc cgacttcaac 2400

gcctctcaga atatcgccaa gtctaccgac atcatccggt gcaccgagat ggccaagaac 2460

aacgatatcg agaagaatgc caaaagacct gccgctacaa agaaggccgg ccaggccaag 2520

aaaaagaagg gtcgacttga cgcgttgata tcaacaagtt tgtacaaaaa agcaggctac 2580

aaagaggcca gcggttccgg acgggctgac gcattggacg attttgatct ggatatgctg 2640

ggaagtgacg ccctcgatga ttttgacctt gacatgcttg gttcggatgc ccttgatgac 2700

tttgacctcg acatgctcgg cagtgacgcc cttgatgatt tcgacctgga catgctgatt 2760

aactctagaa gttccggatc tggtagccag tacctgcccg acaccgacga ccggcaccgg 2820

atcgaggaaa agcggaagcg gacctacgag acattcaaga gcatcatgaa gaagtccccc 2880

ttcagcggcc ccaccgaccc tagacctcca cctagaagaa tcgccgtgcc cagcagatcc 2940

agcgccagcg tgccaaaacc tgccccccag ccttacccct tcaccagcag cctgagcacc 3000

atcaactacg acgagttccc taccatggtg ttccccagcg gccagatctc tcaggcctct 3060

gctctggctc cagcccctcc tcaggtgctg cctcaggctc ctgctcctgc accagctcca 3120

gccatggtgt ctgcactggc tcaggcacca gcacccgtgc ctgtgctggc tcctggacct 3180

ccacaggctg tggctccacc agcccctaaa cctacacagg ccggcgaggg cacactgtct 3240

gaagctctgc tgcagctgca gttcgacgac gaggatctgg gagccctgct gggaaacagc 3300

accgatcctg ccgtgttcac cgacctggcc agcgtggaca acagcgagtt ccagcagctg 3360

ctgaaccagg gcatccctgt ggcccctcac accaccgagc ccatgctgat ggaatacccc 3420

gaggccatca cccggctcgt gacaggcgct cagaggcctc ctgatccagc tcctgcccct 3480

ctgggagcac caggcctgcc taatggactg ctgtctggcg acgaggactt cagctctatc 3540

gccgatatgg atttctcagc cttgctgggc tctggcagcg gcagccggga ttccagggaa 3600

gggatgtttt tgccgaagcc tgaggccggc tccgctatta gtgacgtgtt tgagggccgc 3660

gaggtgtgcc agccaaaacg aatccggcca tttcatcctc caggaagtcc atgggccaac 3720

cgcccactcc ccgccagcct cgcaccaaca ccaaccggtc cagtacatga gccagtcggg 3780

tcactgaccc cggcaccagt ccctcagcca ctagatccag cgcccgcagt gactcccgag 3840

gccagtcacc tgttggaaga tcccgatgaa gagacgagcc aggctgtcaa agcccttcgg 3900

gagatggccg atactgtgat tccccagaag gaagaggctg caatctgtgg ccaaatggac 3960

ctttcccatc cgcccccaag gggccatctg gatgagctga caaccacact tgagtccatg 4020

accgaggatc tgaacctgga ctcacccctg accccggaat tgaacgagat tctggatacc 4080

ttcctgaacg acgagtgcct cttgcatgcc atgcatatca gcacaggact gtccatcttc 4140

gacacatctc tgttttaatg aggatccgca ggcctctgct agcttgactg actgagatac 4200

agcgtacctt cagctcacag acatgataag atacattgat gagtttggac aaaccacaac 4260

tagaatgcag tgaaaaaaat gctttatttg tgaaatttgt gatgctattg ctttatttgt 4320

aaccattata agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca 4380

ggttcagggg gaggtgtggg aggtttttta aagcaagtaa aacctctaca aatgtggtat 4440

tggcccatct ctatcggtat cgtagcataa ccccttgggg cctctaaacg ggtcttgagg 4500

ggtttttgt gcccctcggg ccggattgct atctaccggc attggcgcag aaaaaaatgc 4560

ctgatgcgac gctgcgcgtc ttatactccc acatatgcca gattcagcaa cggatacggc 4620

ttccccaact tgcccacttc catacgtgtc ctccttacca gaaatttatc cttaaggtcg 4680

tcagctatcc tgcaggcgat ctctcgattt cgatcaagac attcctttaa tggtcttttc 4740

tggacaccac taggggtcag aagtagttca tcaaactttc ttccctccct aatctcattg 4800

gttaccttgg gctatcgaaa cttaattaac cagtcaagtc agctacttgg cgagatcgac 4860

ttgtctgggt ttcgactacg ctcagaattg cgtcagtcaa gttcgatctg gtccttgcta 4920

ttgcacccgt tctccgatta cgagtttcat ttaaatcatg tgagcaaaag gccagcaaaa 4980

ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 5040

cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 5100

ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 5160

taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 5220

ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 5280

ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 5340

aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 5400

tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 5460

agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 5520

ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 5580

tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 5640

tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 5700

cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 5760

aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5820

atttcgttca tccatagttg catttaaatt tccgaactct ccaaggccct cgtcggaaaa 5880

tcttcaaacc tttcgtccga tccatcttgc aggctacctc tcgaacgaac tatcgcaagt 5940

ctcttggccg gccttgcgcc ttggctattg cttggcagcg cctatcgcca ggtattactc 6000

caatcccgaa tatccgagat cgggatcacc cgagagaagt tcaacctaca tcctcaatcc 6060

cgatctatcc gagatccgag gaatatcgaa atcggggcgc gcctggtgta ccgagaacga 6120

tcctctcagt gcgagtctcg acgatccata tcgttgcttg gcagtcagcc agtcggaatc 6180

cagcttggga cccaggaagt ccaatcgtca gatattgtac tcaagcctgg tcacggcagc 6240

gtaccgatct gtttaaacct agatattgat agtctgatcg gtcaacgtat aatcgagtcc 6300

tagcttttgc aaacatctat caagagacag gatcagcagg aggctttcgc atgagtattc 6360

aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttcct gtttttgctc 6420

acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgcg cgagtgggtt 6480

acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgct 6540

ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc cgtattgacg 6600

ccgggcaaga gcaactcggt cgccgcatac actattctca gaatgacttg gttgagtatt 6660

caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgctg 6720

ccataaccat gagtgataac actgcggcca acttacttct gacaacgatt ggaggaccga 6780

aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt gatcgttggg 6840

aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg cctgtagcaa 6900

tggcaacaac cttgcgtaaa ctattaactg gcgaactact tactctagct tcccggcaac 6960

agttgataga ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc 7020

cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct cgcggtatca 7080

ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga 7140

gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc tcactgatta 7200

agcattggta accgattcta ggtgcattgg cgcagaaaaa aatgcctgat gcgacgctgc 7260

gcgtcttata ctcccacata tgccagattc agcaacggat acggcttccc caacttgccc 7320

acttccatac gtgtcctcct taccagaaat ttatccttaa gatcccgaat cgtttaaact 7380

cgactctggc tctatcgaat ctccgtcgtt tcgagcttac gcgaacagcc gtggcgctca 7440

tttgctcgtc gggcatcgaa tctcgtcagc tatcgtcagc ttacctttt ggca 7494

<210> 158
<211> 7494
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 158
gcgatcgcgg ctcccgacat cttggaccat tagctccaca ggtatcttct tccctctagt 60

ggtcataaca gcagcttcag ctacctctca attcaaaaaa cccctcaaga cccgtttaga 120

ggccccaagg ggttatgcta tcaatcgttg cgttacacac acaaaaaacc aacacacatc 180

catcttcgat ggatagcgat tttattatct aactgctgat cgagtgtagc cagatctagt 240

aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta 300

cggtaaatgg cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga 360

cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt 420

tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta 480

ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg 540

actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg ctgatgcggt 600

tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc 660

accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat 720

gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct 780

atataagcag agctggttta gtgaaccgtc agatcagatc tttgtcgatc ctaccatcca 840

ctcgacacac ccgccagcgg ccgcgccacc atggccccta agaagaaaag aaaggtggac 900

tacaaggacc acgacggcga ctacaaagat cacgatatcg actacaagga cgacgatgat 960

aagatgagca cccctctgca gcagcctcac cagaaatcca agaagaccag ccagatgata 1020

acaacacgca agttcaagct cgccatcgtt agcgacaacc ggaacgaggc ctacagcttc 1080

atccggaacg agatcagaaa ccagaacaag gccctgaacg ccgcttataa ccacctgtac 1140

ttcgagcaca tcgcaaccga gaaactgaag cacagcgatg aggaatacca gaagcacctg 1200

acaaagtaca gagaggtggc caccaacaag taccaggact acctgaaggt gaaagaaaaa 1260

gtgaacgaca gcaaggacga cgagaagctg cagaagcggg tggataaggc tagagaggcc 1320

tacaacaagg cccaggagaa agtgtacaag atcgagaagg aattcaacaa aaagagcatg 1380

gaaacctacc agaaggtggt cggcctgagc aaacagacaa gaatcggaaa gctgctcaag 1440

agccagttca ccctgcacta cgacaccgaa gatagaatca ccagcacagt gatcagccac 1500

tttaacaatg atatgaagac aggagtgctg agaggcgata gaagcctgag aacatacaag 1560

aacagccatc ctctgctggt tagagccaga tctatgaagg tgtacgagga aaacggcgac 1620

tacttcatca agtgggtcaa ggggatcgtg tttaagatcg tgatcagcgc cggcagcaag 1680

cagaaagcta atatcggcga gctgaaatct gtgcttatca acatcctgaa tggccactat 1740

aaggtgtgcg acagcagcat cagcctgaat aaagacctga tcctgaacct gtctctgaac 1800

atccctgtgt ctaaggagaa cgtgttcgtg cccggcagag tggtgggcgt ggccctgggc 1860

ctgaagatcc cagcctatgt gtccctgaac gacacaccct acatcaagaa gggcatcggc 1920

aacatcgacg atttcctgaa ggtgcggacc caactgcaga gccagagaaa gagactgcaa 1980

aagaccctgg aatgtacctc cggcggcaag ggcagaaaca agaagctgaa gggactggac 2040

agactgaagg ccaaggaaaa aaacttcgtg aacacctaca accacttcct gagcaagaaa 2100

atcatccagt ttgccgtgaa gaacaacgcc ggcgtgatcc acctggccga gctgcagttc 2160

gacaagctga aacataagtc cctgctgcgg aactggtcct actaccaact gcagacaatg 2220

atcgagtaca aggctgaacg ggaaggcatc gaggtgaagt acgtggacgc cagctacacc 2280

agccagacct gtagcaagtg cggccactac gaggaaggcc agcgggtgct gcaagacacc 2340

ttcacctgca agaacaaaga gtgcaagggc tacgtgcaca aggttaacgc cgccttcaac 2400

gcctctcaga atatcgccaa gtctaccgac atcatccggt gcaccgagat ggccaagaac 2460

aacgatatcg agaagaatgc caaaagacct gccgctacaa agaaggccgg ccaggccaag 2520

aaaaagaagg gtcgacttga cgcgttgata tcaacaagtt tgtacaaaaa agcaggctac 2580

aaagaggcca gcggttccgg acgggctgac gcattggacg attttgatct ggatatgctg 2640

ggaagtgacg ccctcgatga ttttgacctt gacatgcttg gttcggatgc ccttgatgac 2700

tttgacctcg acatgctcgg cagtgacgcc cttgatgatt tcgacctgga catgctgatt 2760

aactctagaa gttccggatc tggtagccag tacctgcccg acaccgacga ccggcaccgg 2820

atcgaggaaa agcggaagcg gacctacgag acattcaaga gcatcatgaa gaagtccccc 2880

ttcagcggcc ccaccgaccc tagacctcca cctagaagaa tcgccgtgcc cagcagatcc 2940

agcgccagcg tgccaaaacc tgccccccag ccttacccct tcaccagcag cctgagcacc 3000

atcaactacg acgagttccc taccatggtg ttccccagcg gccagatctc tcaggcctct 3060

gctctggctc cagcccctcc tcaggtgctg cctcaggctc ctgctcctgc accagctcca 3120

gccatggtgt ctgcactggc tcaggcacca gcacccgtgc ctgtgctggc tcctggacct 3180

ccacaggctg tggctccacc agcccctaaa cctacacagg ccggcgaggg cacactgtct 3240

gaagctctgc tgcagctgca gttcgacgac gaggatctgg gagccctgct gggaaacagc 3300

accgatcctg ccgtgttcac cgacctggcc agcgtggaca acagcgagtt ccagcagctg 3360

ctgaaccagg gcatccctgt ggcccctcac accaccgagc ccatgctgat ggaatacccc 3420

gaggccatca cccggctcgt gacaggcgct cagaggcctc ctgatccagc tcctgccct 3480

ctgggagcac caggcctgcc taatggactg ctgtctggcg acgaggactt cagctctatc 3540

gccgatatgg atttctcagc cttgctgggc tctggcagcg gcagccggga ttccagggaa 3600

gggatgtttt tgccgaagcc tgaggccggc tccgctatta gtgacgtgtt tgagggccgc 3660

gaggtgtgcc agccaaaacg aatccggcca tttcatcctc caggaagtcc atgggccaac 3720

cgcccactcc ccgccagcct cgcaccaaca ccaaccggtc cagtacatga gccagtcggg 3780

tcactgaccc cggcaccagt ccctcagcca ctagatccag cgcccgcagt gactcccgag 3840

gccagtcacc tgttggaaga tcccgatgaa gagacgagcc aggctgtcaa agcccttcgg 3900

gagatggccg atactgtgat tccccagaag gaagaggctg caatctgtgg ccaaatggac 3960

ctttcccatc cgcccccaag gggccatctg gatgagctga caaccacact tgagtccatg 4020

accgaggatc tgaacctgga ctcacccctg accccggaat tgaacgagat tctggatacc 4080

ttcctgaacg acgagtgcct cttgcatgcc atgcatatca gcacaggact gtccatcttc 4140

gacacatctc tgttttaatg aggatccgca ggcctctgct agcttgactg actgagatac 4200

agcgtacctt cagctcacag acatgataag atacattgat gagtttggac aaaccacaac 4260

tagaatgcag tgaaaaaaat gctttatttg tgaaatttgt gatgctattg ctttatttgt 4320

aaccattata agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca 4380

ggttcagggg gaggtgtggg aggtttttta aagcaagtaa aacctctaca aatgtggtat 4440

tggcccatct ctatcggtat cgtagcataa ccccttgggg cctctaaacg ggtcttgagg 4500

ggttttttgt gcccctcggg ccggattgct atctaccggc attggcgcag aaaaaaatgc 4560

ctgatgcgac gctgcgcgtc ttatactccc acatatgcca gattcagcaa cggatacggc 4620

ttccccaact tgcccacttc catacgtgtc ctccttacca gaaatttatc cttaaggtcg 4680

tcagctatcc tgcaggcgat ctctcgattt cgatcaagac attcctttaa tggtcttttc 4740

tggacaccac taggggtcag aagtagttca tcaaactttc ttccctccct aatctcattg 4800

gttaccttgg gctatcgaaa cttaattaac cagtcaagtc agctacttgg cgagatcgac 4860

ttgtctgggt ttcgactacg ctcagaattg cgtcagtcaa gttcgatctg gtccttgcta 4920

ttgcacccgt tctccgatta cgagtttcat ttaaatcatg tgagcaaaag gccagcaaaa 4980

ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 5040

cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 5100

ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 5160

taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 5220

ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 5280

ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 5340

aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 5400

tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 5460

agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 5520

ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 5580

tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 5640

tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 5700

cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 5760

aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5820

atttcgttca tccatagttg catttaaatt tccgaactct ccaaggccct cgtcggaaaa 5880

tcttcaaacc tttcgtccga tccatcttgc aggctacctc tcgaacgaac tatcgcaagt 5940

ctcttggccg gccttgcgcc ttggctattg cttggcagcg cctatcgcca ggtattactc 6000

caatcccgaa tatccgagat cgggatcacc cgagagaagt tcaacctaca tcctcaatcc 6060

cgatctatcc gagatccgag gaatatcgaa atcggggcgc gcctggtgta ccgagaacga 6120

tcctctcagt gcgagtctcg acgatccata tcgttgcttg gcagtcagcc agtcggaatc 6180

cagcttggga cccaggaagt ccaatcgtca gatattgtac tcaagcctgg tcacggcagc 6240

gtaccgatct gtttaaacct agatattgat agtctgatcg gtcaacgtat aatcgagtcc 6300

tagcttttgc aaacatctat caagagacag gatcagcagg aggctttcgc atgagtattc 6360

aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttcct gtttttgctc 6420

acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgcg cgagtgggtt 6480

acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgct 6540

ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc cgtattgacg 6600

ccgggcaaga gcaactcggt cgccgcatac actattctca gaatgacttg gttgagtatt 6660

caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgctg 6720

ccataaccat gagtgataac actgcggcca acttacttct gacaacgatt ggaggaccga 6780

aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt gatcgttggg 6840

aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg cctgtagcaa 6900

tggcaacaac cttgcgtaaa ctattaactg gcgaactact tactctagct tcccggcaac 6960

agttgataga ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc 7020

cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct cgcggtatca 7080

ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga 7140

gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc tcactgatta 7200

agcattggta accgattcta ggtgcattgg cgcagaaaaa aatgcctgat gcgacgctgc 7260

gcgtcttata ctcccacata tgccagattc agcaacggat acggcttccc caacttgccc 7320

acttccatac gtgtcctcct taccagaaat ttatccttaa gatcccgaat cgtttaaact 7380

cgactctggc tctatcgaat ctccgtcgtt tcgagcttac gcgaacagcc gtggcgctca 7440

tttgctcgtc gggcatcgaa tctcgtcagc tatcgtcagc ttacctttt ggca 7494

<210> 159
<211> 7512
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 159
gcgatcgcgg ctcccgacat cttggaccat tagctccaca ggtatcttct tccctctagt 60

ggtcataaca gcagcttcag ctacctctca attcaaaaaa cccctcaaga cccgtttaga 120

ggccccaagg ggttatgcta tcaatcgttg cgttacacac acaaaaaacc aacacacatc 180

catcttcgat ggatagcgat tttattatct aactgctgat cgagtgtagc cagatctagt 240

aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta 300

cggtaaatgg cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga 360

cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt 420

tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta 480

ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg 540

actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg ctgatgcggt 600

tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc 660

accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat 720

gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct 780

atataagcag agctggttta gtgaaccgtc agatcagatc tttgtcgatc ctaccatcca 840

ctcgacacac ccgccagcgg ccgcgccacc atggccccta agaagaaaag aaaggtggac 900

tacaaggacc acgacggcga ctacaaagat cacgatatcg actacaagga cgacgatgat 960

aaggacgcat tggacgattt tgatctggat atgctgggaa gtgacgccct cgatgatttt 1020

gaccttgaca tgcttggttc ggatgccctt gatgactttg acctcgacat gctcggcagt 1080

gacgcccttg atgatttcga cctggacatg ctgattaact ctagaagttc cggatctggt 1140

agccagtacc tgcccgacac cgacgaccgg caccggatcg aggaaaagcg gaagcggacc 1200

tacgagacat tcaagagcat catgaagaag tccccccttca gcggccccac cgaccctaga 1260

cctccaccta gaagaatcgc cgtgcccagc agatccagcg ccagcgtgcc aaaacctgcc 1320

ccccagcctt acccccttcac cagcagcctg agcaccatca actacgacga gttccctacc 1380

atggtgttcc ccagcggcca gatctctcag gcctctgctc tggctccagc ccctcctcag 1440

gtgctgcctc aggctcctgc tcctgcacca gctccagcca tggtgtctgc actggctcag 1500

gcaccagcac ccgtgcctgt gctggctcct ggacctccac aggctgtggc tccaccagcc 1560

cctaaaccta cacaggccgg cgagggcaca ctgtctgaag ctctgctgca gctgcagttc 1620

gacgacgagg atctgggagc cctgctggga aacagcaccg atcctgccgt gttcaccgac 1680

ctggccagcg tggacaacag cgagttccag cagctgctga accagggcat ccctgtggcc 1740

cctcacacca ccgagcccat gctgatggaa taccccgagg ccatcacccg gctcgtgaca 1800

ggcgctcaga ggcctcctga tccagctcct gcccctctgg gagcaccagg cctgcctaat 1860

ggactgctgt ctggcgacga ggacttcagc tctatcgccg atatggattt ctcagccttg 1920

ctgggctctg gcagcggcag ccgggattcc agggaaggga tgtttttgcc gaagcctgag 1980

gccggctccg ctattagtga cgtgtttgag ggccgcgagg tgtgccagcc aaaacgaatc 2040

cggccatttc atcctccagg aagtccatgg gccaaccgcc cactccccgc cagcctcgca 2100

ccaacaccaa ccggtccagt acatgagcca gtcgggtcac tgaccccggc accagtccct 2160

cagccactag atccagcgcc cgcagtgact cccgaggcca gtcacctgtt ggaagatccc 2220

gatgaagaga cgagccaggc tgtcaaagcc cttcgggaga tggccgatac tgtgattccc 2280

cagaaggaag aggctgcaat ctgtggccaa atggaccttt cccatccgcc cccaaggggc 2340

catctggatg agctgacaac cacacttgag tccatgaccg aggatctgaa cctggactca 2400

cccctgaccc cggaattgaa cgagattctg gataccttcc tgaacgacga gtgcctcttg 2460

catgccatgc atatcagcac aggactgtcc atcttcgaca catctctgtt ttccggagga 2520

tctagcggag gctcctctgg ctctgagaca cctggcacaa gcgagagcgc aacacctgaa 2580

agcagcgggg gcagcagcgg ggggtcaatg agcacccctc tgcagcagcc tcaccagaaa 2640

tccaagaaga ccagccagat gataacaaca cgcaagttca agctcgccat cgttagcgac 2700

aaccggaacg aggcctacag cttcatccgg aacgagatca gaaaccagaa caaggccctg 2760

aacgccgctt ataaccacct gtacttcgag cacatcgcaa ccgagaaact gaagcacagc 2820

gatgaggaat accagaagca cctgacaaag tacagagagg tggccaccaa caagtaccag 2880

gactacctga aggtgaaaga aaaagtgaac gacagcaagg acgacgagaa gctgcagaag 2940

cgggtggata aggctagaga ggcctacaac aaggcccagg agaaagtgta caagatcgag 3000

aaggaattca acaaaaagag catggaaacc taccagaagg tggtcggcct gagcaaacag 3060

acaagaatcg gaaagctgct caagagccag ttcaccctgc actacgacac cgaagataga 3120

atcaccagca cagtgatcag ccactttaac aatgatatga agacaggagt gctgagaggc 3180

gatagaagcc tgagaacata caagaacagc catcctctgc tggttagagc cagatctatg 3240

aaggtgtacg aggaaaacgg cgactacttc atcaagtggg tcaaggggat cgtgtttaag 3300

atcgtgatca gcgccggcag caagcagaaa gctaatatcg gcgagctgaa atctgtgctt 3360

atcaacatcc tgaatggcca ctataaggtg tgcgacagca gcatcagcct gaataaagac 3420

ctgatcctga acctgtctct gaacatccct gtgtctaagg agaacgtgtt cgtgcccggc 3480

agagtggtgg gcgtggacct gggcctgaag atcccagcct atgtgtccct gaacgacaca 3540

ccctacatca agaagggcat cggcaacatc gacgatttcc tgaaggtgcg gacccaactg 3600

cagagccaga gaaagagact gcaaaagacc ctggaatgta cctccggcgg caagggcaga 3660

aacaagaagc tgaagggact ggacagactg aaggccaagg aaaaaaactt cgtgaacacc 3720

tacaaccact tcctgagcaa gaaaatcatc cagtttgccg tgaagaacaa cgccggcgtg 3780

atccacctgg aagagctgca gttcgacaag ctgaaacata agtccctgct gcggaactgg 3840

tcctactacc aactgcagac aatgatcgag tacaaggctg aacgggaagg catcgaggtg 3900

aagtacgtgg acgccagcta caccagccag acctgtagca agtgcggcca ctacgaggaa 3960

ggccagcggg tgctgcaaga caccttcacc tgcaagaaca aagagtgcaa gggctacgtg 4020

cacaaggtta acgccgactt caacgcctct cagaatatcg ccaagtctac cgacatcatc 4080

cggtgcaccg agatggccaa gaacaacgat atcgagaaga atgccaaaag acctgccgct 4140

acaaagaagg ccggccaggc caagaaaaag aagtaatgag gatccgcagg cctctgctag 4200

cttgactgac tgagatacag cgtaccttca gctcacagac atgataagat acattgatga 4260

gtttggacaa accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga 4320

tgctattgct ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg 4380

cattcatttt atgtttcagg ttcaggggga ggtgtgggag gtttttttaaa gcaagtaaaa 4440

cctctacaaa tgtggtattg gcccatctct atcggtatcg tagcataacc ccttggggcc 4500

tctaaacggg tcttgagggg ttttttgtgc ccctcgggcc ggattgctat ctaccggcat 4560

tggcgcagaa aaaaatgcct gatgcgacgc tgcgcgtctt atactcccac atatgccaga 4620

ttcagcaacg gatacggctt ccccaacttg cccacttcca tacgtgtcct ccttaccaga 4680

aatttatcct taaggtcgtc agctatcctg caggcgatct ctcgatttcg atcaagacat 4740

tcctttaatg gtcttttctg gacaccacta ggggtcagaa gtagttcatc aaactttctt 4800

ccctccctaa tctcattggt taccttgggc tatcgaaact taattaacca gtcaagtcag 4860

ctacttggcg agatcgactt gtctgggttt cgactacgct cagaattgcg tcagtcaagt 4920

tcgatctggt ccttgctatt gcacccgttc tccgattacg agtttcattt aaatcatgtg 4980

agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 5040

taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 5100

cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 5160

tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 5220

gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct 5280

gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 5340

tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 5400

gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta 5460

cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 5520

aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 5580

tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt 5640

ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag 5700

attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat 5760

ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc 5820

tatctcagcg atctgtctat ttcgttcatc catagttgca tttaaatttc cgaactctcc 5880

aaggccctcg tcggaaaatc ttcaaacctt tcgtccgatc catcttgcag gctacctctc 5940

gaacgaacta tcgcaagtct cttggccggc cttgcgcctt ggctattgct tggcagcgcc 6000

tatcgccagg tattactcca atcccgaata tccgagatcg ggatcacccg agagaagttc 6060

aacctacatc ctcaatcccg atctatccga gatccgagga atatcgaaat cggggcgcgc 6120

ctggtgtacc gagaacgatc ctctcagtgc gagtctcgac gatccatatc gttgcttggc 6180

agtcagccag tcggaatcca gcttgggacc caggaagtcc aatcgtcaga tattgtactc 6240

aagcctggtc acggcagcgt accgatctgt ttaaacctag atattgatag tctgatcggt 6300

caacgtataa tcgagtccta gcttttgcaa acatctatca agagacagga tcagcaggag 6360

gctttcgcat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 6420

gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 6480

tgggtgcgcg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt 6540

ttcgccccga agaacgcttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg 6600

tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 6660

atgacttggt tgagtattca ccagtcacag aaaagcatct tacggatggc atgacagtaa 6720

gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac ttacttctga 6780

caacgattgg aggaccgaag gagctaaccg ctttttttgca caacatgggg gatcatgtaa 6840

ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 6900

ccacgatgcc tgtagcaatg gcaacaacct tgcgtaaact attaactggc gaactactta 6960

ctctagcttc ccggcaacag ttgatagact ggatggaggc ggataaagtt gcaggaccac 7020

ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc 7080

gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 7140

ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga 7200

taggtgcctc actgattaag cattggtaac cgattctagg tgcattggcg cagaaaaaaa 7260

tgcctgatgc gacgctgcgc gtcttatact cccacatatg ccagattcag caacggatac 7320

ggcttcccca acttgcccac ttccatacgt gtcctcctta ccagaaattt atccttaaga 7380

tcccgaatcg tttaaactcg actctggctc tatcgaatct ccgtcgtttc gagcttacgc 7440

gaacagccgt ggcgctcatt tgctcgtcgg gcatcgaatc tcgtcagcta tcgtcagctt 7500

acctttttgg ca 7512

<210> 160
<211> 7512
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 160
gcgatcgcgg ctcccgacat cttggaccat tagctccaca ggtatcttct tccctctagt 60

ggtcataaca gcagcttcag ctacctctca attcaaaaaa cccctcaaga cccgtttaga 120

ggccccaagg ggttatgcta tcaatcgttg cgttacacac acaaaaaacc aacacacatc 180

catcttcgat ggatagcgat tttattatct aactgctgat cgagtgtagc cagatctagt 240

aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta 300

cggtaaatgg cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga 360

cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt 420

tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta 480

ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg 540

actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg ctgatgcggt 600

tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc 660

accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat 720

gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct 780

atataagcag agctggttta gtgaaccgtc agatcagatc tttgtcgatc ctaccatcca 840

ctcgacacac ccgccagcgg ccgcgccacc atggccccta agaagaaaag aaaggtggac 900

tacaaggacc acgacggcga ctacaaagat cacgatatcg actacaagga cgacgatgat 960

aaggacgcat tggacgattt tgatctggat atgctgggaa gtgacgccct cgatgatttt 1020

gaccttgaca tgcttggttc ggatgccctt gatgactttg acctcgacat gctcggcagt 1080

gacgcccttg atgatttcga cctggacatg ctgattaact ctagaagttc cggatctggt 1140

agccagtacc tgcccgacac cgacgaccgg caccggatcg aggaaaagcg gaagcggacc 1200

tacgagacat tcaagagcat catgaagaag tccccccttca gcggccccac cgaccctaga 1260

cctccaccta gaagaatcgc cgtgcccagc agatccagcg ccagcgtgcc aaaacctgcc 1320

ccccagcctt acccccttcac cagcagcctg agcaccatca actacgacga gttccctacc 1380

atggtgttcc ccagcggcca gatctctcag gcctctgctc tggctccagc ccctcctcag 1440

gtgctgcctc aggctcctgc tcctgcacca gctccagcca tggtgtctgc actggctcag 1500

gcaccagcac ccgtgcctgt gctggctcct ggacctccac aggctgtggc tccaccagcc 1560

cctaaaccta cacaggccgg cgagggcaca ctgtctgaag ctctgctgca gctgcagttc 1620

gacgacgagg atctgggagc cctgctggga aacagcaccg atcctgccgt gttcaccgac 1680

ctggccagcg tggacaacag cgagttccag cagctgctga accagggcat ccctgtggcc 1740

cctcacacca ccgagcccat gctgatggaa taccccgagg ccatcacccg gctcgtgaca 1800

ggcgctcaga ggcctcctga tccagctcct gcccctctgg gagcaccagg cctgcctaat 1860

ggactgctgt ctggcgacga ggacttcagc tctatcgccg atatggattt ctcagccttg 1920

ctgggctctg gcagcggcag ccgggattcc agggaaggga tgtttttgcc gaagcctgag 1980

gccggctccg ctattagtga cgtgtttgag ggccgcgagg tgtgccagcc aaaacgaatc 2040

cggccatttc atcctccagg aagtccatgg gccaaccgcc cactccccgc cagcctcgca 2100

ccaacaccaa ccggtccagt acatgagcca gtcgggtcac tgaccccggc accagtccct 2160

cagccactag atccagcgcc cgcagtgact cccgaggcca gtcacctgtt ggaagatccc 2220

gatgaagaga cgagccaggc tgtcaaagcc cttcgggaga tggccgatac tgtgattccc 2280

cagaaggaag aggctgcaat ctgtggccaa atggaccttt cccatccgcc cccaaggggc 2340

catctggatg agctgacaac cacacttgag tccatgaccg aggatctgaa cctggactca 2400

cccctgaccc cggaattgaa cgagattctg gataccttcc tgaacgacga gtgcctcttg 2460

catgccatgc atatcagcac aggactgtcc atcttcgaca catctctgtt ttccggagga 2520

tctagcggag gctcctctgg ctctgagaca cctggcacaa gcgagagcgc aacacctgaa 2580

agcagcgggg gcagcagcgg ggggtcaatg agcacccctc tgcagcagcc tcaccagaaa 2640

tccaagaaga ccagccagat gataacaaca cgcaagttca agctcgccat cgttagcgac 2700

aaccggaacg aggcctacag cttcatccgg aacgagatca gaaaccagaa caaggccctg 2760

aacgccgctt ataaccacct gtacttcgag cacatcgcaa ccgagaaact gaagcacagc 2820

gatgaggaat accagaagca cctgacaaag tacagagagg tggccaccaa caagtaccag 2880

gactacctga aggtgaaaga aaaagtgaac gacagcaagg acgacgagaa gctgcagaag 2940

cgggtggata aggctagaga ggcctacaac aaggcccagg agaaagtgta caagatcgag 3000

aaggaattca acaaaaagag catggaaacc taccagaagg tggtcggcct gagcaaacag 3060

acaagaatcg gaaagctgct caagagccag ttcaccctgc actacgacac cgaagataga 3120

atcaccagca cagtgatcag ccactttaac aatgatatga agacaggagt gctgagaggc 3180

gatagaagcc tgagaacata caagaacagc catcctctgc tggttagagc cagatctatg 3240

aaggtgtacg aggaaaacgg cgactacttc atcaagtggg tcaaggggat cgtgtttaag 3300

atcgtgatca gcgccggcag caagcagaaa gctaatatcg gcgagctgaa atctgtgctt 3360

atcaacatcc tgaatggcca ctataaggtg tgcgacagca gcatcagcct gaataaagac 3420

ctgatcctga acctgtctct gaacatccct gtgtctaagg agaacgtgtt cgtgcccggc 3480

agagtggtgg gcgtggccct gggcctgaag atcccagcct atgtgtccct gaacgacaca 3540

ccctacatca agaagggcat cggcaacatc gacgatttcc tgaaggtgcg gacccaactg 3600

cagagccaga gaaagagact gcaaaagacc ctggaatgta cctccggcgg caagggcaga 3660

aacaagaagc tgaagggact ggacagactg aaggccaagg aaaaaaactt cgtgaacacc 3720

tacaaccact tcctgagcaa gaaaatcatc cagtttgccg tgaagaacaa cgccggcgtg 3780

atccacctgg ccgagctgca gttcgacaag ctgaaacata agtccctgct gcggaactgg 3840

tcctactacc aactgcagac aatgatcgag tacaaggctg aacgggaagg catcgaggtg 3900

aagtacgtgg acgccagcta caccagccag acctgtagca agtgcggcca ctacgaggaa 3960

ggccagcggg tgctgcaaga caccttcacc tgcaagaaca aagagtgcaa gggctacgtg 4020

cacaaggtta acgccgcctt caacgcctct cagaatatcg ccaagtctac cgacatcatc 4080

cggtgcaccg agatggccaa gaacaacgat atcgagaaga atgccaaaag acctgccgct 4140

acaaagaagg ccggccaggc caagaaaaag aagtaatgag gatccgcagg cctctgctag 4200

cttgactgac tgagatacag cgtaccttca gctcacagac atgataagat acattgatga 4260

gtttggacaa accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga 4320

tgctattgct ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg 4380

cattcatttt atgtttcagg ttcaggggga ggtgtgggag gtttttttaaa gcaagtaaaa 4440

cctctacaaa tgtggtattg gcccatctct atcggtatcg tagcataacc ccttggggcc 4500

tctaaacggg tcttgagggg ttttttgtgc ccctcgggcc ggattgctat ctaccggcat 4560

tggcgcagaa aaaaatgcct gatgcgacgc tgcgcgtctt atactcccac atatgccaga 4620

ttcagcaacg gatacggctt ccccaacttg cccacttcca tacgtgtcct ccttaccaga 4680

aatttatcct taaggtcgtc agctatcctg caggcgatct ctcgatttcg atcaagacat 4740

tcctttaatg gtcttttctg gacaccacta ggggtcagaa gtagttcatc aaactttctt 4800

ccctccctaa tctcattggt taccttgggc tatcgaaact taattaacca gtcaagtcag 4860

ctacttggcg agatcgactt gtctgggttt cgactacgct cagaattgcg tcagtcaagt 4920

tcgatctggt ccttgctatt gcacccgttc tccgattacg agtttcattt aaatcatgtg 4980

agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 5040

taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 5100

cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 5160

tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 5220

gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct 5280

gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 5340

tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 5400

gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta 5460

cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 5520

aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 5580

tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt 5640

ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag 5700

attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat 5760

ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc 5820

tatctcagcg atctgtctat ttcgttcatc catagttgca tttaaatttc cgaactctcc 5880

aaggccctcg tcggaaaatc ttcaaacctt tcgtccgatc catcttgcag gctacctctc 5940

gaacgaacta tcgcaagtct cttggccggc cttgcgcctt ggctattgct tggcagcgcc 6000

tatcgccagg tattactcca atcccgaata tccgagatcg ggatcacccg agagaagttc 6060

aacctacatc ctcaatcccg atctatccga gatccgagga atatcgaaat cggggcgcgc 6120

ctggtgtacc gagaacgatc ctctcagtgc gagtctcgac gatccatatc gttgcttggc 6180

agtcagccag tcggaatcca gcttgggacc caggaagtcc aatcgtcaga tattgtactc 6240

aagcctggtc acggcagcgt accgatctgt ttaaacctag atattgatag tctgatcggt 6300

caacgtataa tcgagtccta gcttttgcaa acatctatca agagacagga tcagcaggag 6360

gctttcgcat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 6420

gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 6480

tgggtgcgcg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt 6540

ttcgccccga agaacgcttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg 6600

tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 6660

atgacttggt tgagtattca ccagtcacag aaaagcatct tacggatggc atgacagtaa 6720

gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac ttacttctga 6780

caacgattgg aggaccgaag gagctaaccg ctttttttgca caacatgggg gatcatgtaa 6840

ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 6900

ccacgatgcc tgtagcaatg gcaacaacct tgcgtaaact attaactggc gaactactta 6960

ctctagcttc ccggcaacag ttgatagact ggatggaggc ggataaagtt gcaggaccac 7020

ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc 7080

gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 7140

ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga 7200

taggtgcctc actgattaag cattggtaac cgattctagg tgcattggcg cagaaaaaaa 7260

tgcctgatgc gacgctgcgc gtcttatact cccacatatg ccagattcag caacggatac 7320

ggcttcccca acttgcccac ttccatacgt gtcctcctta ccagaaattt atccttaaga 7380

tcccgaatcg tttaaactcg actctggctc tatcgaatct ccgtcgtttc gagcttacgc 7440

gaacagccgt ggcgctcatt tgctcgtcgg gcatcgaatc tcgtcagcta tcgtcagctt 7500

acctttttgg ca 7512

<210> 161
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 161
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa taagctctcg gggtgtgga 239

<210> 162
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 162
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tactgtaaaa gatgtaaag 239

<210> 163
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 163
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tatgttttat gttactgta 239

<210> 164
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 164
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tacagtaaca taaaacata 239

<210> 165
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 165
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tggcaaagga gcacatcag 239

<210> 166
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 166
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tgccaaagca gatgtgttt 239

<210> 167
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic

polynucleotide"

<400> 167
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tgtcctgtag acaggcatg 239

<210> 168
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 168
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tggtggtgtg gggggggag 239

<210> 169
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 169
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tccacactca cgcatgcct 239

<210> 170
<211> 239

<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 170
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tcagcacagc agcttcccc 239

<210> 171
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 171
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa tgcccggatg agggagagc 239

<210> 172
<211> 239
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 172
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt 60

ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg 120

gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc 180

aaagggtcaa aaagatccta catggatgtg aggattgaaa taccctccc cccccacac 239

<210> 173
<211> 506
<212> PRT
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polypeptide"

<400> 173
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
1 5 10 15

Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
20 25 30

Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
35 40 45

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

Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
65 70 75 80

Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
85 90 95

Lys Glu Lys Val Asn Asp Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
100 105 110

Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
115 120 125

Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
130 135 140

Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
145 150 155 160

```
Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
                165                 170                 175

Ile Ser His Phe Asn Asn Asp Met Lys Thr Gly Val Leu Arg Gly Asp
            180                 185                 190

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
        195                 200                 205

Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
    210                 215                 220

Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
225                 230                 235                 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn
                245                 250                 255

Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu
            260                 265                 270

Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe
        275                 280                 285

Val Pro Gly Arg Val Val Gly Val Ala Leu Gly Leu Lys Ile Pro Ala
    290                 295                 300

Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn
305                 310                 315                 320

Ile Asp Asp Phe Leu Lys Val Arg Thr Gln Leu Gln Ser Gln Arg Lys
                325                 330                 335

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Asn
            340                 345                 350

Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe
        355                 360                 365

Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala
```

```
    370                 375                 380

Val Lys Asn Asn Ala Gly Val Ile His Leu Ala Glu Leu Gln Phe Asp
385                 390                 395                 400

Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
                405                 410                 415

Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys
            420                 425                 430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
        435                 440                 445

Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
    450                 455                 460

Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Ala Phe Asn Ala
465                 470                 475                 480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
                485                 490                 495

Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
            500                 505


<210> 174
<211> 56
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
      oligonucleotide"

<400> 174
ctacgaattc gtttgccatt actttagttc ccattccaaa aaccggatcc atatgg          56


<210> 175
<211> 242
<212> RNA
<213> Artificial Sequence
```

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 175
gggaaacagc uagaauguaa cuuaaaguag gucaauguuu aaauucgaug uugcaauuug 60

uuuggacaag uggauuaaaa cguuccuuga aaaucauaua aagcagccag uuuacgggcu 120

ugggcgaauu ugcguccaaa gggugaggcc agguguaagu aagaaccuac aaaagcacuc 180

accaaagggu caaaaagauc cuacauggau gugaggauug aaauuggaau gggaacuaaa 240

gu 242

<210> 176
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<220>
<221> modified_base
<222> (1)..(8)
<223> a, c, t, g, unknown or other

<400> 176
nnnnnnnn 8

<210> 177
<211> 199
<212> PRT
<213> Homo sapiens

<400> 177
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1 5 10 15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
20 25 30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
35 40 45

```
Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
                165                 170                 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
            180                 185                 190

Ile Leu Gln Asn Gln Gly Asn
        195
```

<210> 178
<211> 173
<212> PRT
<213> Gordonia sp.

<400> 178

```
Met Ser Gly Met Lys Ile Arg Ala Tyr Ala Pro Ala Gly Gly Pro Ser
1               5                   10                  15

Leu Ala Glu Val Ala Leu Ser Asn Val Thr Ala Leu Cys Asn Arg Trp
            20                  25                  30
```

```
Leu Asp Gln Pro Ala Ser Glu Val Trp Arg Glu Gln Ala Ala Asn Glu
        35                  40                  45

Ile Leu Ala Ala Met Asp Thr Asp Pro Val Ser Asp Leu Gly Ile Val
    50                  55                  60

Gly Leu Asn Val Lys Asn Gly Glu Ala Thr Leu Asp Val Ala Ser Cys
65                  70                  75                  80

Gly Pro Ala Ser Gln Ala Thr Leu Leu Ala Val Val Ala Ala Ile Gly
                85                  90                  95

Asp Leu Leu His Thr Glu Asn Pro Pro Asn Tyr Val Glu Phe Glu Val
            100                 105                 110

Ile Gln Gly Thr Arg Ser Tyr Ile Ala His Val Arg Thr Gly Asp Arg
        115                 120                 125

Pro Ser Pro His Thr Leu Arg Arg Gln Ala Glu Gln Arg Ile Thr Ala
    130                 135                 140

Ala Leu Asp Val Leu Asp Asp His Ala Asp Asp Pro Thr Glu Val Ile
145                 150                 155                 160

Ala Arg Ala Arg Ala Val Leu Arg Glu Gln Ala Glu Gln
                165                 170
```

<210> 179
<211> 135
<212> PRT
<213> Gordonia sp.

<400> 179

```
Met Thr Ala Thr Pro Pro Ala Ala Glu Thr Val Trp Ala Val Gln Phe
1               5                   10                  15

Thr Gly Val Gly Gly Glu Pro Val Val His Thr Arg Asp Leu Arg Gly
            20                  25                  30

Val Pro Phe Ala Asp Asp Ala His Ala Ala Ala His Ala Thr His Leu
        35                  40                  45
```

```
Gln Gly Leu Gly Arg Arg Asp Ala Arg Thr Val Tyr Gln Ala Asn Gly
    50                  55                  60

Ser Thr Gly Trp Met Asp Pro Thr Pro Gln Pro Pro Gln Gly Gln Pro
65                  70                  75                  80

Pro Asp Arg Ala Ala Asp Val Thr Glu Val Ala Ser Ala Leu His Arg
                85                  90                  95

His Ser Gly Gly Arg Leu Ala Val Thr Asp Ala His Ser Ile Ala Ala
            100                 105                 110

Val Ala Leu Arg Met Ala Ala Ala Val Thr Glu Ala Arg Ala Pro Ala
        115                 120                 125

Ala Arg Ser Arg His Ser Ile
    130                 135
```

<210> 180
<211> 191
<212> PRT
<213> Gordonia sp.

<400> 180

```
Met Thr Asp Pro Thr Ala Ser Ala Gly Thr His Pro Ala Ala Leu Ala
1               5                   10                  15

Ala Val Ala Val Ala Asn Ser Thr Ala Ala Thr Val Gln Ala Ile Cys
            20                  25                  30

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

Ala Leu Pro His Arg Thr Gln His Ala Gly Arg Arg Phe Arg Ile Val
    50                  55                  60

Ala Glu His Asp Asp Arg Val Ala Ile Glu Phe Thr Asp Arg Gly Asp
65                  70                  75                  80

Arg Met Leu Ala Tyr Pro Glu Glu Ile Thr Ala Asp Thr Tyr Gly Tyr
                85                  90                  95
```

```
Pro Glu Gly Ala His Leu Ile Arg Ala Met Tyr Ala Glu Ala Ala Glu
            100                 105                 110

Ser Leu Ser His Arg Ile Phe Met Leu Glu Phe Thr Arg Glu Trp Trp
        115                 120                 125

Leu Glu Leu Ser Glu Gly Arg Gly Gly Pro Leu Asp Thr Ala Pro Gln
    130                 135                 140

Ser Val Val Leu Leu Gly Ser Arg Met Val Arg Ala Ala Asp Gly Gly
145                 150                 155                 160

Arg Ala Val Arg Leu Tyr Arg Arg Ala Ala Asp Gly Ala Pro Asp Val
                165                 170                 175

Gly Met Leu Arg Ser Ala Ala Arg Asp Phe Val Gly Cys Leu His
            180                 185                 190
```

<210> 181
<211> 234
<212> PRT
<213> Gordonia sp.

<400> 181

```
Met Thr Ala Ala Ala Asp Pro Val Asp Met Phe Val Val Val Ile Glu
1               5                   10                  15

His Arg Tyr Gly Thr Asn Val Gly Val His Ala Ser Arg Gly Asp Ala
            20                  25                  30

Val Thr Glu Val Ala Gln Phe Ala Arg Asn Trp Trp Asn Asp Gln Pro
        35                  40                  45

Pro Ala Thr Met Ser Pro Asp Ser Arg Arg Pro Glu Ile Pro Ala Asp
    50                  55                  60

Asp Gly Ala Val Ile Asp Ala Tyr Phe His Ala Met Ala Gly Gln Glu
65                  70                  75                  80

Ser Tyr Thr Ile Thr Ser Thr Pro Leu Ala Arg Ala Thr Val Asp Gly
                85                  90                  95
```

```
Leu Met Phe Ser Ala Ala Gly Ala Ala Ala Pro Pro Ala Ala Ser Thr
            100                 105                 110

Ile Asp Ala Val Arg Ala Glu Tyr Thr Ala Gln Lys Ala Arg Gln Asp
        115                 120                 125

Gln Ala Arg Glu Arg Ala Arg Tyr Leu Ala Val Leu Leu Leu Val Asp
    130                 135                 140

Val Leu Ala Arg Arg Leu Pro Asp Val Lys Gln Leu Leu Val Glu Gly
145                 150                 155                 160

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

Thr Pro Asp Thr Glu Ile Ser Gln Ile Asn Ala Asp Ile Gly Ala Leu
            180                 185                 190

Pro Thr Asp Val Gly Tyr Lys Val Trp Gly His Ile Cys Thr Arg Ala
        195                 200                 205

Ser Glu Asp Arg Asn Gln Leu Val Val Asp Val Ala Glu Leu Gln Lys
    210                 215                 220

Tyr Ala Ala Asn Gly Gly Asp Arg Arg Lys
225                 230
```

<210> 182
<211> 99
<212> PRT
<213> Micrococcus sp.

<400> 182

```
Met Thr Thr Phe Tyr Tyr Ser Val Asn Glu Met Ala Arg Gln Ala Gly
1               5                   10                  15

Phe Thr Ser Pro Ser Thr Trp Thr Arg Ile Lys Asn Asp Leu Arg Ala
            20                  25                  30

Pro Asp Ala Val Met Val Ser Lys Ile Gln Gly Ala Gly Trp Thr Leu
        35                  40                  45
```

Asp Ser Trp Glu Asp Ala Ala Asn Asn Glu Leu His Thr Arg Arg Glu
    50   55   60

Pro Leu Met Lys Ala Ile Thr Arg Leu Arg Leu Glu Glu Glu Ala Leu
65   70   75   80

Arg Gln Glu Tyr Gly Val Thr Asp Gln Asp Glu Pro Ala Asp Gln Lys
                85   90   95

Ala Glu Gly

<210> 183
<211> 149
<212> PRT
<213> Micrococcus sp.

<400> 183
Met Asn Ala Ile Glu Leu Lys Cys Arg Arg Glu Ala Leu Gly Leu Ser
1   5   10   15

Arg Asp Ala Leu Ala Asp Thr Leu Gly Val Ala Glu Lys Ser Ile Thr
            20   25   30

Arg Trp Glu Phe Gly Lys Asn Pro Pro Arg Asp Trp Ser Trp Ile Asp
        35   40   45

Ser Ala Met Thr Arg Leu Glu Asp Tyr Arg Glu Arg Leu Val Glu Glu
    50   55   60

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

Met Leu Thr Tyr Ala Ser Arg Gly Ser Phe Tyr Arg Trp Leu Pro Glu
                85   90   95

Met Glu Glu Pro Asp Met Gly Asp Gly Ile Glu Val Ile Pro Val Glu
            100   105   110

Leu His Arg Ala Ala Thr Ala Glu Ala Ala Arg Arg Leu Arg Val Ser
        115   120   125

His Gly Leu His Ala Val Ile Glu Ala Val Pro Thr Pro Gln Glu Gly
130 135 140

Glu Gly Ala Gly Ser
145

<210> 184
<211> 245
<212> PRT
<213> Micrococcus sp.

<400> 184
Met Ser Val Arg Arg Trp Val Ser Ser Val Val Ala Val Val Ala Ala
1 5 10 15

Leu Gly Val Val Gly Ala Val Gly Gly Gly His Val Gln Gly Val Gln
20 25 30

Val Pro Glu Val Lys Leu Pro Lys Val Ala Leu Pro Ser Gly Val Glu
35 40 45

Leu Pro Ser Leu Gly Asp Thr Ser Pro Thr Pro Ala Ala Gly Pro Val
50 55 60

Gly Gly Gln Leu Glu Gln Leu Ala Leu Thr Ser Ala Ala Pro Ala Ser
65 70 75 80

Ala Tyr Thr Arg Asp Ala Phe Gly Gln Arg Trp Ala Asp Val Asp Arg
85 90 95

Asn Gly Cys Asp Thr Arg Asn Asp Val Leu Arg Arg Asp Leu Thr Gln
100 105 110

Val Gln Ile Lys Pro Gly Thr Gln Gly Cys Lys Val Leu Ser Gly Gln
115 120 125

Leu Val Asp Pro Tyr Ser Gly Ala Thr Ile Pro Phe Ser Ser Gln Asp
130 135 140

Ser Gln Ala Val His Ile Asp His Thr Val Ser Leu Ala Asp Ala Trp
145 150 155 160

Ala Ser Gly Ala Trp Ala Trp Asp Glu Ser Gln Arg Thr Ala Phe Ala
165 170 175

Asn Asp Pro Ala Asn Leu Leu Ala Val Asp Gly Pro Ala Asn Thr Ser
180 185 190

Lys Ser Asp Ala Thr Ala Ala Asp Trp Leu Pro Asp Thr Val Ala Gly
195 200 205

Arg Cys Glu Leu Val Glu His Gln Val Val Val Lys Ala Lys Trp Gly
210 215 220

Leu Ser Val Thr Glu Arg Glu Arg Ala Ala Met Arg Arg Val Leu Ala
225 230 235 240

Ser Cys Pro Ser Gly
245

<210> 185
<211> 99
<212> PRT
<213> Micrococcus sp.

<400> 185
Met Thr Thr Phe Tyr Tyr Ser Val Asn Glu Met Ala Arg Gln Ala Gly
1 5 10 15

Phe Thr Ser Pro Ser Thr Trp Thr Arg Ile Lys Asn Asp Leu Arg Ala
20 25 30

Pro Asp Ala Val Met Val Ser Lys Ile Gln Gly Ala Gly Trp Thr Leu
35 40 45

Asp Ser Trp Glu Asp Ala Ala Asn Asn Glu Leu His Thr Arg Arg Glu
50 55 60

Pro Leu Met Lys Ala Ile Thr Arg Leu Arg Leu Glu Glu Glu Ala Leu
65 70 75 80

Arg Gln Glu Tyr Gly Val Thr Asp Gln Asp Glu Pro Ala Asp Gln Lys
85 90 95

Ala Glu Gly

<210> 186
<211> 123
<212> PRT
<213> Micrococcus sp.

<220>
<221> MOD_RES
<222> (107)..(107)
<223> Any amino acid

<400> 186
Met Asn Ala Ile Glu Leu Lys Cys Arg Arg Glu Ala Leu Gly Leu Ser
1 5 10 15

Arg Asp Ala Leu Ala Asp Thr Leu Gly Val Ala Glu Lys Ser Ile Thr
20 25 30

Arg Trp Glu Phe Gly Lys Asn Pro Pro Arg Asp Trp Ser Trp Ile Asp
35 40 45

Ser Ala Met Thr Arg Leu Glu Asp Tyr Arg Glu Arg Leu Val Glu Glu
50 55 60

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

Met Leu Thr Tyr Ala Ser Arg Gly Ser Phe Tyr Arg Trp Leu Pro Glu
85 90 95

Met Glu Glu Pro Asp Met Gly Asp Gly Ile Xaa Val Ile Glu Ala Val
100 105 110

Pro Thr Pro Gln Glu Arg Glu Gly Ala Gly Ser
115 120

<210> 187
<211> 245
<212> PRT
<213> Micrococcus sp.

<400> 187
Met Ser Val Arg Arg Trp Val Ser Ser Val Val Ala Val Val Ala Ala
1   5   10   15

Leu Gly Val Ala Gly Ala Val Gly Gly Gly His Val Gln Gly Val Gln
            20   25   30

Val Pro Glu Val Lys Leu Pro Lys Val Ala Leu Pro Ser Gly Val Glu
        35   40   45

Leu Pro Ser Leu Gly Asp Thr Ser Pro Thr Pro Ala Ala Gly Pro Val
    50   55   60

Gly Gly Gln Leu Glu Gln Leu Ala Leu Thr Ser Ala Ala Pro Ala Ser
65   70   75   80

Ala Tyr Thr Arg Asp Ala Phe Gly Gln Arg Trp Ala Asp Val Asp Arg
                85   90   95

Asn Gly Cys Asp Thr Arg Asn Asp Val Leu Arg Arg Asp Leu Thr Gln
            100   105   110

Val Gln Ile Lys Pro Gly Thr Gln Gly Cys Lys Val Leu Ser Gly Gln
        115   120   125

Leu Val Asp Pro Tyr Ser Gly Ala Thr Ile Pro Phe Ser Ser Gln Asp
    130   135   140

Ser Gln Ala Val Gln Ile Asp His Thr Val Ser Leu Ala Asp Ala Trp
145   150   155   160

Ala Ser Gly Ala Trp Ala Trp Asp Glu Ser Gln Arg Thr Ala Phe Ala
                165   170   175

Asn Asp Pro Ala Asn Leu Leu Ala Val Asp Gly Pro Ala Asn Asn Ser
            180   185   190

Lys Ser Asp Ala Thr Ala Ala Asp Trp Leu Pro Asp Thr Val Ala Gly
        195   200   205

Arg Cys Glu Leu Val Glu His Gln Val Val Val Lys Ala Lys Trp Gly

```
    210                 215                 220

Leu Ser Val Thr Glu Arg Glu Arg Ala Ala Met Arg Arg Val Leu Ala
225                 230                 235                 240

Ser Cys Pro Ala Gly
                245
```

<210> 188
<211> 99
<212> PRT
<213> Micrococcus sp.

<400> 188

```
Met Thr Thr Phe Tyr Tyr Ser Val Asn Glu Met Ala Arg Gln Ala Gly
1               5                   10                  15

Phe Thr Ser Pro Ser Thr Trp Thr Arg Ile Lys Asn Asp Leu Arg Ala
            20                  25                  30

Pro Asp Ala Val Met Val Ser Lys Ile Gln Gly Ala Gly Trp Thr Leu
        35                  40                  45

Asp Ser Trp Glu Asp Ala Ala Asn Asn Glu Leu His Thr Arg Arg Glu
    50                  55                  60

Pro Leu Met Lys Ala Ile Thr Arg Leu Arg Leu Glu Glu Glu Ala Leu
65                  70                  75                  80

Arg Gln Glu Tyr Gly Val Thr Asp Gln Asp Glu Pro Ala Asp Gln Lys
                85                  90                  95

Ala Glu Gly
```

<210> 189
<211> 149
<212> PRT
<213> Micrococcus sp.

<400> 189

```
Met Asn Ala Ile Glu Leu Lys Cys Arg Arg Glu Ala Leu Gly Leu Ser
1               5                   10                  15
```

Arg Asp Ala Leu Ala Asp Thr Leu Gly Val Ala Glu Lys Ser Ile Thr
            20   25   30

Arg Trp Glu Phe Gly Lys Asn Pro Pro Arg Asp Trp Ser Trp Ile Asp
        35   40   45

Ser Ala Met Thr Arg Leu Glu Asp Tyr Arg Glu Arg Leu Val Glu Glu
    50   55   60

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

Met Leu Thr Tyr Ala Ser Arg Gly Ser Phe Tyr Arg Trp Leu Pro Glu
                85   90   95

Met Glu Glu Pro Asp Met Gly Asp Gly Ile Glu Val Ile Pro Val Glu
            100   105   110

Leu His Arg Ala Ala Thr Ala Glu Ala Ala Arg Arg Leu Arg Val Ser
        115   120   125

His Gly Leu His Ala Val Ile Glu Ala Val Pro Thr Pro Gln Glu Arg
    130   135   140

Glu Gly Ala Gly Ser
145

<210> 190
<211> 245
<212> PRT
<213> Micrococcus sp.

<400> 190
Met Ser Val Arg Arg Trp Val Ser Gly Val Val Ala Val Val Ala Ala
1   5   10   15

Leu Gly Val Ala Gly Ala Val Gly Gly Gly His Val Gln Gly Val Gln
            20   25   30

Val Pro Glu Val Lys Leu Pro Lys Val Ala Leu Pro Ser Gly Val Glu
        35   40   45

```
Leu Pro Ser Leu Gly Asp Thr Ser Pro Thr Pro Ala Ala Gly Pro Val
    50                  55                  60

Gly Gly Gln Leu Glu Gln Leu Ala Leu Thr Ser Ala Ala Pro Ala Ser
65                  70                  75                  80

Ala Tyr Thr Arg Asp Ala Phe Gly Gln Arg Trp Ala Asp Val Asp Arg
                85                  90                  95

Asn Gly Cys Asp Thr Arg Asn Asp Val Leu Arg Arg Asp Leu Thr Gln
            100                 105                 110

Val Gln Ile Lys Pro Gly Thr Gln Gly Cys Lys Val Leu Ser Gly Gln
        115                 120                 125

Leu Val Asp Pro Tyr Ser Gly Ala Thr Ile Pro Phe Ser Ser Gln Asp
    130                 135                 140

Ser Gln Ala Val His Ile Asp His Thr Val Ser Leu Ala Asp Ala Trp
145                 150                 155                 160

Ala Ser Gly Ala Trp Ala Trp Asp Glu Ser Gln Arg Thr Ala Phe Ala
                165                 170                 175

Asn Asp Pro Ala Asn Leu Leu Ala Val Asp Gly Pro Ala Asn Thr Ser
            180                 185                 190

Lys Ser Asp Ala Thr Ala Ala Asp Trp Leu Pro Gly Thr Val Ala Gly
        195                 200                 205

Arg Cys Glu Leu Val Glu His Gln Val Val Val Lys Ala Lys Trp Gly
    210                 215                 220

Leu Ser Val Thr Glu Arg Glu Arg Ala Ala Met Arg Arg Val Leu Ala
225                 230                 235                 240

Ser Cys Pro Ala Gly
                245
```

<210> 191

<211> 206
<212> PRT
<213> Paeniglutamicibacter sp.

<400> 191
Met Leu Gly Gln Gly Phe Asp Gln Gly Leu Glu Phe Gly Phe Val Leu
1   5   10   15

Gly Lys Gly Leu Val Gln Gln Leu Leu Ala Cys Pro Val Gln Gly His
            20   25   30

Arg Val Val Val Gly Leu Ala Asp Ile Asp Ala Asp Glu His Val Asp
        35   40   45

Gly Phe Val Val Leu Asp His Cys Ala Pro His Ser Trp Gly His Arg
    50   55   60

Arg Val Ala Ser Asp Glu Thr Val Leu Ala Cys Pro Ala Ala Pro Arg
65   70   75   80

Leu Gly Ile His Val Thr Thr Asp Asn Trp Thr Gly Ala Val Pro Ser
                85   90   95

Arg Ser Gly Pro Tyr Leu Arg Tyr Pro Arg Cys Leu Ser Val Pro Val
            100   105   110

Thr Ala Pro Pro Gly Ser Trp Met Thr Gly Gly Val Asn His Ala Gly
        115   120   125

Thr Asp Arg Pro Ala Ile Pro Ile Leu Ala His Gly Ala Arg Leu Arg
    130   135   140

Asp Tyr Lys Lys Gly Asn Gly Gly Gly Ala Asp Ala Pro Leu Asn Ala
145   150   155   160

Ala Ser Cys Gly Gly Gln Trp Arg Met Ala Cys Asp Arg Thr Ser Val
                165   170   175

Ala Cys Gly Arg Gln Arg Pro Leu Ser Ser Gly Ser Arg Arg Leu Pro
            180   185   190

Arg His Trp Ser Leu Pro His Pro Gly Cys Thr Phe Ile Pro

195 200 205

<210> 192
<211> 106
<212> PRT
<213> Streptomyces sp.

<220>
<221> MOD_RES
<222> (65)..(66)
<223> Any amino acid

<400> 192
Met Gly Ser Ser Ala Val Arg Ala Leu Arg Val Leu Pro Pro Gly Ala
1 5 10 15

Val Asp Thr Thr Pro Pro Pro Ala Ala Asn Arg Pro Ala Ser Gly Pro
20 25 30

Ala Ala Ala Pro Ala Gly Pro Val Lys Thr Arg Glu Thr Ala Ser Ala
35 40 45

Gly Tyr Arg Leu Ala Leu Asp Ala His Arg Ala Ala Val Asp Pro Ala
50 55 60

Xaa Xaa Ser Arg Ile Pro Ala Asp Ala Gln Ala Ala Phe Ala Asp Tyr
65 70 75 80

Arg Ala Gln Ala Arg Glu Thr Ala Ser Ala Val Gln Asp Ala Ala Ile
85 90 95

Arg Arg Ala Arg Val Glu Arg Thr Ala Ala
100 105

<210> 193
<211> 244
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 193

gggaaacagc uagaauguaa cuuaaaguag gucaauguuu aaauucgaug uugcaauuug 60

uuuggacaag uggauuaaaa cguuccuuga aaaucauaua aagcagccag uuuacgggcu 120

ugggcgaauu ugcguccaaa gggugaggcc agguguaagu aagaaccuac aaaagcacuc 180

accaaagggu caaaaagauc cuacauggau gugaggauug aaauguacga gauauaggaa 240

gcug 244

<210> 194
<211> 242
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic polynucleotide"

<400> 194
gggaaacagc uagaauguaa cuuaaaguag gucaauguuu aaauucgaug uugcaauuug 60

uuuggacaag uggauuaaaa cguuccuuga aaaucauaua aagcagccag uuuacgggcu 120

ugggcgaauu ugcguccaaa gggugaggcc agguguaagu aagaaccuac aaaagcacuc 180

accaaagggu caaaaagauc cuacauggau gugaggauug aaauacuaca acagccacaa 240

cg 242

<210> 195
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 195
aggagcggac agcagcttcc tatatctcgt acaggaccgt caacatggtt 50

<210> 196
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 196
agccatgata tagacgttgt ggctgttgta gtaggaccgt caacatggtt 50

<210> 197
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 197
agcggtggga aaggaatccc 20

<210> 198
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 198
ttcgattccg gagagggagc ct 22

<210> 199
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 199
aggagcggac agcagcttcc tatatctcgt acagggaact caacatggtt 50

<210> 200
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 200
agccatgata tagacgttgt ggctgttgta gtagggaact caacatggtt 50

<210> 201
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 201
gtcacatcct acatggatgt gaggattgaa at 32

<210> 202
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 202
gtcacatcct acatggatgt gaggattgaa at 32

<210> 203
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 203
gtcacatcct acatgggtgt gtggattgaa at 32

<210> 204
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 204
gtcacatcct acatggatgt gaggattgaa at 32

<210> 205
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 205
gtcacaccct acatgggtgt aaggattgaa at 32

<210> 206
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 206
gttwaaacat aacaatagat gtwttgaaat 30

<210> 207
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 207
gtcacatcct acatgggtgt gtggattgaa at 32

<210> 208
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 208
gttaaaacaa aacaatagat gtattgaaat 30

<210> 209
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 209
gtttaaactt aacaatagat gtattgaaat 30

<210> 210
<211> 31
<212> DNA
<213> Bacillus sp.

<400> 210
gtttactatt aacattggat gtatttaaat t 31

<210> 211
<211> 37
<212> DNA
<213> Gordonia sp.

<400> 211
gtcgtgagta cgacgccctc acgctggcgt tgcgacc 37

<210> 212
<211> 29
<212> DNA
<213> Micrococcus sp.

<400> 212
ggaagccccg cgcacgcagg gatgagccc 29

<210> 213
<211> 29
<212> DNA
<213> Micrococcus sp.

<400> 213
ggaagccccg cgcacgcagg gatgagccc 29

<210> 214
<211> 29
<212> DNA
<213> Paeniglutamicibacter sp.

<400> 214
gtccttcccg ggcgcgcgcg gaatggctc 29

<210> 215
<211> 29
<212> DNA
<213> Streptomyces sp.

<400> 215
gtacggccca cacgcgcagg ggacgaccg 29

<210> 216
<211> 29
<212> DNA
<213> Micrococcus sp.

<400> 216
ggaagccccg cgcacgcagg gatgagccc 29

<210> 217
<211> 29
<212> DNA
<213> Bacillus sp.

<400> 217
gtcggccccg cgcacgcgga gatgagccc 29

<210> 218
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 218
gtttgacact aacataagat gtatttaaat 30

<210> 219
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 219
gtttaacact aacataagat gtatttaaat 30

<210> 220
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 220
gtttaacatt aacataagat gtatttaaat 30

<210> 221
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 221
gtcgcacctt atataggtgc gtggattgaa at 32

<210> 222
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 222
gtttastact aacataagat gtatttaaat 30

<210> 223
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 223
gtttaatact aacataagat gtatttaaat 30

<210> 224
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 224
gtcacatcct acatggatgt gaggattgaa at 32

<210> 225
<211> 22
<212> DNA
<213> Bacillus sp.

<400> 225
tactcatgaa cacctcatgw gt 22

<210> 226
<211> 33
<212> DNA
<213> Bacillus sp.

<400> 226
gtttactayt aacataagat gtatttaaat ggt 33

<210> 227
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 227
gtttastayt aacataagat gtatttaaat 30

<210> 228
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 228
gtttagtact aacataagat gtatttaaat 30

<210> 229
<211> 29
<212> DNA
<213> Bacillus sp.

<400> 229
gttttgaata aactatgtag aatgtgaat 29

<210> 230
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 230
gtttaacact aacataagat gtatttaaat 30

<210> 231
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 231
gtttaacact aacataagat gtatttaaat 30

<210> 232
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 232
gtttaacact aacataagat gtatttaaat 30

<210> 233
<211> 29
<212> DNA
<213> Bacillus sp.

<400> 233
gttttgaata aactatgtag aatgtgaat 29

<210> 234
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 234
gtttaacact aacataagat gtatttaaat 30

<210> 235
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<221> modified_base
<222> (1)..(19)

<223> a, c, t, g, unknown or other

<400> 235
nnnnnnnnnn nnnnnnnnn 19

<210> 236
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 236
gtttagtact aacataagat gtatttaaat 30

<210> 237
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 237
gtttaacact aacataagat gtatttaaat 30

<210> 238
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<221> modified_base
<222> (1)..(19)
<223> a, c, t, g, unknown or other

<400> 238
nnnnnnnnnn nnnnnnnnn 19

<210> 239
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 239
gtttartact aacataagat gtatttaaat 30

<210> 240
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 240

gtttaacact aacataagat gtatttaaat 30

<210> 241
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 241
gttaaaacat aacaatagat gtattgaaat 30

<210> 242
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 242
gttaaaacat aacaatagat gtattgaaat 30

<210> 243
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 243
gtttaacact aacataagat gtatttaaat 30

<210> 244
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 244
gtttracact aacataagat gtatttaaat 30

<210> 245
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 245
gtttaacatt aacataagat gtatttaaat 30

<210> 246
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 246

gtcacatcct acatggatgt gaggattgaa at 32

<210> 247
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 247
gtcacatcct acatggatgt gtggattgaa at 32

<210> 248
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 248
gtttagtact aacataagat gtatttaaat 30

<210> 249
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 249
gtttaacatt aacataagat gtatttaaat 30

<210> 250
<211> 33
<212> DNA
<213> Bacillus sp.

<400> 250
gtcacatcct acatggatgt gaggattgaa att 33

<210> 251
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<221> modified_base
<222> (1)..(19)
<223> a, c, t, g, unknown or other

<400> 251
nnnnnnnnnn nnnnnnnnn 19

<210> 252
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 252
gtttaacatt aacataagat gtatttaaat 30

<210> 253
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 253
gtttaacatt aacataagat gtatttaaat 30

<210> 254
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 254
gtttastayt aacataagat gtatttaaat 30

<210> 255
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 255
gtttaacact aacataagat gtatttaaat 30

<210> 256
<211> 30
<212> DNA
<213> Bacillus sp.

<220>
<221> modified_base
<222> (7)..(7)
<223> a, c, t, g, unknown or other

<400> 256
gkttganact aayataagat gtatttaaat 30

<210> 257
<211> 30
<212> DNA

<213> Bacillus sp.

<400> 257
gtttaacact aacataagat gtatttaaat 30

<210> 258
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 258
gtttactayt aacataagat gtatttaaat 30

<210> 259
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 259
gtttastayt aacataagat gtatttaaat 30

<210> 260
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 260
gtttaacact aacataagat gtatttaaat 30

<210> 261
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 261
gtttaatact aacataagat gtatttaaat 30

<210> 262
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 262
gtttaacact aacataagat gtatttaaat 30

<210> 263
<211> 30
<212> DNA

<213> Bacillus sp.

<400> 263
gtttartact aacataagat gtatttaaat 30

<210> 264
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 264
gtttagtayt aacataagat gtatttaaat 30

<210> 265
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 265
gtttaacact aacataagat gtatttaaat 30

<210> 266
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 266
gtttaatact aacataagat gtatttaaat 30

<210> 267
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 267
gtttascatt aacataagat gtatttaaat 30

<210> 268
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 268
gtttgacact aacataagat gtatttaaat 30

<210> 269
<211> 29
<212> DNA

<213> Bacillus sp.

<400> 269
rtttgatatc aactatgtgg aatgtaaat 29

<210> 270
<211> 29
<212> DNA
<213> Bacillus sp.

<400> 270
rtttgatatc aactatgtgg aatgtaaat 29

<210> 271
<211> 29
<212> DNA
<213> Bacillus sp.

<400> 271
rtttgatatc aactatgtgg aatgtaaat 29

<210> 272
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 272
gtttaacact aacataagat gtatttaaat 30

<210> 273
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 273
gtttartact aacataagat gtatttaaat 30

<210> 274
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 274
gtttagtact aacataagat gtatttaaat 30

<210> 275
<211> 29
<212> DNA

<213> Bacillus sp.

<400> 275
gttttgaata aactatgtag aatgtgaat 29

<210> 276
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 276
gtttaacact aacataagat gtatttaaat 30

<210> 277
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 277
gtttaacact aacataagat gtatttaaat 30

<210> 278
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 278
gtttaacact aacataagat gtatttaaat 30

<210> 279
<211> 34
<212> DNA
<213> Bacillus sp.

<400> 279
agtaacaata cgtatagcta cgttaattgt agga 34

<210> 280
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 280
gtttaacact aacataagat gtatttaaat 30

<210> 281
<211> 30
<212> DNA

<213> Bacillus sp.

<400> 281
gtttaacact aacataagat gtatttaaat 30

<210> 282
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 282
gtttagtact aacataagat gtatttaaat 30

<210> 283
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 283
gtttastact aacataagat gtatttaaat 30

<210> 284
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 284
gtttagtact aacataagat gtatttaaat 30

<210> 285
<211> 29
<212> DNA
<213> Bacillus sp.

<400> 285
gttttaatca caatatgtat tgttagaat 29

<210> 286
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 286
gtttagyact aacataagat gtatttaaat 30

<210> 287
<211> 30
<212> DNA

<213> Bacillus sp.

<400> 287
gtttastact aacataagat gtatttaaat 30

<210> 288
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 288
gtttaaacat aacaatagat gtattgaaat 30

<210> 289
<211> 29
<212> DNA
<213> Bacillus sp.

<400> 289
rttttaaata caatatgtat tgttagaat 29

<210> 290
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 290
gtttastayt aacataagat gtatttaaat 30

<210> 291
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 291
gtttastayt aacataagat gtatttaaat 30

<210> 292
<211> 29
<212> DNA
<213> Bacillus sp.

<400> 292
gtttgatatc aactatatgg aatgtaaat 29

<210> 293
<211> 30
<212> DNA

<213> Bacillus sp.

<400> 293
gtttastayt aacataagat gtatttaaat 30

<210> 294
<211> 32
<212> DNA
<213> Bacillus sp.

<400> 294
gtcacatcct acatgggtgt gaggattgaa at 32

<210> 295
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 295
gtttaahatt aacataagat gtatttaaat 30

<210> 296
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 296
gtttastayt aacataagat gtatttaaat 30

<210> 297
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 297
gtttastayt aacataagat gtatttaaat 30

<210> 298
<211> 21
<212> DNA
<213> Bacillus sp.

<400> 298
ggtcatgttg tcattacatt g 21

<210> 299
<211> 18
<212> DNA

<213> Bacillus sp.

<400> 299
atcccaaaaa cgttaggt 18

<210> 300
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 300
gtttaacact aacataagat gtatttaaat 30

<210> 301
<211> 29
<212> DNA
<213> Bacillus sp.

<400> 301
gtcttaatta caatatgtat tgttagaat 29

<210> 302
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 302
gtttaacact aacataagat gtatttaaat 30

<210> 303
<211> 19
<212> DNA
<213> Bacillus sp.

<220>
<221> modified_base
<222> (5)..(5)
<223> a, c, t, g, unknown or other

<400> 303
gatanaaagt gtttcatat 19

<210> 304
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<221> modified_base
<222> (1)..(19)
<223> a, c, t, g, unknown or other

<400> 304
nnnnnnnnnn nnnnnnnnn 19

<210> 305
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 305
gtttactrtt aacawtggat gtatttaaat 30

<210> 306
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 306
gttaaaacaa aacaatagat gtattgaaat 30

<210> 307
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 307
gttaaaacaw aacaatagat gtattgaaat 30

<210> 308
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 308
gttwaaacat aacaatagat gtwttgaaat 30

<210> 309
<211> 30
<212> DNA
<213> Bacillus sp.

<400> 309
gttwaaacat aacaatagat gtwttgaaat 30

<210> 310
<211> 146
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<400> 310
acggucuuug uguuacaccc uucaaagcau ccaggaugcu uauguaucaa uguuuuuuua 60

acauucgaua cuaccuuaug gagcguuuac acuugguaaa uucaccuugu auuacuucuc 120

cauuaaagca auagauguau ugaaau 146

<210> 311
<211> 143
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220>
<221> modified_base
<222> (52)..(59)
<223> a, c, t, g, unknown or other

<400> 311
tcgtcggcag cgtcagatgt gtataagaga cagaacgacg gccagtgaat tnnnnnnnna 60

ctacaacagc cacaacgtct atatcatggg atcctctaga gtcgacctgc tgtctcttat 120

acacatctcc gagcccacga gac 143

<210> 312
<211> 148
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
polynucleotide"

<220>

<221> modified_base
<222> (52)..(56)
<223> a, c, t, g, unknown or other

<400> 312
tcgtcggcag cgtcagatgt gtataagaga cagaacgacg gccagtgaat tnnnnntgga 60

atgggaacta aagtaatggc aaacgaattc gtaggatcct ctagagtcga cctgctgtct 120

cttatacaca tctccgagcc cacgagac 148

<210> 313
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 313
actacaacag ccacaacgtc tatatcatgg 30

<210> 314
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 314
acacctgtaa tcccagca 18