



US007846686B2

(12) **United States Patent**  
**Kramer**

(10) **Patent No.:** **US 7,846,686 B2**

(45) **Date of Patent:** **Dec. 7, 2010**

(54) **MICROGININ PRODUCING PROTEINS AND NUCLEIC ACIDS ENCODING A MICROGININ GENE CLUSTER AS WELL AS METHODS FOR CREATING NOVEL MICROGININS**

(65) **Prior Publication Data**

US 2009/0220962 A1 Sep. 3, 2009

(30) **Foreign Application Priority Data**

Dec. 2, 2005 (EP) ..... 05026396

(75) Inventor: **Dan Kramer**, Berlin (DE)

(51) **Int. Cl.**

**C12P 1/00** (2006.01)

**C12N 1/20** (2006.01)

**C12N 16/00** (2006.01)

**C07H 21/04** (2006.01)

(73) Assignees: **Cyano Biotech GmbH**, Berlin (DE);  
**Humboldt-Universitaet zu Berlin**,  
Berlin (DE)

(\*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 130 days.

(52) **U.S. Cl.** ..... **435/41**; 435/69.1; 435/252.3; 435/252.35; 435/320.1; 435/183; 536/23.2

(58) **Field of Classification Search** ..... None  
See application file for complete search history.

(21) Appl. No.: **12/095,787**

*Primary Examiner*—Nashaat T Nashed

(74) *Attorney, Agent, or Firm*—Rothwell, Figg, Ernst & manbeck, P.C.

(22) PCT Filed: **Dec. 1, 2006**

(86) PCT No.: **PCT/EP2006/011563**

(57) **ABSTRACT**

The invention provides for nucleic acid molecules enabling the synthesis of microginin and microginin analogues. The invention also provides for methods for identifying microginins as well creating microginins which may not be found in nature.

§ 371 (c)(1),

(2), (4) Date: **Jun. 26, 2008**

(87) PCT Pub. No.: **WO2007/062867**

PCT Pub. Date: **Jun. 7, 2007**

**11 Claims, 2 Drawing Sheets**

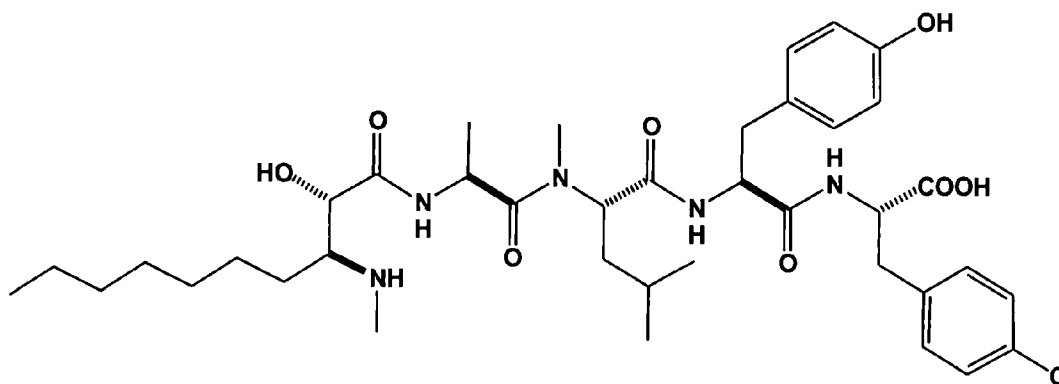


Fig. 1:

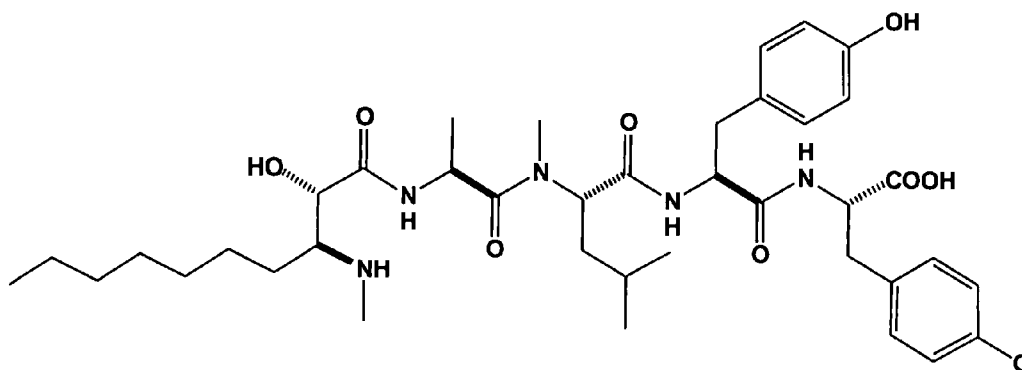
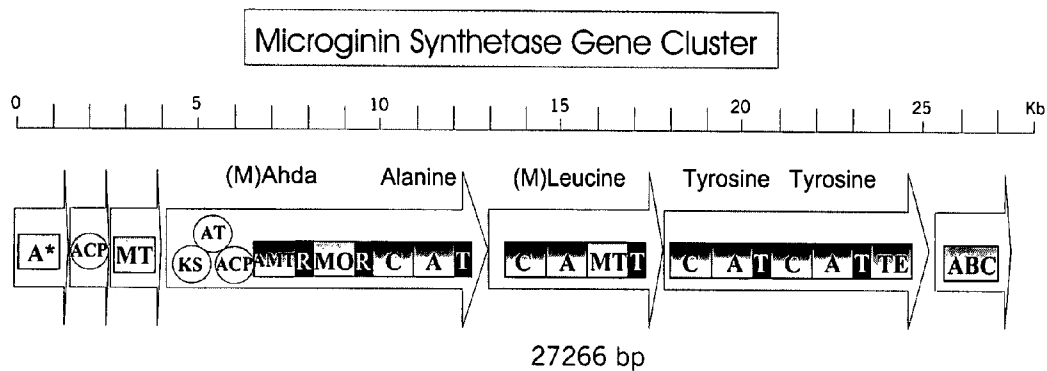
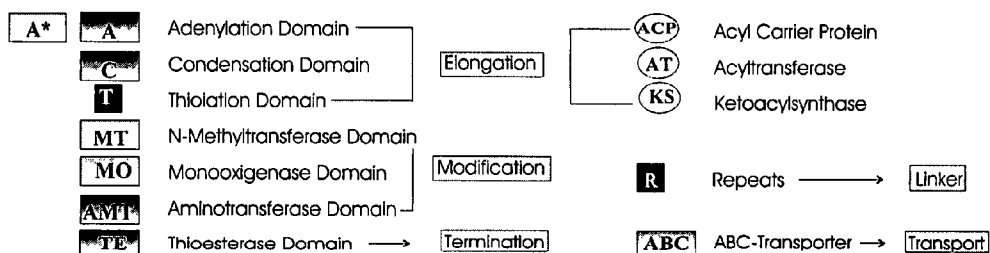


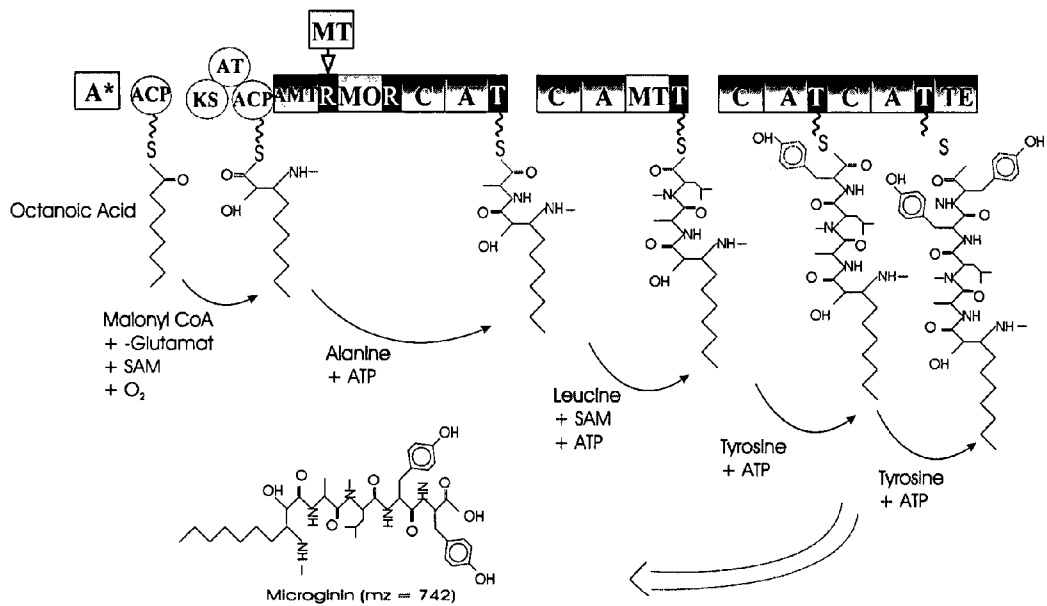
Fig. 2



Modular Organisation of the Biosynthetic Genes and Proteins resp.



Biosynthetic Pathway of Microginin



**MICROGININ PRODUCING PROTEINS AND  
NUCLEIC ACIDS ENCODING A MICROGININ  
GENE CLUSTER AS WELL AS METHODS  
FOR CREATING NOVEL MICROGININS**

CROSS REFERENCE TO RELATED  
APPLICATION

This application is a 35 U.S.C. 371 National Phase Entry Application from PCT/EP2006/011563, filed Dec. 1, 2006, which claims the benefit of European Patent Application No. 05026396.1 filed on Dec. 2, 2005, the disclosure of which is incorporated herein in its entirety by reference.

TECHNICAL FIELD

The present invention relates to the fields of chemistry, biology, biochemistry, molecular biology. The invention provides for novel nucleic acid molecules enabling the synthesis of microginin and microginin analogues. Microginin finds an application in therapeutics. The invention thus extends into the field of mammalian therapeutics and drug development.

INTRODUCTION

Cyanobacteria and Microginin

Cyanobacteria are gram-negative bacteria. Due to their ability to perform photosynthesis they were long thought to belong to the plant kingdom and were formerly classified as blue-green algae. Cyanobacteria have adapted to almost all ecological niches. Most of strains known up to date are found in fresh water lakes and oceans. In the last few years cyanobacteria have been recognised as a source for biologically active natural compounds.

Cyanobacteria are a group of microscopic organisms somewhere "in between" algae and bacteria and they are found in freshwater and marine areas throughout the world. Scientifically, they are considered to be bacteria, but because they can perform photosynthesis, they also used to be classified as "blue-green algae".

Cyanobacterial peptides (cyanopeptides) are among the most ubiquitously found potentially hazardous natural products in surface waters used by humans. Though these substances are natural in origin, eutrophication (i.e. excessive loading with fertilising nutrients) has caused massive cyanobacterial proliferation throughout Europe. Thus, cyanopeptides now occur with unnatural frequency and concentration.

A large group among the diverse cyanopeptides are the oligopeptides (peptides with a molecular weight of <2 KD). But while specific cyanopeptides—e.g. microcystins and nodularins—are well studied and recognised as being causative for many animal poisonings and human illness, a substantial and increasing body of evidence points toward a decisive role of other potentially toxic cyanopeptides in the causation of both acute and chronic human illnesses.

Freshwater and marine cyanobacteria are known to produce a variety of bioactive compounds, among them potent hepatotoxins and neurotoxins. Many of the toxic species of cyanobacteria tend to massive proliferation in eutrophicated water bodies and thus have been the cause for considerable hazards for animal and human health. One of the most widespread bloom-forming cyanobacteria is the genus *Microcystis*, a well-known producer of the hepatotoxic peptide microcystin. Microcystins are a group of closely related cyclic heptapeptides sharing the common structure. So far, more than 80 derivatives of microcystins have been identified, varying largely by the degree of methylation, peptide sequence, and toxicity.

The traditional botanical code describes the genus *Microcystis* as a cocal, unicellular cyanobacterium that grows as mucilaginous colonies of irregularly arranged cells (under natural conditions, while strain cultures usually grow as single cells). According to this tradition, morphological criteria such as size of the individual cells, colony morphology, and mucilage characteristics are used for species delimitation within *Microcystis* (i.e., morphospecies). Microcystin-producing strains as well as strains that do not synthesize microcystin have been reported for all species within the genus *Microcystis*. However, whereas most field samples and strains of *Microcystis aeruginosa* and *Microcystis viridis* studied to date were found to contain microcystins, strains of *M. wesenbergii*, *M. novaceckii*, and *M. ichthyoblabe* have only sporadically been reported to contain microcystins.

Beside microcystins, various other linear and cyclic oligopeptides such as anabaenopeptins, aeruginosins, microginins and cyanopeptolins are found within the genus *Microcystis* (Namikoshi, M., and K. L. Rinehart. 1996. Bioactive compounds produced by cyanobacteria. J. Ind. Microbiol. 17:373-384.).

Similar to microcystins, these peptides possess unusual amino acids like 3-amino-6-hydroxy-2-piperidone (Ahp) in cyanopeptolins, 2-carboxy-6-hydroxyoctahydroindol (Choi) in aeruginosin-type molecules or 3-amino-2 hydroxy-decanoic acid (Ahda) in microginins and numerous structural variants also exist within these groups. These peptides show diverse bioactivities, frequently protease inhibition (Namikoshi, M., and K. L. Rinehart. 1996. Bioactive compounds produced by cyanobacteria. J. Ind. Microbiol. 17:373-384).

The occurrence of both microcystins and other oligopeptides such as anabaenopeptins, microginins and cyanopeptolins in natural *Microcystis* populations was recently demonstrated. It is well known that the species and genotype composition in natural *Microcystis* populations is heterogeneous, and both microcystin- and non-microcystin-containing strains have been isolated from the same sample. Just as strains producing microginin and strains not producing microginin have been found. These results suggest a considerable diversity of genotypes with different oligopeptide patterns in natural *Microcystis* populations.

By typing single *Microcystis* colonies, it was possible in 1999 to show for the first time that the actual peptide diversity in a natural population of this genus is extremely high. Many of the substances detected belong to well-known groups of cyanobacterial peptides like microcystins, anabaenopeptins, microginins, cyanopeptolins, and aeruginosins, of which many have been discovered in *Microcystis* spp. In addition, numerous unknown components have been detected in such colonies. However, the origin of these unknown components has yet to be investigated, since besides the observed epiphytic cyanobacteria and algae, heterotrophic bacteria are also known to be present in *Microcystis* colonies. Chemical screening of cyanobacterial samples (both from field samples and from culture strains) has demonstrated a wide variety of substances: e.g. an almost monospecific bloom of *Planktothrix agardhii* contained as many as 255 different substances, most of which were oligopeptides.

Thus, it may be concluded, that the situation with respect to the assignment of the capability of microginin production to certain species and strains, i.e. also a true understanding of the genotypes and species involved as well as their evolution has to date, not been possible. In fact PEPCY a research project supported by the European Commission concluded that present information shows that one species or "morphotype" (i.e. individuals with the same morphological characteristics) may comprise a range of genotypes that encode for different

“chemotypes” (i.e. morphologically indistinguishable individuals containing different cyanopeptides).

#### ACE Inhibitors and Microginin

ACE catalyses the conversion of angiotensin I into angiotensin II within the mammalian renin-angiotensin system, leading to arterial stenosis, which in turn causes an increase of blood pressure. ACE inhibitors counteract this process and therefore play a role in human medicine as blood pressure lowering agents. Microginin is an important drug candidate for ACE inhibition. So far only 30 structural variants of microginin are known, making clinical development difficult.

Microginins are characterized by a decanoic acid derivate, 3-amino-2-hydroxy-decanoic acid (Ahda) at the N-terminus and a predominance of two tyrosine units at the C-terminus. They vary in length from 4 to 6 amino acids with the variability occurring at the C-terminal end (Microginins, zinc metalloprotease inhibitors from the cyanobacterium *Microcystis aeruginosa*, 2000, Tetrahedron 56:8643-8656). In the past it has only been possible by means of synthesis of 3-amino-2-hydroxy-decanoic acid to chemically generate microginin variants (J Org. Chem. 1999 Apr. 16; 64(8):2852-2859. Acylnitrene Route to Vicinal Amino Alcohols. Application to the Synthesis of (-)-Bestatin and Analogues. Bergmeier S C, Stanchina D M.) Alternatively cyanobacterial strains were screened for microginin activity, which was tedious and time consuming. It has so far not been possible to screen for strains efficiently due to the lack of species understanding and a methodology of efficiently distinguishing microginin producers from non-producers (see above). Further it was not possible to easily and efficiently alter and thus develop microginins in order to provide for a variety of lead compounds from which better ACE-inhibitors may be developed.

#### BRIEF DESCRIPTION OF THE INVENTION

From *Microcystis aeruginosa* a cluster of genes, spanning about 30 kbps has been isolated encoding a hybrid synthetase composed of non-ribosomal peptide synthetases (NRPS), polyketide synthases (PKS) and tailoring enzyme which as the inventors show is responsible for the biosynthesis of microginin. The strain from which this nucleic acid was first isolated by G. C. Kürzinger from Lake Pehlitz 1977].

The inventors provide for a biological system enabling not only the production of microginins, the heterologous expression of microginin, but also a system for modifying microginin and thus developing so far unknown variants of microginin. The invention further provides for nucleic acids and methods for identifying strains which have the ability to produce microginin.

In particular the invention relates to one or more nucleic acids encoding a microginin synthetase enzyme complex with the following activities: an adenylation domain (A\*) wherein, the adenylation domain comprises a peptide sequence according to SEQ ID NO. 1, an acyl carrier protein (ACP), an elongation module (EM) of polyketide synthases (PKS) comprising the following activities: (i) ketoacylsynthase (KS), (ii) acyl transferase (AT) (iii) acyl carrier protein (ACP2), an aminotransferase (AMT), three to five elongation modules (EM) of non-ribosomal peptide synthetases (NRPS) comprising the following activities: (i) condensation domain (C), (ii) adenylation domain (A), (iii) thiolation domain (T) and a thioesterase (TE).

#### DETAILED DESCRIPTION OF THE INVENTION

As outlined above the invention in particular relates to one or more nucleic acids encoding a microginin synthetase enzyme complex with the following activities: an adenylation

domain (A\*) wherein, the adenylation domain comprises a peptide sequence according to SEQ ID NO. 1, an acyl carrier protein (ACP), an elongation module (EM) of polyketide synthases (PKS) comprising the following activities: (i) ketoacylsynthase (KS), (ii) acyl transferase (AT) (iii) acyl carrier protein (ACP 2), an aminotransferase (AMT), three to five elongation modules (EM) of non-ribosomal peptide synthetases (NRPS) comprising the following activities: (i) condensation domain (C), (ii) adenylation domain (A), (iii) thiolation domain (T) and a thioesterase (TE).

The inventors have found that microginin is the product of non-ribosomal synthesis. It is important to understand that microginin as previously identified in nature may also in part have been the product of ribosomal synthesis and further processed via various enzymatic reactions.

It is important to note that the nucleic acid claimed herein, i.e. a microginin synthetase enzyme complex may also be present in organisms other organisms than *Microcystis* sp., such as Nostoc, Anabaena, Plankthotrix or Oscillatoria. The term microginin shall thus not limit the invention to such nucleic acids producing synthetase enzyme complexes resulting in peptides officially termed “microginin”.

Herein, an adenylation domain (A\*) is understood to activate octanoic acid as an acyl adenylate and an acyl carrier protein (ACP) is understood to bind the octanoic acid adenylate as a thioester.

An elongation module (EM) of polyketide synthases (PKS) is also known e.g. from the Jamaicamide synthetase gene cluster isolated from *Lyngbya majuscula* (Chem. Biol. Vol. 11, 2004 pp 817-833. Structure and Biosynthesis of the Jamaicamides, new mixed polyketide-peptide neurotoxin from the marine cyanobacterium *Lyngbya majuscula*) herein comprises at least the following activities: (i) ketoacylsynthase (KS), (ii) acyl transferase (AT) and (iii) acyl carrier protein (ACP2). The AT is responsible for the recognition of malonyl-CoA, the KS is responsible for the Claisen-type-condensation of the activated octanoic acid adenylate with malonyl-CoA and the ACP2 is responsible for binding of the resulting decanoic acid. An aminotransferase (AMT) performs the  $\beta$ -amination of the decanoic acid.

The nucleic acid according to the invention may have three to five elongation modules (EM) of non-ribosomal peptide synthetases (NRPS) comprising at least the following activities: (i) condensation domain (C), (ii) adenylation domain (A), (iii) thiolation domain (T). The A is responsible for the activation of carboxyl groups of amino acids, the T is responsible for the binding and the transport of the activated intermediate, the C is responsible for the condensation of the activated amino acids with the growing peptide chain.

Finally the nucleic acid according to the invention shall contain a thioesterase (TE) activity which performs the cleavage of the final product from the synthetase complex.

One may envision that the nucleic acid according to the invention is present in a vector or a bacterial chromosome, in which case one may envision that the portions designated above while being in one cell need not all, be in, or on, one molecule. It is essential to the invention however, that a cell meant to produce microginin synthetase enzyme complex contains the activities designated above in order to produce an enzyme complex according to the invention which in turn may produce a microginin. Thus, the invention also encompasses derivatives of the nucleic acid molecule as outlined above having the function of a microginin synthetase enzyme complex.

The molecule is characterized by a special adenylation domain (A\*) which is unusual in that it is not similar to known adenylation domains found in other molecules encoding non-ribosomal enzyme complexes such as the microcystin synthetase gene cluster (Chem. Biol. Vol. 7 2000, pp 753-764: Structural organisation of microcystin synthesis in *Microcys-*

*tis aeruginosa* PCC 7806: In integrated peptide-polyketide-synthetase system) Molecules encompassed herein are those which carry this adenylation domain (A\*) as depicted in SEQ ID NO. 1 and at least an ACP whereby this ACP may stem from another known non-ribosomal enzyme complex, at least one EM of PKS whereby this EM may stem from another known non-ribosomal enzyme complex comprising at least the following activities: (i) KS, (ii) AT (iii) ACP, an AMT whereby this AMT may stem from another known non-ribosomal enzyme complex three to five EMs comprising at least the following activities: (i) C, (ii) A, (iii) T whereby these EMs may stem from another known non-ribosomal enzyme complex and a TE whereby this TE may stem from another known non-ribosomal enzyme complex. Chimeras whereby parts of the above are on one or more vectors and or integrated in chromosomes are equally encompassed by the invention as long as all the components are in one cell.

The invention also pertains to isolated nucleic acid molecules encoding a microginin synthetase enzyme complex comprising an adenylation domain which is 85% identical to SEQ ID NO. 1, more preferred 90% identical to SEQ ID NO. 1 most preferred 95% identical to SEQ ID NO. 1. Sequence identity herein is in percent of total sequence of the adenylation domains when aligned with conventional nucleotide alignment software, such as the best fit and or pileup programs of the GCG package

The invention also pertains to a microginin synthetase enzyme protein complex with the following activities: an adenylation domain (A\*) wherein, the adenylation domain comprises a peptide sequence according to SEQ ID NO. 1, an acyl carrier protein (ACP), an elongation module (EM) of polyketide synthases (PKS) comprising the following activities: (i) ketoacylsynthase (KS), (ii) acyl transferase (AT) (iii) acyl carrier protein (ACP 2), an aminotransferase (AMT), three to five elongation modules (EM) of non-ribosomal peptide synthetases (NRPS) comprising the following activities: (i) condensation domain (C), (ii) adenylation domain (A), (iii) thiolation domain (T) and a thioesterase (TE).

The invention in particular also relates to a nucleic acid molecule encoding an adenylation domain (A\*) wherein, the adenylation domain comprises a peptide sequence according to SEQ ID NO. 1.

The invention in particular also relates to a peptide molecule, an adenylation domain (A\*) wherein, the molecule comprises a peptide sequence according to SEQ ID NO. 1.

The invention in particular also relates to a nucleic acid molecule encoding an adenylation domain (A\*) wherein, the molecule comprises a nucleic acid sequence according to SEQ ID NO. 25.

In a preferred embodiment of the invention the nucleic acid additionally and optionally comprises sequences encoding the following activities or domains: a monooxygenase (MO), an integrated N-methyltransferase domain (MT) within one or more elongation modules (EM) of NRPS, a non-integrated N-methyltransferase (MT), a modifying activity (MA) wherein, said MA is selected from the group comprising the following activities: halogenase, sulfatase, glycosylase, racemase, O-methyltransferase and C-methyltransferase, two or more peptide repeat spacer sequences (SP) consisting of one or more repeats of being either glycine rich or proline and leucine rich, located adjacently upstream and downstream of the MO and/or another MA.

Herein MO is an enzyme catalyzing the hydroxylation of the decanoic acid, an integrated N-methyltransferase domain (MT) within one or more elongation modules (EM) of NRPS catalyses the methylation of the amide bond by the respective module and a non-integrated N-methyltransferase (MT) catalyzes the methylation of an amino group of the microginin. The term modifying enzyme stands for numerous enzymes such enzymes may add groups or create bonds, in a preferred

embodiment MA is selected from the group comprising the following activities: halogenase, sulfatase, glycosylase, racemase, O-methyltransferase and C-methyltransferase.

Nucleic acids encoding two or more peptide repeat spacer sequences (SP) consisting of one or more repeats being either glycine rich or proline and leucine rich have astonishingly been found by the inventors to aid in integration of novel MAs into existing microginin synthetase enzyme complexes. By means of placing such SPs adjacently to MAs the inventors are able to create microginin synthetase enzyme complexes (MSEC) comprising activities previously not found in MSECs. This in turn allows for the creation of novel microginins with potentially novel therapeutic properties. Thus the invention relates to nucleic acids encoding two or more peptide repeat spacer sequences (SP) consisting of one or more repeats being either glycine rich or proline and leucine rich may be positioned adjacently to a MA such as but not limited to a halogenase, a sulfatase, a glycosylase, a racemase, an O-methyltransferase or a C-methyltransferase. These SPs aid in ensuring that the "foreign" activity "works" in the enzyme complex. The inventors have found, that this is due to the lack of secondary structures in the SP peptide chains.

The nucleic acid according to the invention in a preferred embodiment optionally comprises the following sequences, nucleic acid sequences encoding protein sequences as follows:

An adenylation domain (A\*) according to SEQ ID NO. 1, an acyl carrier protein (ACP) according to SEQ ID NO. 2, an elongation module of polyketide synthases responsible for the activation and the condensation of malonyl-Co A: (i) ketoacylsynthase domain (KS) according to SEQ ID NO. 3, (ii) acyl transferase domain (AT) according to SEQ ID NO. 4, an acyl carrier protein domain (ACP 2) according to SEQ ID NO. 5, an aminotransferase (AMT) according to SEQ ID NO. 6, an elongation modules of non-ribosomal peptide synthetases responsible for the activation and condensation of alanin: (i) condensation domain (C) according to SEQ ID NO. 7, (ii) adenylation domain (A) according to SEQ ID NO. 8, (iii) thiolation domains (T) according to SEQ ID NO. 9, an elongation modules of non-ribosomal peptide synthetases responsible for the activation and condensation of leucin: (i) condensation domain (C2) according to SEQ ID NO. 10, (ii) adenylation domain (A 2) according to SEQ ID NO. 11, (iii) thiolation domain (T 2) according to SEQ ID NO. 12, an elongation modules of non-ribosomal peptide synthetases responsible for the activation and condensation of tyrosine 1: (i) condensation domain (C 3) according to SEQ ID NO. 13, (ii) adenylation domain (A 3) according to SEQ ID NO. 14 (iii) thiolation domain (T 3) according to SEQ ID NO. 15, an elongation modules of non-ribosomal peptide synthetases responsible for the activation and condensation of tyrosine 2: (i) condensation domain (C4) according to SEQ ID NO. 16, (ii) adenylation domain (A 4) according to SEQ ID NO. 17, (iii) thiolation domain (T 4) according to SEQ ID NO. 18, a thioesterase (TE) according to SEQ ID NO. 19, a monooxygenase (MO) according to SEQ ID NO. 20, two or more peptide repeat spacer sequences (SP1/SP2) according to SEQ ID NO. 21 and 22, an integrated N-methyltransferase domain (MT) within the elongation module (EM) of the NRPS responsible for the activation and condensation of leucin according to SEQ ID 23 and a non-integrated N-methyltransferase (MT 2) according to SEQ ID NO. 24.

As outlined above, the minimal requirement according to the invention is a nucleic acid encoding a microginin synthetase enzyme complex with the following activities: an adenylation domain (A\*) wherein, the adenylation domain comprises a peptide sequence according to SEQ ID NO. 1, an ACP according to SEQ ID NO. 2, an elongation module (EM) of polyketide synthases (PKS) comprising the following activities: (i) ketoacylsynthase (KS) according to SEQ ID

NO. 3, (ii) acyl transferase (AT) according to SEQ ID NO 4, (iii) acyl carrier protein (ACP 2) according to SEQ ID NO. 5, an aminotransferase (AMT) according to SEQ ID NO. 6, three to five elongation modules (EM) of non-ribosomal peptide synthetases (NRPS) comprising the following activities: (i) condensation domain (C) according to SEQ ID NO. 7, (ii) adenylation domain (A) according to SEQ ID NO. 8, (iii) thiolation domain (T) according to SEQ ID NO. 9 and a thioesterase (TE) according to SEQ ID NO. 10. A molecule comprising the above sequences is preferred herein.

The invention explicitly also relates to analogs hereto, additionally comprising, e.g. other activities and/or spacer regions both transcribed and non-transcribed.

It is apparent to those skilled in the art, that amino acids may be exchanged maintaining the enzymatic activity required. Thus, the invention also relates to molecules with sequences which are not identical to those outlined above however, altered only in so far as the enzymatic activity desired is retained.

The nucleic acid according to the invention may contain nucleic acids selected from the group comprising: an adenylation domain (A\*) according to SEQ ID NO. 25, an acyl carrier protein (ACP) according to SEQ ID NO. 26, an elongation module of polyketide synthases encoding for the activation and the condensation of malonyl-Co A: (i) ketoacyl-synthase domain (KS) according to SEQ ID NO. 27, (ii) acyl transferase domain (AT) according to SEQ ID NO. 28, (iii) acyl carrier protein domain (ACP 2) according to SEQ ID NO. 29, an aminotransferase (AMT) according to SEQ ID NO. 30, an elongation modules of non-ribosomal peptide synthetases encoding for the activation and condensation of alanin: (i) condensation domain (c) according to SEQ ID NO. 31, (ii) adenylation domain (A) according to SEQ ID NO. 32, (iii) thiolation domain (T) according to SEQ ID NO. 33, an elongation modules of non-ribosomal peptide synthetases encoding for the activation and condensation of leucin: (i) condensation domain (C 2) according to SEQ ID NO. 34, (ii) adenylation domain (A 2) according to SEQ ID NO. 35, (iii) thiolation domain (T 2) according to SEQ ID NO. 36, elongation modules of non-ribosomal peptide synthetases encoding for the activation and condensation of tyrosine 1: (i) condensation domains (C3) according to SEQ ID NO. 37, (ii) adenylation domains (A 3) according to SEQ ID NO. 38, (iii) thiolation domains (T 3) according to SEQ ID NO. 39, elongation modules of non-ribosomal peptide synthetases encoding for the activation and condensation of tyrosine 2: (i) condensation domains (C4) according to SEQ ID NO. 40, (ii) adenylation domains (A 4) according to SEQ ID NO. 41, (iii) thiolation domains (T 4) according to SEQ ID NO. 42, a thioesterase (TE) according to SEQ ID NO. 43, a monooxygenase (MO) according to SEQ ID NO. 44, two or more peptide repeat spacer sequences (SP1/2) according to SEQ ID NO. 45 and 46, an integrated N-methyltransferase domain (MT) within the elongation module (EM) of the NRPS encoding for the activation and condensation of leucin according to SEQ ID 47 and a non-integrated N-methyltransferase (MT 2) according to SEQ ID NO. 48.

As outlined above, the minimal requirement according to the invention is a nucleic acid encoding a microginin synthetase enzyme complex with the following activities: an adenylation domain (A\*) wherein, the adenylation domain is a nucleic acid sequence according to SEQ ID NO. 25, an ACP with a nucleic acid sequence according to SEQ ID NO. 26, an elongation module (EM) of polyketide synthases (PKS) comprising the following activities: (i) ketoacylsynthase (KS) with a nucleic acid sequence according to SEQ ID NO. 27, (ii) acyl transferase (AT) with a nucleic acid sequence according to SEQ ID NO 28, (iii) acyl carrier protein (ACP 2) with a nucleic acid sequence according to SEQ ID NO. 29, an aminotransferase (AMT) with a nucleic acid sequence according

to SEQ ID NO. 30, three to five elongation modules (EM) of non-ribosomal peptide synthetases (NRPS) comprising the following activities: (i) condensation domain (C) with a nucleic acid sequence according to SEQ ID NO. 31, (ii) adenylation domain (A) with a nucleic acid sequence according to SEQ ID NO. 32, (iii) thiolation domain (T) with a nucleic acid sequence according to SEQ ID NO. 33 and a thioesterase (TE) with a nucleic acid sequence according to SEQ ID NO. 43. A molecule comprising the above sequences is preferred herein.

The invention also relates to nucleic acid molecules with sequences which are not identical to those outlined above however, altered only in so far as the enzymatic activity desired is retained. A particular one skilled in the art will know that positions in nucleic acid triplets may "wobble" and these positions may thus be altered with no influence on the peptide sequence. Further multiple amino acids are encoded by more than one DNA triplet. One skilled in the art will know that one may alter such triplets maintaining the amino acid sequence. Thus said sequences are equally encompassed by the invention.

The invention also pertains to isolated nucleic acid molecules encoding a microginin synthetase enzyme complex comprising an adenylation domain which is 85% identical to SEQ ID NO. 25, more preferred 90% identical to SEQ ID NO. 1 most preferred 95% identical to SEQ ID NO. 1. Sequence identity herein is in percent of total sequence of the adenylation domains when aligned with a conventional amino acid alignment software such as the best fit and or pileup programs of the GCG package.

In a preferred embodiment the one or more nucleic acids according to the invention are organized in sequence parts encoding the microginin synthetase enzyme complex in an upstream to downstream manner as depicted in FIG. 1. In a particularly preferred embodiment the activities and domains are arranged as shown and on one molecule.

The nucleic acid molecule may be part of a vector. Such vectors are in particular, bacterial artificial chromosomes (BAC), Cosmids or Fosmids, and Lambda vectors. Preferred plasmid vectors which are able to replicate autonomously in cyanobacteria are derived from the pVZ vectors. Preferred fosmid vectors which are able to replicate autonomously in cyanobacteria are derived from the pCC1FOS™ and pCC2FOS™ vectors (Epicentre Biotechnologies). The integration of the nucleic acid according to the invention into the vector is a procedure known to those skilled in the art (Molecular Cloning: A Laboratory manual, 1989, Cold Spring Harbour Laboratory Press) or in the manuals of manufactures of kits for creation of genomic libraries (e.g. Epicenter Biotechnologies).

In a preferred embodiment the invention concerns a microorganism transformed with a nucleic acid according to the invention. The nucleic acid according to the invention may be integrated into the chromosome of the host organism or may be present on a separate vector (see also examples). It is preferred that the phototrophic cyanobacterial host organism is selected for the group comprising: *Synechocystis* sp., *Synechococcus* sp., *Anabaena* sp., *Nostoc* sp., *Spirulina* sp., *Microcystis* sp . . . . Cells are cultured as follows:

Media: Bg 11 (for cultivation of cyanobacteria)

Aeration: air containing 0.3-3.0% carbon dioxide

Light intensity: 40-100  $\mu\text{E}/\text{m}^2\cdot\text{s}$  (diameter of illuminated culture vessels of photobioreactor  $d=4-12$  cm)

Cell density at harvest:  $\text{OD}_{750} 1-2$

And if the host is *Microcystis aeruginosa*:

Light quality: Additional red light illumination with 25  $\mu\text{E}/\text{m}^2\cdot\text{s}$  for 24-48 hours before harvesting.

It is preferred that the heterotrophic host organism is selected for the group comprising: *E. coli* and *Bacillus* sp. due to a more suitable GC content and codon usage than other heterotrophic bacteria.

In case of using *E. coli* for the heterologues expression of the microginin synthetase a phosphopanthetein transferase (Ppt) has to be co-expressed in order to enable the synthesis of microginin. The co-expression of the Ppt from a microginin producing strain would be preferred. Other Ppt's with a broad specificity even from heterotrophic organisms like *Bacillus* sp. are also suitable.

In one embodiment of the invention the invention relates to a method of producing a microginin, comprising culturing a cell under conditions under which the cell will produce microginin, wherein said cell comprises a nucleic acid encoding a recombinant microginin, according to the invention, and wherein said cell does not produce the microginin in the absence of said nucleic acid.

The inventors have identified nucleic acid sequences which for the first time make it possible to detect nucleic acids encoding a microginin synthetase enzyme complex. This has been extremely difficult, due to the fact that other gene clusters which encode non-ribosomal protein producing complexes share sequence similarity with the present cluster claimed herein. Such primers or probes according to the invention are selected from the group of, a) nucleic acid according to SEQ ID NO. 49 (Primer A), b) nucleic acid according to SEQ ID NO. 50 (Primer B), c) nucleic acid according to SEQ ID NO. 51 (Primer C), d) nucleic acid according to SEQ ID NO. 52 (Primer D), e) nucleic acid according to SEQ ID NO. 53 (Primer E), f) nucleic acid according to SEQ ID NO. 54 (Primer F), g) nucleic acid according to SEQ ID NO. 55 (Primer G), h) nucleic acid according to SEQ ID NO. 56 (Primer H), i) nucleic acid according to SEQ ID NO. 57 (Primer I) and j) nucleic acid according to SEQ ID NO. 58 (Primer J). It is known to one skilled in the art that such primers or probes may be altered slightly and still accomplishes the task of specifically detecting the desired target sequence. Such alterations in sequence are equally encompassed by the invention. The primers or probes according to the invention may be applied in hybridization reactions and/or amplification reactions. Such reactions are known to one skilled in the art.

The invention also concerns a method for detecting a microginin synthetase gene cluster in a sample wherein, one or more of the nucleic acids according to the invention are, applied in an amplification and/or a hybridization reaction.

In a preferred embodiment of the method according to the invention primers D and F or H and J or E and I or E and A are added to a PCR reaction mixture comprising a sample and wherein, presence of an amplification product represents presence of microginin synthetase gene cluster and absence of an amplification product represents absence of a microginin synthetase gene cluster. As can be seen from the examples (example 3 below), certain combinations are preferred. Samples may be isolated DNA, prokaryotic cells stemming from plates or liquid cultures.

When performing an amplification reaction with primers D and F the most preferred amplification conditions are as follows: a) denaturing, b) 48° C. annealing and c) elongation (product size: 675 bp). These temperatures may vary a bit in the range of 2-8 degrees C.

When performing an amplification reaction with primers H and J the most preferred amplification conditions are as follows: a) denaturing, b) 54° C. annealing and c) elongation (product size: 1174 bp). These temperatures may vary a bit in the range of 2-8 degrees C.

When performing an amplification reaction with primers E and I the most preferred amplification conditions are as fol-

lows: a) denaturing, b) 56° C. annealing and c) elongation (product size: 1279 bp). These temperatures may vary a bit in the range of 2-8 degrees C.

When performing an amplification reaction with primers E and A the most preferred amplification conditions are as follows: a) denaturing, b) 57° C. annealing and c) elongation (product size: 621 bp). These temperatures may vary a bit in the range of 2-8 degrees C. Molarity is most commonly 0.2-1.0 μM for the primers. Buffers and other reagents depending on polymerase used.

When performing hybridisation reactions the above nucleic acids are usually labeled. Such labels may be radioactive or non-radioactive, such as fluorescent. The nucleic acid primers or probes may be applied, e.g. for the screening of libraries.

The invention also relates to antibodies against a peptide according to SEQ ID NO. 1 (A\*).

The creation of such antibodies is known to one skilled in the art. The antibodies may be polyclonal or monoclonal. Such antibodies may be labeled or non-labeled, they may also be altered in other form, such as humanized.

The inventors have astonishingly found that newly identified peptide repeat spacer sequences (SP) may be placed adjacently to MAs I in order to create novel hybrid gene clusters. These SPs act by spacing the novel activity or domain so that it is functionally active in the microginin synthetase enzyme complex.

The invention thus, further relates to nucleic acids encoding a peptide repeat spacer sequence (SP) wherein, the peptide sequence comprises at least 4 glycine amino acids per single repeat unit (SRU) or, at least 5 proline and/or leucine amino acids per SRU. A SRU within the SP is between 7 and 15 amino acids in length and, the SP comprises between 2 and 10 SRUs.

The invention further relates to peptides of a peptide repeat spacer sequence (SP) wherein, the peptide sequence comprises at least 4 glycine amino acids or, at least 5 proline and/or leucine amino acids, the single repeat unit (SRU) within the SP is between 7 and 15 amino acids in length and, the SP comprises between 2 and 10 SRU. In a preferred embodiment of the invention the SRU is between 9 and 13 amino acids in length in a particularly preferred embodiment the SRU is eleven amino acids in length. In a preferred embodiment the SP comprises between 3 and 9 SRU.

In a preferred embodiment the nucleic acid encoding the peptide repeat spacer sequence (SP) according to the invention, encodes a peptide SRU as shown in SEQ ID NO. 20 or SEQ ID NO. 21. In a further embodiment the peptide repeat spacer sequence (SP) according to the invention, comprises or contains a sequence as shown in SEQ ID NO. 20 or SEQ ID NO. 21. In a further embodiment the nucleic acid according to the invention has a sequence as laid down in SEQ ID NO. 43 or SEQ ID NO. 44.

Not only by means of the above mentioned SPs but in particular because of these the inventors are able to create enzyme complexes resulting in microginin variants which may not be found in nature. This is an essential aspect of the present invention. The invention provides for, for the first time a simple method of producing recombinant microginin variants comprising, modifying the nucleic acid according to the invention in vitro or in vivo, growing a recombinant cell comprising said recombinantly modified nucleic acid encoding a microginin synthetase under conditions which lead to synthesis of a microginin and, recovering the synthesized microginin.

In a preferred embodiment of said method according to the invention, said modifying of said nucleic acid may be an action selected from the group of one or more of the following actions: a) inactivation of one or more of the MTs present, b) substitution of one or more of the MTs present with a halo-

genase, a sulfatase, a glycosylase, a racemase, an O-methyltransferase or a C-methyltransferase, c) inactivation of the MO, d) substitution of the MO with a halogenase, a sulfatase, a glycosylase, a racemase, an O-methyltransferase or a C-methyltransferase, e) inactivation of the AMT, f) substitution of the AMT with a halogenase, a sulfatase, a glycosylase, a racemase, an O-methyltransferase or a C-methyltransferase, g) inactivation of the PKS module, h) substitution of the entire PKS module with an alternative PKS module and/or substitution of one or more of the domains (KS, AT, ACP) therein, i) inactivation of the A\* domain, j) substitution of the A\* domain with alternative A domains, k) inactivation of one or more of the NRPS modules and l) substitution of one or more of the NRPS modules with alternative NRPS modules and/or substitution of one or more of the domains (C, A, T) therein.

Halogenases, sulfatases, glycosylases, racemases, O-methyltransferases or C-methyltransferases are known from prokaryotes. These enzymes are encoded by genes of the secondary metabolism in particular NRPS/PKS systems.

Alternative PKS-systems, entire modules as well as single domains (KS, AT, ACP) are found in cyanobacteria as well as Actinomycetes, Myxobacteria, *Bacillus* among the bacteria.

Alternative NRPS-systems, entire modules as well as single domains (C, A, T) are found in cyanobacteria as well as Actinomycetes, Myxobacteria, *Bacillus* among the bacteria.

In a preferred embodiment the above are from cyanobacteria.

It is important to note, that said inactivation and/or substitution may done in many ways, e.g. inactivation may imply deleting the complete activity or domain, or may imply inactivation by means of a single nucleotide exchange.

The methods are known to those skilled in the art and comprise basic molecular biological methods such as DNA isolation, restriction digestion, ligation, transformation, amplification etc.

In a preferred embodiment said alternative modules or domains which are used for substitution of the original module or domain, additionally may comprise one or more SP nucleic acids according to the invention located adjacently upstream of the module or domain used for substitution and one or more SP nucleic acids according the invention located adjacently downstream of the module or domain used for substitution. Thus, in this embodiment of the invention a construct is made comprising the domain which is to be entered into the original nucleic acid according to the invention, further comprising one or more SPs located adjacently in an upstream and downstream manner. This construct is then ligated into the original microginin synthetase encoding nucleic acid. The resultant construct is then brought into a host by means of transformation for either a) integration into the host chromosome or b) with a self-replicating vector.

The polypeptides, i.e. proteins can be any of those described above but with not more than 10 (e.g., not more than: 10, nine, eight, seven, six, five, four, three, two, or one) conservative substitutions. Conservative substitutions are known in the art and typically include substitution of, e.g. one polar amino acid with another polar amino acid and one acidic amino acid with another acidic amino acid. Accordingly, conservative substitutions preferably include substitutions within the following groups of amino acids: glycine, alanine, valine, proline, isoleucine, and leucine (non polar, aliphatic side chain); aspartic acid and glutamic acid (negatively charged side chain); asparagine, glutamine, methionine, cysteine, serine and threonine (polar uncharged side chain); lysine, histidine and arginine; and phenylalanine, tryptophane and tyrosine (aromatic side chain); and lysine, arginine and histidine (positively charged side chain). It is well

known in the art how to determine the effect of a given substitution, e.g. on pK<sub>i</sub> etc. All that is required of a polypeptide having one or more conservative substitutions is that it has at least 50% (e.g., at least: 55%; 60%; 65%; 70%; 75%; 80%; 85%; 90%; 95%; 98%; 99%; 99.5%; or 100% or more) of the ability of the unaltered protein according to the invention.

In preferred embodiments the polynucleotides, i.e. nucleic acids of the present invention also comprise nucleic acid molecules which are at least 85%, preferably 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to those claimed herein.

The determination of percent identity between two sequences is accomplished using the mathematical algorithm of Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90: 5873-5877. Such an algorithm is incorporated into the BLASTN and BLASTP programs of Altschul et al. (1990) *J. Mol. Biol.* 215: 403-410. BLAST nucleotide searches are performed with the BLASTN program, score=100, word length=12, to obtain nucleotide sequences homologous to the nucleic acids according to the invention. BLAST protein searches are performed with the BLASTP program, score=50, wordlength=3, to obtain amino acid sequences homologous to the EPO variant polypeptide, respectively. To obtain gapped alignments for comparative purposes, Gapped BLAST is utilized as described in Altschul et al. (1997) *Nucleic Acids Res.* 25: 3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs are used.

## FIGURES

FIG. 1 depicts the structure of microginin.

FIG. 2 depicts the microginin synthetase gene cluster and the biosynthetic pathway of microginin.

## EXAMPLES

### Example 1

#### Method for Detecting Gene Clusters According to the Invention

Strains carrying a gene cluster encoding a microginin synthetase complex can be distinguished from strains not carrying such a gene cluster performing a PCR reaction using RedTaq ReadyMix PCR Reaction Mix with MgCl<sub>2</sub> (Sigma) and primer pairs and the corresponding annealing temperatures as described. In particular the PCR conditions are as follows: an initial denaturation for 1 minutes at 95° C., followed by 30 cycles of denaturation at 95° C. for 30 seconds, elongation at said annealing temperatures for 30 seconds and extension at 72° C. for 1 kb of product size.

### Example 2

#### Method for Optimised Cultivation of Microginin Producing *Microcystis* spp

Strains. Media: Bg 11 (for cultivation of cyanobacteria)

Aeration: air containing 0.3-3.0% carbon dioxide

Light intensity: 40-100 pE/m<sup>2</sup>\*s (diameter of illuminated culture vessels of photobioreactor d=4-12 cm)

Light quality: Additional red light illumination with 25 μE/m<sup>2</sup>\*S for 24-48 hours before harvesting.

Cell density at harvest: OD<sub>750nm</sub> 1-2

Tables

TABLE 1

SEQ ID NO. 1	MTINYDGLQEPFNKFS TLVELLRYRASSQPERLAYIFLRDGEIE EARLTYGELDQKARAAAYLQSLAEAGERGLLLYPGLDFISAF FGCLYAGVVAI PAYPPRRNQNLRLQAI IADSQARFTFTNAALF PSLKNQWAKDPELGAMEWIVTDEIDHHLREDWLEPTLEKNSLAF LQYTSGSTGTPKGVVMVSHHLL INSADLRDGRWGHQDQSVMTWL PTFHDMGLI YGVI QPLYKGF L CYMSPAS FMERPLRWLQALS DK KATHS AAPNFAYDLCVRKIPPEKRATLDL SHWCMALNGAEPVRA EVLKKFAEAFQVSGFKATALCPGYGLAEATLKVTAVSYDSPPYF YPVQANALEKNKI VGATE TD TNVQTLVGCWTT IDTQIVIVNPE TLKPCSP EIVGEI WVSGSTIAQGYWGK PQETQETFOAYLADTGA GPFLRTGDLGFIKDGELFITGRLEKIELIRGRNNYPQDI ELTVQ NSHPALRPS CGAAFTVNGEKEKLVVVQEVERTWLRKVDIDEVK RAIKAVVQYEDLQVYAIALIR TGS LPKTSSGKIQRRSCR AKFL EGSLEILG
SEQ ID NO. 2	MSTEI PNDKKQPTLTKIQNLVAYMTEMMEVDEDEIDLSPVFDE YGLDSSMAVALIADLEDWLRDLHRTLIYDYPTLEKLAKQVSEF ACP
SEQ ID NO. 3	MEPIAII GLACRFP GADNPEAFWQLMRNGVD AIADIPPERWDIE RFYDPTPATAKMYSRQGGFLKNVDQFDPQFFRISPLEATYLDP QORLLLEV TWEAL ENAAI VPETLAGSQSGVFIGISDVVYHRLAY QSPTNLTAYVGTGNSTSI AANRLSYLFDLRGSP LAVDTACSSSL VAVHLACQSLQSQENLCLVGGVNLILSPETT VVFSQARMIAPD SRCKTFDARADGVVRS EGGCVVVLKRLRDAIQDGDRI LAVIEGS AVNQDGLSNGLTAPNGPAQQA V IRQALANAQVKPAQI SYVEAHG TGT ELGDP I EVKSLKAVLGEKRSLDQTCWLGSVKNTI GHLEAAA GMAGL I KVVLC LQHQE I PPNLHFQTLNPIYISLADTAFAIPTQAQ PWR TKPPKSGENGVERRLAGLS SFGFGGTNSHVIL
SEQ ID NO. 4	VFLFAGQGSQYVGMGRQLYETQPIFRQTLDRCAEILRPHLDQPL LEILYPADPEAETASFYLEQTAYTQPTLFAFEYALAQLRWSWGI EPAAVIGHSVGEYVAATVAGALSLEEGTLIAKRAKLMQSLPKN GTMIAVFAAEERVKAVIEPYRTDVAIAAVNGPENFVISGKAPII ABEI IHLTAAGIEVRLPLKVS HAFHSHLLEPILDSLEQEAASIS YQPLQIPLVANLTGEVLP EGATIEARYWRNHARNPVQFYGSIQT LIEQKPSLFLVSPKPTLSRLGQQCCPERSTTWLFSLALPQQEE EQSLLNSLAILYDSQGA E
SEQ ID NO. 5	ITLQTLVGNLLQLSPADVNHTPFLEMGADSI VMVEAVRRIENT YNVKI AMRQLFEBLSTLDALATYL ACP 2
SEQ ID NO. 6	KEMLYPIVAQRSQGSRIWDVDGNEYIDMTMGQGVTLFGHQPDFI MSALQSQLTEG IHLNPRSP I VGEVAALICELTGAERACPCNSGT EAVMAAIRIARATTGRSKIALFEGSYHGHDGTLFRNQI IDNQL HSFPLALGVPPSLSSDVVLDYGS AEALNYLQ TGGQDLAAVLVE PIQSGNPLLPQOQFLQSLRQITSQMGIALIFDEMITGFRSHPGG AQALFGVQADIATYGVKVVAGMPIGV IAGKAHYLDSIDGGMWRY GDKSYPGVDRTPFGGTFNQHPLAMVAARAVLTHLKEQPGGLQOQ LTERTALADTLNHYFQAEVPIKIEQFS SFFRFALS GNLDLLF YHMVEKGIYVWEWRKHFLSTAHT EADLAQFVQAVKDSITELR
SEQ ID NO. 7	GGDQVPLTEAQRQLWILAQLGDNGSVAYNQSVTLQLSGPLNPVA MNQAIQQISDRHEALRTKINAQGDSQEILPQVEINCPILDFSLD QASAQQAQEWLKEESEKPFDL SQGLVLRWHLLKLEPELHLLVL TAHHIISDGWSMGVILRELGE LYS AKCQGV TANLKTQKPRELI EWQSQPSQGEELKKQAYWLATLADPPVNLNPTDKPRPALPSYQ ANRRSLTLD SQFTEK LKQPSRKQCTLLM TLVSYN I LVHRLTG QDDILVGLPASGRGLLDS EGMVGYCTHFLPIRSQLA
SEQ ID NO. 8	TYSELNCRANQLALHYLQKLGVGPEVLVGI LVERSLEMI VGLLG ILKAGGAYVLPDPDYP PERLQFMLED SQFLLTQQHLLSFAQ SSETATPKIICLDSYQIISQAKNINPENS VTTSNLAVIYTSG STGKPKGVMMNHVAISNKLLWQD TYPLTTEDCILQKTPFSFDV SVWELFWPPLLNGARLVFAKPNGHK DASYLVNLIQEQQVTTLHFV SSMQLFLTEKDV EKNLSKRVICSGEALS ELQERFFARLVCE LHNLYGPT EAAIHVTFWQCQSDSNLKTVP IGRPIANIQIYILDS HLQVPVIGVIGELHIGVGLARGYLNRP ELTAEKFIANPPASLD PPLTPLDKGGESEYKTFKKGGEQPSRLYKTDGLARYLPDGKIEY LGRIDNQV KIRGFRIELGEIEAVLLSHPQVREAVV
SEQ ID NO. 9	EATAAIFGQVLKLEKVIYDNFFEIGGNSLQATQVISRLRESFA LELPLRRLFEQPTVADLALAV T

TABLE 1-continued

SEQ ID NO. 10	PRDQQLPLSFAQSR LWFLYQLEGATGTYNMTGALSLSGPLQVEA LKQALRTIIQRHEPLR TSFQSV DGVVQVINPYPVWELAMVDLT GKETEAEKLAYQESQTPFDL TNSPLL RVTL LKQLQPEKHILLINM HHIISDGWSIGV FVRELSHL YRAFVAGKEPTLPILPIQYADFAV WQREW LQGVLA AQLEYWKRQLADAPLLELPTDRPRPAIQTFQ GKTERFELDRKLTQELKALSQQSGCTLFMTLLAAGVVLRSYSG QTDIVIGSAIANRNRQDIEGLI GFFVNTLALRLDLS
SEQ ID NO. 11	TYGELNHRANQLAHYLQSLGVTKQIVGVYLSERSLEMAIGFLGI LKAGAAALPIDPEYPSVRTQPILEDTQLSLLTQAE LAEKLPQT QNKIICLDRDWEI TSQPQTNLDLKI EPNNLAIYCIYTSGSTGQP KGVLSHQALLNLI FWHQQAFEIGPLHKATQVAGI AFDATVWEL WPYLTGACINLVQPNILLSPTDLDRDWLLNREITMSFVPTPLAE KLLSLDWPNHSC LKTL LGGDKLHFYPAASLPFQV INNYGPTEN TVVATSGLVKSSSSHFGTPTI GRPIANVQIYLLDQNLQPVPIG VPGELHLGGAGLAQGYLNRP ELTAEKFIANPDPPLTLPDKGGE EPSLKYKTGDLARYLPDGNVEFLGRIDNQVKIRGFRIETGEIEA VLSQYFLLAESV V
SEQ ID NO. 12	AQLTQIWSEVLG LERIGVKDNFFELGGHSLLATQVLSRINSAFG LDLSVQIMFESPTIAGIAGYI T2
SEQ ID NO. 13	ARDGHLPLSFAQQR LWFHLHYLSPDSRSYNTLEILQIDGNLNTV LEQSLGELINRHEIFRTTFFTVSGEPIQKIALPSRFQKVDNYQ DLDENEQSAKIQVAE LEAGQAFDLTVGPLIQKLLQLSPQKSV LLLKMHIIYD GWSFGILIR ELSALYEAFKLNLANPLALS IQY ADFAVWQRQYLSGEVLDKQLNYWQEQLATVSLPTLPTDRRPA IQTFQGGVERFQLDQNVTOGLKKGQDQVATLFMTLLAGFGVLL SRYSGQSDLMV GSP IANRNQAAIEPLI GFFANTLALRINLS
SEQ ID NO. 14	TYELNHRANQLAHYLQTLGVGAEVLVGI SLERSLEMI IGLLGI LKVGAGYLP LDPDYPTERLQMLLED SQVPFLI THSSLLAKLPPS QATLILCDHIQEQISQYSPDNLQOQLT PANLANVIYTSGSTGKP KGMVVEHKGVLNLA LAQIQSFAVNHNRSRVLQFASFPD ACISEI LMTFGSGATLYLAQKDALLPGQPLIERLVKNGI THVTLPPSALV VLPQEPLRNLETLI VAGEACSLDLVKQWS IDRNFFNAYGPT EAS VCATIGQCYQDDLVKVTIGKAIANVQIYILDAFLQPVPIGVSGEL YIGVGVARGYLNRP ELTQEFIANPFPNDPSRLYKVTGDLARY LPDGNIEYLGRIDNQVKIRGFRIELGEIEAVLSQCPDVQNTAV
SEQ ID NO. 15	EILAQIWGQVLKIERVSREDNFFELGGHSLLATQVMSRLRETFQ VELPLRSLFTAPTIAELALTI T3
SEQ ID NO. 16	NDSANLPLSFAQQR LWFLDQLEPN SFYHVGAVRLEGTLNITALEQSLKEIINRHEALRTNFI TIDGQATQIHPTINWRLSVVD CQ NLTDTSLEIAEAEKPFNLAQDCLFRATLFVRSPL EYHLLVTMH HIVSDGWSIGVFFQELTHLYAVYVNLQGLPSLTPIKIQYADFAVW QRNLQGEILSNQNLNYWREQLANAPAFPLP TDRPRPAIQTFIG SHQEFKLSQPLS QKLNQLSQKHGVTLFMTLLA AFTLLYRYTGO ADILVGSPIANRNRKEIBGLIGFFVNTLVLR LSLD
SEQ ID NO. 17	TYAELNHQANQLVHYLQTLGIGPEVLVAI SVERSLEMI IGLLAI LKACGAYLP LAPDYPTERLQFMLED SQASFLI THSSLLEKLPSS QATLILCDHIQEQISQYSPDNLQOQLT PANLANVIYTSGSTGKP KGMVVEHRLVNLASSQIQSFAVNHNRSRVLQFASFPD ACISEI LMTFGSGATLYLAQKNDLLPGQPLMERLEKNKI THVTLPPSALA VLPKPLPNLQTLTI VAGEACPLDLVKQWSVGRNFFNAYGPTETS VCATIGQCYQDDLVKVTIGKAIANVQIYILDAFLQPVPIGVSGEL YIGVGVARGYLNRP ELTAEKFIANPFPDPLTPLKKGDKSYET FKKGEEQPSKLYKTGDLARYLPDGNIEYLGRIDNQVKIRGFRIE LGEIEAVLSQCPDVQNTAV
SEQ ID NO. 18	LQLAQIWSEILGINNIGIQENFFELGGHSL LAVSLINRIEQKLD KRLPLTSLFQNGTIIASLAQLL T4
SEQ ID NO. 19	TPFFAVHPIGGNVLCYADLARNLGTQKPPYGLQSLG LSELEKTV ASI EEMAMIYIEAIQTVQASGPYYLGGWSMGGVIAFEIAQQLLT QGQEVALLALIDSYSPSLLNSVNREKNSANLSTEEFNEDIN IAY SFIRDLASIFNQEISFSGSEL AHFTSDELLDKFITSWQETNLLP SDFGKQVKTWFVQINHQALSSYSPKTYLGRSVFLGAEDSSI KNP GWHQ



TABLE 1-continued

	TTGAAGGCTCCTATCATGGACATGCGGACGGAAACCCCTTTTAGG
	AACCAAATTTATGATAACCAACTCCACTCTTTTCCCCTAGCTCT
	AGGCGTCCCCCAGCCTTAGTTCAGATGTGGTGGTATTGGACT
	ATGGCAGTGGGAGCTCTGAACATTTACAAACCCAGGGGCAG
	GATTTAGCGGCGTCTTAGTAGAACCAATTCAAAGTGGCAATCC
	TCTACTCAACCCCAACAATTTCTCCAAGTCTGCGACAATTA
	CCAGTCAAATGGGCATTGCCCTGATTTTGTGATAATGATTACG
	GGTTTTTCGATCGCACCAGGGGAGCGCAAGCTTTATTTGGAGT
	ACAGGCGGATATTGCCACTATGGCAAAGTAGTTGCGGGAGGAA
	TGCCCATTTGGAGTTATTGACAGTAAAGGCCATTATCTGGACAGC
	ATTGACGGGGGAAATGTCGGTATTAGCGATAAATCTATCTCTGG
	GGTGGACAGAACCCTTTTGGGGAAACCTTTAATCAGCATCCGT
	TAGCAATGGTAGCGGTAGGGCTGTCTGACCCATTTAAAGGAG
	CAGGGGCCAGGTCTGCAACAACAATTAACGTAACGCACTGCGGC
	CTTAGCCGATACACTG
	AATCATTATTTTCAAGCCGAAGAAGTTCCTATTAATAATCGAACA
	GTTTAGTCTTCTTCCGGTTTGGCCCTCTCTGGCAATTTGGATT
	TACTTTTCTATCACATGGTAGAAAAAGGTATTTATGTCTGGGAA
	TGGCGTAAACATTTCTTTCACCCGCCATACGGAAGCCGATCT
	TGCCCAATTTGTCCAAGCGGTTAAGGATAGCATCACAGAATTGC
	GT
SEQ ID	GGGGGGATCAAGTCCCTCTCACCGAAGCCCAACGACAACGTG
No. 31	GATTTTGGCTCAATTAGGAGACAACGGCTCTGTGGCTATAAACC
C nucl	AATCAGTGACATTGCAATTAAGTGGCCCATTAATCCCCTCGCA
acid	ATGAATCAAGCTATTCAACAATAACAGCATCGCCATCAAGCGTT
	ACGAACCAAAATTAATGCCAGGGAGATAGTCAAGAATCCTGC
	CCCAGTCTGAATTAATCTC
	CCTATCTAGACTTCAAGCTTTGACCAAGCTTCGGCCCAACAGCA
	AGCAGAACAAATGGTTAAAGGAAGAAAGTGAATAACCTTTGATT
	TGAGCCAGGGTCTCTCGTGGCTTGGCATCTACTCAAATTAGAA
	CCAGAATTAACATTTGTTAGTATTAACGGCCCATCACATTATCAG
	TGACGGTGGTCAATGGGGTAATCCTTCGGGAATTAGGAGAGT
	TATATTCAGCCAATGTCAGGGTGTACGGCTAATCTTAAACC
	CCAAAACAGTTTCGAGAATTGATTGAATGGCAAAGCCAGCCAAG
	CCAAGGGGAAGAACTGAAAAACAGCAAGCCTATTGGTTAGCAA
	CCCTTGGC
	GATCCCCCTGTTTGAATTTACCCACTGACAAAACCTCGTCCAGC
	TTTACCAGTTACCAAGCTAATCGTCGAAGTCTAATTTAGATA
	GCCAAATTTACAGAAAACTTAAAGCAATTTAGTCTGAACAGGGC
	TGTACTTGTCTGATGACCTGTTATCGGTTTATAACATTTCTCGT
	TCATCGTTTACGGGACAGGATGATATTTCTGGTGGTCTGCCAG
	CCTCTGGACGGGGCTTTTAGATAGTGAAGGTATGGTGGGTTAT
	TGCACCCATTTTTTACCAATTCGAGTCAATTAGCA
SEQ ID	ACTTACAGTGAATTAATTTGTCGAGCCAATCAGTTAGCACATTA
No. 32	TTTACAAAAATTAGGAGTTGGGCGAGGCTTAGTCCGTTATTT
A nucl	TGGTCAACGTTCTTTTGAAGATGATGTCGGATTGTTAGGGATT
acid	CTCAAGGCTGGGGGAGCCTATGTACCTCTGACTGACTATCC
	CCCTGAACGCTCTCAATTTATGTTAGAAGATAGTCAATTTTTTTC
	TCCTCTTAAACCAACAGCATTACTGGAATCTTTTGTCTAGTCT
	TCAGAAACGGCTTACTCCCAAGATTTATTTGTTGGATAGCGACTA
	CCAAATTAATTTCCAGGCAAGAAATTAATCCGAAAAATTCAG
	TCACAACGAGTAATCTTGCCATGTAATTTATACCTCTGGTTCG
	ACAGGTAAACCGAAGGGCGTGAATAAATCATGTTGCTATTAG
	TAATAAATTTGTTATGGGTACAAGACTTATCTCTAACCACAG
	AAGACTGATTTTACAAAAAATCCCTTTAGTTTGTGATGTTCA
	GTGTGGGAATTAATCTGGCCCTTACTAAACGAGCGGTTTGGT
	TTTTGCAAAGCCGAATGGCCATAAAGATGCCAGTTACTTAGTCA
	ATCTGATCAAGAGCAACAAGTAAACAACGCTACATTTTGTGTCT
	TCTATGCTACAGCTTTTCTGACAGAAAAAGAGCTAGAAAAATG
	TAATAGTCTTAAACGAGTCAATTTGATGTTGTAAGCCCTTCTT
	TAGAGCTTCAAGAACGTTTTTGTCTGTTAGTCTGTGAATTA
	CACAATCTTTATGGACCGACAGAAAGCCGCTATTATGTCACATT
	TTGGCAATGTCATCAGATAGCAATTTGAAAAACAGTACCATTG
	GTCGGCCGATCGCTAATATCCAAATTTACATTTTAGACTCTCAT
	TTACAGCGTACCTATTGGAGTAAATCGGAGAATTGCAACATTGGT
	GGGTTGGTTTGGCGCGGGGTTATTTAAACAGCCCTGAGTTAAC
	GGCGGAGAAATTTATTGCAAAATCCGTTTGGCTTCCCTTGATCCCC
	CCCTAACCCCTTGATAAGGGGGGAGATGAGAGCTATAAAACT
	TTTAAAAAGGGGGGAGAGCAACCATCAAGATTGTATAAA
	ACGGGAGATTAGTCTGTTATTTACCCTGAGGCAAGATTGAGTA
	TCTAGGGCGCATTTGATAATCAGTAAAAATTCGCGGTTTCCCGA
	TTGAATTTGGGGAAATTTGAAGCGGTTTTGCTATCCCATCCCCAG
	GTACGAGAAGCGGTCGTT

TABLE 1-continued

SEQ ID	GAGGCGATCGCCGCTATTTTTGGTCAAGTTTTAAACTGGAAAA
No. 33	AGTGGGAATTTATGATAACTTTTTTGGATCGGCGGTAATCTT
T nucl	TGCAAGCCACTCAAGTTATTTTACGCTTACGAGAAAGTTTTGCC
acid	CTAGAGTTGCCCTTGCCTCGCTGTTTGAACAACCGACTGTGGC
	GGATTTGGCTTTTAGCCGTA
SEQ ID	CCTCGTGTGGCCAAATACCCTCTCCTTTGCCAGTCCGCACT
No. 34	CTGGTCTTGTATCAATTAGAAGGACCCAGGAACTTATAACA
C2	TGACAGGGCCCTTGAGTTTAAAGGGCCCTCTCAGGTCGAAGCC
nucl	CTCAACAAGCCCTAAGAACTATCATTCAACGCCATGAGCCATT
acid	CGGTACCAGTTTCCAATCGGTTGACGGGTTCCAGTGCAGGTGA
	TTAATCCCTATCTGTTTGGGAATTAGCGATGGTTGATTTGACA
	GGAAAGGAGACAGAAAGCAGAAAAATTTGCCCTATCAG
	GAATCCCAAAACCCGTTTGAATTTGACCAATAGTCCCTTTGTTGAG
	GGTAACGCTCCTCAAATTACAGCCAGAAAAGCATATTTTATTA
	TTAATATGCACCATATATTTTCCGATGGCTGGTCAATCGGTGTT
	TTTTGTCGTGAATTTGCCATCTCTATAGGGCTTTTGTGGCGGG
	TAAAGAACCAACTTTACCAGTTTACCAATTGATATGCGGATT
	TTGCCGTTTGGCAGCGAGAGTGGTTACAGGGTAAAGGTTTAGCG
	GCTCAATTTGGAATATTGGAAGCGCAATTTGGCAGATGCTCCTCC
	TCTGCTGGAAGTCCCACTGATCGCCCTCGTCCCCTCAATCCAAA
	CCTTTCAAGGCAAGACAGAAAGATTGAGCTAGATAGGAAACTG
	ACCAAGAATTAAGGCATTAAGT
	CAACAGTCCGGTTGTACTTTATTTATGACTTTGTTGGCCGCTTT
	TGGGGTGGTTTTATCCGTTATAGTGGCCAGACTGATATCGTCA
	TTGGTTCGGCGATCGCCAACCGTAATCGCCAAAGACATTGAGGGG
	TTAATTTGGCTTTTTTGTAACTTTGGCGTTGAGGTTAGATTT
	ATCA
SEQ ID	ACCTATGGAGAATTAACCATCGCCCAATCAATTAGCTCACTA
No. 35	TCTTCAGTCTGTTAGGAGTCAACAAGAACAAATCGTCCGGGTTT
A2	ATCTGGAACGTTCCCTTGAATGGCGATCGGATTTTAGGTATT
nucl	CTCAAAGCAGGAGCCGCTATCTCCCATTTGATCTTGAATATCC
acid	CTCAGTACGCCCAATTTATTTCTGAAATACCCAACTTTTCGC
	TTCTTTAATCAGGCAAGAACTGGCAGAAAACTGCCCCAGACT
	CAAAAACAAATTTATCTGTCTAGATCGGACTGGCCA
	GAAATTAACCTCCCAACCCAGACAAACCTAGACTTAAAGATAGA
	ACCTAATAACCTAGCC
	TATGCTATCTATCTTCTGGTTCACAGGACAAACCAGGAGT
	ACTGATTTCCCATCAAGCCCTACTCAACTTAATTTTCTGGCATC
	AACAAGCGTTTGGATTGGCCCTTACATAAAGCGCACCAAGTG
	GCAGGATGCTTTGATGCAACGGTTTGGGAATTTGGGCCCTA
	TCTGACCACAGGAGCTGATTAATCTGGTTCCTCCAAATATTC
	TGCTCTCAGCAGGATTTACGGGATTTGGGATTTGGTCTTAAACCGA
	ATTACCATGAGTTTGTGCCAACTCCTTTAGCTGAAAAATTAAT
	ATCCTTGGATTGGCTAACCAATCTTGTCTAAAAACCTGTTAC
	TGGGAGGTGACAAACTTCAATTTTATCTGCTGCTGCTCCCTCC
	TTTCAGGTCATTAACTATGGCCCAACGGAATAACAGTGGT
	TGCGACTCTGGACTGGTCAATCATCTTCACTCATCACTTTG
	GAACTCCGACTATTGGTCTCCATTCAGCAAGCCCAATCTAT
	TTATTAGACAAAACCTACAACCTGCTCCCATTTGGTGTACCAGG
	AGAATTAACATTTAGTGGGGCGGTTTGGCGCAGGGCTATCTCA
	ATGCTCTGAGTAAACGGCTGAAAAATTTATTCGCAACCCCTTT
	GATCCCCCTAACCCCTTGATAAGGGGGGAGAGAACCCTC
	AAAACCTATAAAAACG
	GGAGACTTAGCCGTTATTTACCCGATGGCAATGTAGAATTTTT
	GGGACGTTATGCAATCAGGTAAAAATTCGGGGTTTTCGCATCG
	AACTGGGGAATCGAAGCCGTTTTAAGTCAATATTTCTATTA
	GACTGAAAGTGTAGTC
SEQ ID	GCTCAACTGACTCAAATTTGGAGTGAAGTTTTGGGACTGGAACG
No. 36	CATTGGCGTTAAGGACAACTTTTTGAATTTGGGAGGACATTCTC
T2	TTTTGGCTACCCAGGTTTTATCAAGAAATTAATTCAGCCCTTTGGA
nucl	CTTGATCTTTCTGTGCAAAATATGTTGAATCACCACAGCATCGC
acid	GGGCATTGCGGGTTATATT
SEQ ID	GCTAGAGACGGTCAATTTACCCTCTCTTTTGTCTCAACACGTTT
No. 37	ATGGTTTTTACATTATCTTTCCCTGATAGTCTGTTCTTACAATA
C3	CCCTGGAAATATGCAAAATGATGGGAATCTCAATCTGACTGTG
nucl	CTAGAGCAGAGTTTGGGGAAATTAATTAACCCGCAATGAAATTTT
acid	TAGAACAACATTTCCCACTGTTTTCAGGGGAACCGATCAGAAAA
	TTGCACTTCTAGTCTGTTTTCAGTTAAAAGTTGATTAATATCAA
	GATTTAGACGAAAAATGAACAATCAGCTAAAATTAACAAGTAGC
	AGAAATGGAAGCAGGACAAGCTTTTGAATTTAACCGTGGGGCCAC
	TGATTCAGTTTAAAGCTATTGCAATTTGAGTCCCGAAGTCCGGT
	CTGCTGTTGAAAAATGACCATATTTATCTATGATGGCTGGTCTTT
	TGGGATCTGATTCGGGAATTTACGGCTCTATACGAAGCAATTTT
	TAAAGAACTTAGCCAATCTCTCCCTGCGTGTCTATTGATGAT

TABLE 1-continued

		GCAGATTTTGGCGTTTGGCAACGTCAATATCTCTCAGGTGAGGT CTTAGATAAACAACTCAATATTTGGCAAGAACAGTTAGCAACAG TCTCTCTGTCTTACTCTTTACCAACGATAGACCCCGTCCGCGG ATACAAACTTTTCAGGGAGGAGTTGAGCGTTTCACTGGATCA AAATGTCACCTCAAGGTCTTAAAAAGTTAGGTCAAGATCAGGTTG CAACCCTGTTTATGACGTTGTTGGCCGTTTTCGCGCTTTTGTCTA TCTCGTTATAGTGGTCAATCTGATCTGATGGTGGGTTCTCCGAT CGCTAATCGTAATCAAGCAGCGATCGAACCTTTAATTTGGCTTT TTGCTAACACTTTGGCTTTAAGAAATTAATTTATCA
SEQ ID	ACATACACTGAATTAACCATCGCGCTAATCAGTTAGCCCATTA	
No. 38	TTTACAAACTTTAGGCGTGGGAGCAGAAGTCTTAGTCGGTATTT CCCTAGAACGTTCTTTAGAGATGATTATCGGCTTATTAGGGATT CTCAAGGTAGGTGGTCTTATCTTCTCTTGTATCCAGACTATCC CACTGAGCGTCTTTCAGTTGATGTTAGAAGACAGTCAAGTTCTT TTTTGATTACCCACAGTCTTTTATTAGCAAAATGGCTCCCTCT CAAGCAACTCTGATTTGTTTAGATCATATCCAAGAGCAGATTTT TCAATATCTCCAGATAAATCTTCAATGTCAGTTAACTCTCGCCA ATTTAGCTAACGTTATTTATACCTCTGGCTTACGGGTAAGCCT AAAGGGGTGATGGTTGAAACATAAAGGTTTAACTTAGCTCT TGCTCAAAATCAATCTTTTGCAGTCAACCATAAACAGTCTGGTGC TGCAATTTGCTCTTTTAGTTTGTAGTCTGTATTTAGAAAT TTGATGACCTTTGGTCTCGGAGCGAGCTTTATCTTGCAAAAA AGATGCTTTTATGGCCAGGTCAGCCATTAATTTGAACGGTTAGTAA AGAATGGAATTACTCATGTGACTTTGCCGCTTTCAGCTTTAGTG GTTTTTACCCAGGAACCGTTACGCACTTAGAACCTTAAATTTG GGCGGTGAGGCTTGTCTTCTTGTAGTTAGTGAACAAATGGTCAA TCGATAGAACTTTTCAATGCTATGGGCCAACGGAAGCGAGT GTTTGTGCCACTTTGGACAATGTTATCAAGATGATTTAAAGGT GACGATTTGTAAGGCTATGCCAATGTCCAAATTTATATTTAG ATGCCTTTTTACAGCCGGTCCCGTGGGAGTGTGAGGAGGTTA TACATTTGGTGGAGTTGGGGTGGCAAGGGGCTATTTAAATCGTCC TGAATTAACCCAAAGAAAATTTATGCTAATCCTTTTAGTAACG ACCCAGATTTCTCGGCTTATAAACTGGCGACTTAGCGGTTAT TTACCCGATGCGTAATTAAGAAATTTAGGACGCATGACAATCA GGTAAAAATTCGGCGTTTTCGCATGAGTTAGGAGAAATGAAG CGGTTCTGAGTCAATGTCCCGATGTGCAAAATACGGCGGTG	
SEQ ID	GAAATCTGGCTCAAATATGGGGCAAGTCTCAAGATAGAAA	
No. 39	AGTCAGCAGAGAAGAT AATTTCTTGAATTTGGGGGGCATTCCCTTTTAGCTACCCAGGT AATGTCCCGTCTGGCGTGAACCTTTCAAGTCGAATTAACCTTGC GTAGTCTCTTTACCGCTCCACATTTGCTGAATTTGGCCCTAACA ATT	
T3		
nucl		
acid		
SEQ ID	AACGACAGTGTAACTCCCGTTATCTTTTGGCTCAACAACGTTT	
No. 40	ATGGTTTCTGGATCAATTAGAACCTTAACAGCGCTTTTATCATG TAGGGGGAGCCGTTAAGACTAGAAGGAACATTAATATTAAGTCC TTAGAGCAAAGCTTAAAGAAAATTAATTAATCGTATGAAGCTTT ACGCACAAATTTTAAACAGATGATGGTCAAGCCACTCAAAATTA TTACCCCTACTATTAATTTGGCGATGTCTGTGTTGATGTTCAA AATTTAACCGCACTCAATCTCGGAAATTTGCGGAAGCTGAAA GCCCTTAACTCTGGCTCAAGATTGCTTATTTCTGCTACTTTTAT TCGTGCGATCACCGCTAGAATAACAATATGCTGATTTTGGCGTCC CATATTTGTTAGCGATGGTGGTCAATTTGGAGTATTTTTCAGA ACTAATCATCTTTACGCTGTCTATAATCAGGGTTTACCCTCAT CTTTAACCGCTATTAATAACAATATGCTGATTTTGGCGTCCG CAACGGAATTTGTTACAAGGTGAAATTTAAGTAATCAATTTGAA TTATTTGGCGCAACAATTAGCAATGCTCTGCTTTTTTACCTT TACCGACAGATAGACCTAGGCCCGCAATCCAAACTTTTATTGGT TCTCATCAAGAAATTTAACTTTCTCAGCCATTAAAGCCAAAAT GAATCAACTAAGTCAGAAGCATGGAGTGACTTTATTTATGACTC TCTTGGCTGCTTTTGTCACTTACTTTACCGTTATACAGGACAA GCAGATATTTAGTTGGTCTCTCTATGCTAACCGTAAATCGTAA GGAAATTTAGGGGATTAATCGGCTTTTTTGTTAATACATTAGTTC TGAGATTTAGTTAGAT	
SEQ ID	ACCTATGCTGAATTAATCAAGCTAATCAGTTAGTCCATTA	
No. 41	CTTACAACTTTAGGAATTTGGGCCAGAGGCTTAGTCTGCTATTT CAGTAGAACGTTCTTTAGAAATGATTATCGGCTTATTAGCCATT CTCAAGGCGTGTGGTCTTATCTCCCTCTTGTCTCTGACTATCC CACTGAGCGTCTTTCAGTTTCAATGTTAGAAGATAGTCAAGCTTCT TTTTGATTACCCACAGTCTTTTATTAGAAAATTTGCCTTCTTCT CAAGCGACTCTAATTTGTTTAGATCACATCCAAGAGCAGATTTT TCAATATCTCCCGTAACTCTTCAAAGTGAAGTTAACTTCCAT ATTTGGCTAACGTTATTTACACCTCTGGCTCTACGGGTAAGCCT AAAGGGGTGATGGTTGAAACATCGGGGCTTAGTTAACTTAGCGAG TTCTCAAAATCAATCTTTTGCAGTCAAAAATAACAGTCTGGTAC TGCAATTTGCTCTCTTTAGTTTGTAGTCTGTATTTAGAAAT	

TABLE 1-continued

		TTGATGACCTTTGGTCTGGAGCGACTCTTATCTTGCTCAAAA AAAATGATTTATTGCCAGGTGACCCATTAAATGGAAAAGGTTAGAAA AGAATAAAATACCCTAGTTACTTTACCCCTTCAGCTTAGCT GTTTACCAAAAAACCGTTACCCAACTTACAACTTTAATTGT GGCGGGTGGAGCTTGTCTCTGGATTTAGTCAACAATGGTTCAG TCGGTAGAAAATTTTTCAATGCTATGGCCCGCAAGAAAGAGT GTTTGTGCCACGATGGCAATGTTATCAAGATGATTTAAAGGT CACGATTTGGTAAGGCGATCGCTAATGTCCAATTTATATTTTGG ATGCCTTTTTACAACAGTACCCTACCGAGTCAACATGGGAATTA TACATTTGGTGGAGTTCGAGTTGGAGGGGTTATCTAAATCGTCC TGAATTAACGGCGGAAAGATTTATTTCCATCTTTTGTATCCCC CCCTAACCCCTTTAAAAGGGGGAGATAAGAGCTATGAAACT TTTTAAAAGGGGGAAGCAACCA TCAAACTCTATAAAACGGGAGATTTAGCTCGTTATTTACCCGA TGGCAATATTGAATTTTAGGACGCTATGACCAATCAGGTAAGAA TTCCGCGTTTTCCGATTTAGTTAGGAGAAATGAAGCGGTTCTG AGTCAATGTCCCGATGTGCAAAATACGGCGGTG
5		
10		
15		
SEQ ID	TTACAATTAGCTCAAATCTGGTTCAGAGATTTTAGGCATTAATA	
No. 42	TATTTGGTATTGAGAAAATCTTTTGAATTTAGGCGGTCATCTT TATTAGCAGTCACTGATCAATCGATTTTGAACAAAATTTAGT AAACGTTTACCATTAAACAGTCTTTTCAAAATGGAACCATAGC AAGTCTAGCTCAATTAAGT	
T4		
nucl		
acid		
20		
25		
30		
35		
40		
45		
50		
55		
SEQ ID	AGCGGGTCTCAAGACCAAAAAACGATACAGTTAGCCCTCTACTA	
No. 44	CTTTGGTAGCTATGAAGCGAATTTAACCCGAATAAATAAAT TACTGTTTGAAGGAGCTAAAATTTGGCGATCGCGCTGGTTTTACG GCCCTTTGGATTCTGAACTGATTTTCCAGCTTTTGGTGGGTT TTCTCCCAATCTTTCGTTTGGCGGGGCTTTAGCACGGGAAA CCAAACAGATTCAGTGCATCAGGACAGTGTGGTTTTTACCCGTA CATAATTCATCCGAGTCCGCAAGAAATGGCCAGTGGTGGACAA TCTTTCCAGGGCCGCTTGGTATTGCTTTTGCATCGGTTGGC ATCCCAAGGATTTTGTCTTGGCTCCCAAGTCTTTGGCCAACT CGGGAATTTGATGTTTCAAGAAATTTGAAACCGTCCAGGATTTG GCGAGGGGAAGCGATCACCGTCCAGAGCGAAAGGGTCAA AGGGTAGAGGTTAAAACCTATCCCAACCGATGCACTCCAGTT ACCCAGCTGGATTAATTTGCAATAAATCCGATGCTATGATCA GAGCAGGGGCGATCGGTGCTAATATCTTACCAATCTGATGGGG CAAAGCGTGGAAAGTTTAGCCCGTAAATTTGCGCTATATCGTCA ATCTTTGGCAGAGCATGGTTATGATCCCGCTCGGGAACGGTGA CAGTTCTCTGCATACTTTTGTGGCAAGGATTTAGAACAAAGTT CGAGAACAGGCTCGCCAACTTTGGGCAATCTCACCTCTCT TGTGGACTCTTGCAGAACATGGTCAAGAGCCAGGGCATGAAAG TGGATTTGAACAATAAGAGACGAGATCGGGACTTTCTCTCT GCTTCTGCTATAAACGCTATACAGAAACAGTGTCTTTAATTTGG CACACCCGAATCTGTCGTCAAATTTAGTCAATTTGAGTCCAGTCCA TCGGTGTGGATGAAGTGGCTTGTTTTATTGATTTTGGGGTAGAT GAACAAACAGTTTGGCAATTTACCCTATCTCCAGTCCCTAA AGACTTATATCAA	
SEQ ID	ATTGATCCCCCCTAACCCCTTGATAAGGGGATTTGATCCCCC	
No. 45	CCCTAACCCCTTGATAAGGGGATTTGATCCCCCCTAACCCCT TTGATAAGGGG	
SP 1		
nucl		
acid		
60		
65		
SEQ ID	CCTTATCAAGGGGGTTAGGGGGGATCAATCCCCCTTATCAAGG	
No. 46	GGGGTTAGGGGGGAT CAATCCCTTATCAAGGGGGTTAGGGGGTATCAATCCCTTATCAAGG TCAAGGGGGTTAGGGGGTATCAATCCCTTATCAAGGGGGTT TAGGGGGGATCAATCCCTTATCAAGGGGGTTAGGGGGGAT TTAGGGGGGATCAATCCCTTATCAAGGGGGTTAGGGGGGAT	
SP 2		
nucl		
acid		

TABLE 1-continued

TCAAGTC

SEQ ID CCTGCTTCAGAAATGCGAGAGTGGGTCGAAAACACTGTTAGTCG  
 NO. 47 CATCTTGGGTTTCCAACCGAAGCGGGTTAGAAATGGTTGTG  
 MT GTACAGGTTTGTACTCTCCAGGGTAGCAAGCATTTGTCTTGAA  
 nucl TATTGGGCAACGGATTATTTCCAAAGGGGCGATCCAGTATGTTGA  
 acid ACGGGTTTGCATGCGGTTGAAGGTTTAGAACAGGTTAAATAC  
 GCTGTCAAATGGCAGATAATTTTGAAGGTATTGCCCTACATCAA  
 TTTGATACCGTCTTAAATTCGATTATTCAGTATTTTCCAG  
 TGTGGATTATCTGTTACAGGTGCTTGAAGGGGCGATCAACGTCA  
 TTGGCGAGCGAGGTACAGATTTTGTGCGGGATGTGCGGAGTTTA  
 CCCCTATTAGAGCCATATCATGCGGCTGTGCAATTAGCCCAAGC  
 TTCTGACTCGAAAACCTGTTGAAACAATGGCAACAACAGGTGCGTC  
 AAAGTGTAGCAGGTGAAGAAGAACTGGTCAATTGATCCACATTG  
 TTCCTGGCTTTAAACAACATTTTCCGCAAAATTAGCTGGGTAGA  
 AATTC AACCGAAGCGGGGTGTGGCTCACAAATGAGTTAACTCAAT  
 TTCGCTATGATGTCACTCTCCATTTAGAGACTATCAATAATCAA  
 GCATTATTGAGCGGCAATCCAACGGTAATTACCTGGTTAAATG  
 GCAACTTGCCAACTGCTTTAACACAAAATAAAGATAAAATTTAT  
 TAACAGACAAAACCTGAATTTGGGGAAATTCGTGGTATTCCTAAT  
 CAGCGAGTTGAAGAGGCTCTAAAAATTTGGGAATGGGTGGAAAA  
 TGCCCTGATGTTGAAACGGTTGAACAAC TCAAAAAACTTCTCA  
 AACAAACAGTAGATACTGGTATTAATCCTGAAACAGGTTTGGCAA  
 TTAGCTGAGTCTCTCGGTTACACCGCTCACCTTAGTTGGTGGGA  
 AAGTAGTCAAGACGGTTCCTTTGATGTCAATTTTTCAGCGGAAT  
 CAGAAGCGGAGGACTCAAAAAATTAACCCTTTCAAACCTTGCT  
 TTCTGGGATGAAAAACCTTTAAATAAAGCCCTGGAGTGACTA  
 TACTAACAAACCTCTGCGCGGTAAAGTAGTCAAAAAATTAATTC  
 CT

SEQ ID ATGACAATATATGGCAAATCTATGTCTCATTACTATGATCTAGT  
 NO. 48 GGTAGGACATAAAGGTTATAACAAGATTACGCCACTGAAGTAG  
 MT 2 AATTCATTACAAATTTAGTTGAGACTTACACAAC TGAAGCCAAA  
 nucl TCTATCCTATACTTGGCTGTGGTACGGGTTATCATGCCGCTCT  
 acid TTTAGCACAGAAAGGTTATCTGTACATGGTGTGATCTCAGTG  
 CTGAAATGTAGAGCAGGCTAAAAC TCGCATGAAAGTAGAAACA  
 ATAGCTTCTAATCTGAGTTTTCTCAAGGAAATTTTGTGAAAT  
 CCGTTTAAATCGTCAAGTTAATGTGTTCTTGCTCTATTTTCATG  
 TGGTTAACTATCAAACGACCAATCAAATTTACTGGCAACGTTT  
 GCAACGGTTAAAACCATTTAAAAGCTGGGGGATTTTATTTG  
 TGATGTCTCTATGGGCTTACGTA CTGGGGAAATTTAAGAGTC  
 GGCCCTACGGCATCAATATGGGTTTAGAGGATAATTCCAATGGT  
 AACGAAGTAACCTATATAGTGAAC TAAATTTTAAACCCATGA  
 AAAATAGTGGAAAGTTACTCACAAATTTATGGGTAACAAATCAAG  
 AAAATCAACTCTAGAGAATTCACGGGAAACACATCTTCAGCGC  
 TATCTTTTCAAGCCTGAAGTTGAATTTGGCTGATGCTTGTGA  
 ACTAAGTGTCTTGTAGCGATGCGCTGGCTTGAACAACGTCCTT  
 TGACAAACATTCCTTGTCTTCAGTTTGTGTTTATTGGGCAT  
 AAAACAACCCATTACGCTTAA

SEQ ID CCGACCTGTGATAAAACAATTC  
 NO. 49  
 Primer  
 A

SEQ ID CKNCCDGTDAARAANARYTC  
 NO. 50  
 Primer  
 B

SEQ ID TTCAATATCCTGGGGATA  
 NO. 51  
 Primer  
 C

SEQ ID YTCDATRTCYTGNNGRTA  
 NO. 52  
 Primer  
 D

SEQ ID CGTTGGTTACAGGCCCTTTCT  
 NO. 53  
 Primer  
 E

TABLE 1-continued

SEQ ID MGNTGGYTNCARGCNYTNWS  
 NO. 54  
 Primer  
 F

SEQ ID TTAGACTTAAGCCATTGG  
 NO. 55  
 Primer  
 G

SEQ ID YTN GAYYTNWSNCA YTTG  
 NO. 56  
 Primer  
 H

SEQ ID CATAGAAGAATCGAGACCATATTC  
 NO. 57  
 Primer  
 I

SEQ ID CATNSWNSWRCTCNARNCCRTAYTC  
 NO. 58  
 Primer  
 J

SEQ ID MTTQTASSANALASFNQFLRDVKAIAPQYWYPTVSNKRSFSEVI  
 NO. 59 RSWGMLSLLIPLIVGLVAVTAFNSFVNRLDIVIIQEKDASQFA  
 ABC STLTVYAI GLICVTL LAGFTKDIRKKIALDWYQWLNTQIVEKYF  
 Trans- SNRAYKINFQSDIDNPDQRLAQEIEPIATNAISFSATFLEKSL  
 porter EMLTFLVVVVISRQIAIPLMFYTIIGNFIAAYLNLQELSKINQA  
 QLQSKADYNYALTHVRTHAESIAPFRGEKEEQNIIQRRFQEVIN  
 DTKNKINWEKNEIFSRGYRSVIOQFFPLVLGPLYIKGEIDYQ  
 VEQASLASFMFASALGELITEFGTSGRFSYVERLNEFSNALET  
 VTKQAENVSTITTI EENHFAFEHVTLETPDYEKVIVEDLSLTVQ  
 KGEGLLIVGPSGRGKSSLLRLAAGLWLNAGTGLRVRPPEEILFL  
 PQRPYII LGTLREQLLYPLTNSEMSNTLQAVLQVNLQVNLNR  
 VDDFDSEKWPENILSLGEQRLAPARLLVNSPSPFTILDEATSAL  
 DLTNEGLIYEQLQTRKTTFISVGHRESLFNYHQWVLELSADS SW  
 ELLSVQDYRLKKAGEMFTNASSNNSITPDITIDNGSEPEIVYSL  
 EGFSHQEMKLLTDLSSLIRS KASRGKVI TAKDGF TYLYDKNPQ  
 ILKWL R

SEQ ID ATGACAACCCAAACAGCTTCTAGTGCCAAATGCCCTTGCTTCCTT  
 NO. 60 TAAACCAATTTTAAAGGATGTAAGGCGATCGCCCAACCTATT  
 ABC GGTATCCACTGTATCAATAAAAAGAGCTTTTCTGAGGTTATT  
 Trans- CGTTCCTGGGAAATGCTATCACTGCTTATCTTTTGTATGTTGGG  
 porter ATTAGTCGCCGTACGGCTTTTAAATAGTTTGTAACTGCTGCT  
 Nucl TAAATGATGTCATTATTCAGAAAAGATGCTCTCAATTTGCC  
 acid AGTACATTAACGTCTATGCGATCGGATTAATCTGTGTAACGCT  
 GCTGGCAGGGTTCACTAAAGATATTTCGCAAAAAAATTCGCCCTAG  
 ATGGTATCAATGGTTAAACACCCAGATTGTAGAGAAATATTTT  
 AGTAATCGTGCCATTATAAAAATTAAC TTTCAATCTGACATTGA  
 TAAACCCGATCAACGCTTAGCCAGGAAATGAACCGATCGCCA  
 CAAACGCATTAGTTTCTCGGCCACTTTTGGAAAAAAGTTTG  
 GAAATGCTAACTTTTGTAGTGGTGGTCAATTTCTCGACA  
 GATTGCTATTCGCTAATGTTTACACGATTATCGGTAATTTTA  
 TTGCCGCTATCTAATCAAGAATTAAGCAAGATCAATCAGGCA  
 CAACTGCAATCAAAGCAGATTATAACTATGCTTAAACCCATGT  
 TCGGACTCATGCGGAATCTATTGCTTTTTCGGGGAGAAAAAG  
 AGGAACAAAATATTATTAGCGACGTTTTCAGGAAGTTATCAAT  
 GATACGAAAAATAAAATTAAC TGGGAAAAAGGAAATGAAATTTT  
 TAGT  
 CGGGCTATCGTTCGCTCATTGCTTTTTCCTTTTGTAGTCTT  
 TGGCCCTTTGTATATAAAGGAGAAATGATTATGGACAAGTTG  
 AGCAAGCTTCATTAGCTAGTTTATGTTGATCGCCCTGGGA  
 GAATTAATTACAGAAATTTGGTACTTCAGGACGTTTCTTAGTTA  
 TGTAGAACGTTTAAATGAATTTCTAATGCTTTAGAACTGTGA  
 CTAACAAGCCGAGAATGTGAGCAAAATTAACAACATAGAAAGAA  
 AATCATTTGCTTTGAAACGCTCACCC TAGAAACCCCTGACTA  
 TGAAAAGGTGATTGTTGAGGATTTATCTCTTACTGTTCAAAAAG  
 GTGAAGGATTTATGATGTGCGGGCCAGTGGTTCGAGGTAAGAGT  
 TCTTTATTAAGGGCGATCGCCGGTTTATGGAATGCTGGCACTGG  
 GCGTTTGTAGTGCCTCCCTTAGAAGAAATTTCTTTTGGCCCC  
 AACCTCCACATTTTGGGAACCTTACGCGAAACATTTGCTG  
 TATCCTTAACCAATAGTGAGATGAGCAATACCGAACTCAAGC  
 AGTATTACAACAAGTCAATTTGCAAAATGTGCTAAATCGGGTGG  
 ATGACTTTGACTCCGAAAAACCTGGGAAAAACATTTCTCTCCCTC  
 GGTGAACACACGCTTAGCCTTTGCTCGATGTTTAGTGAATTC

TABLE 1-continued

	TCCGAGTTTACCATTTTAGATGAGGCGACCAGTGCCTTAGATT TAAACAAATGAGGGGATTTTATACGAGCAATTACAAACTCGCAAG ACAACCTTTATTAGTGTGGGTCATCGAGAAAGTTGTTAATTA CCATCAATGGGTTTTAGAACTTCTGCTGACTCTAGTTGGGAAC TCTTAAGCGTTCAAGATTATCGCCTTAAAAAGCGGAGAAAATG TTTACTAATGCTTCGAGTAAACAATCCATAACACCCGATATTAC TATCGATAATGGATCAGAACCAGAAATAGTCTATTCTCTTGAAG GATTTTCCCATCAGGAAATGAACTATTAACAGACCTATCACCTC TCTAGCATTGCGAGTAAAGCAGTTCGAGGGAAGGTGATTACAGC CAAGGATGGTTTTACCTACCTTTATGACAAAAATCCTCAGATAT TAAAGTGGCTCAGAACTTAA	5 10
		15
In one embodiment the entire gene cluster is transformed and expressed in a heterologous system. SEQ ID NO. 61 encompasses the genes of said cluster.		
1-27260	ATGACTATTAECTATGGTGATCTGCAAGAACCCCTTTA	20
Microginin-Cluster	ATAAATTCTCAACCCCTAGTTGAATTACTCCGTTATCG GGCAAGCAGTCAACCGGAACGCCCTCGCCTATATTTTT	
1-1743	CTGCCGAGACGGAGAAATCGAAGAAGCTCGTTTAACT	25
Adenylation-Protein (A*)	ATGGGAACTGGATCAAAGGCTAGGGCGATCGCCGC TTATCTACAATCCTTAGAAGCCGAGGGCGAAAGGGGT	
1892-2158	TTACTGCTCTATCCCCAGGACTAGATTTTATTTTCAG	30
Acyl-Carrier-Protein (ACP)	CTTTTTTTGGTTGTTTATATGCGGGAGTCGTTGCCAT TCCCCTATCCACCCGACGGAATCAAACCTTTTTG	
2204-3016	CGTTTACAGGCGATTATTGCCGATTCTCAAGCCCGAT	35
Methyltransferase (MT)	TTACCTTCACCAATGCCGCTCTATTTCCAGTTTAAA AAACCAATGGGCTAAAGACCCGTAATTAGGAGCAATG	
3464-13123	GAATGGATTGTTACCGATGAAATTGACCATCACCTCA	40
PKS/NRPS (KS-AT-ACP-AMT-MO-C-A-T)	GGGAGGATTGGCTAGAACCAACCCCTCGAAAAAACAG TCTCGCTTTTCTACAATACACCTCTGGTTCAACGGGA ACTCAAAGGGAGTAATGGTCAGTACCATAATTTGT	
13120-17832	TGATTAATTCAGCCGATTTAGATCGTGGTTGGGGCCA	45
NRPS 2 (C-A-Mt-T)	TGATCAAGATAGCGTAATGGTCACTGGCTACCGACC TTCCATGATATGGGCTGATTTATGGGTTATTTCAGC	
17836-25194	CTTTGTACAAAGGATTTCTTTGTTACATGATGTCCCC	50
NRPS 3 (C-A-T-C-A-T)	TGCCAGCTTTATGGAACGACCGTTACGTTGGTTACAG GCCCTTCTGATAAAAAAGCAACCCATAGTGCGCCCC	
25257-27260	CCAACTTTGCCCTACGATCTTTGTGTGGGAAAATTCC	55
ABC-Transporter (ABC)	CCCTGAAAAACGGGCTACGTTAGACTTAAGCCATTGG AGGTACTTAAAAAGTTTGGCGAGGCTTTTCAAGTTTC TGGTTTCAAAGCCACAGCCCTTTGTCTGGCTACGGT TTAGCAGAAGCCACCCCTGAAAGTTACGGCGGTTAGTT ATGACAGTCCCCCTTACTTTTATCCCGTTACGGCTAA TGCTTTAGAAAAAATAAGATTGTGGGAGCCACTGAA	60 65

-continued

ACCGATACCAATGTGCAGACCCCTCGTGGGCTGCGGCT GGACAACGATTGATACTCAAATCGTCATTGTCAATCC TGAAACCCCTGAAACCTTGCTCCCTGAAATGTGCGGC GAAATTTGGGTATCAGGTTCAACAATCGCCCAAGGCT ATTGGGGAAAACCTCAAGAGACTCAGGAAACCTTTCA AGCTTATTTGGCAGATACAGGAGCCGGGCTTTTCTG CGAACAGGAGACTTGGGCTTCATTAAGATGGTGAAT TGTTTATCACAGGTCGGCTCAAGGAAATTATTCTGAT TCGAGGACGCAATAATTATCCCCAGGATATTGAATTA ACCGTCCAAAATAGTCATCCCGCTCTGCGTCCCAGTT GTGGGGCTGCTTTTACCCTTGAAAAAAGGGCGAAGA AAAGCTCGTGGTCGTTCAAGGAGTGGAGCGCACCTGG CTCCGTAAAGTAGATATAGATGAGGTA AAAAGAGCCA TTCGTAAGCTGTTGTC CAGGAATATGATTTACAGGT TTATGCGATCGCCTGATCAGGACTGGCAGTTTACCA AAAACCTCTAGCGGTA AAAATCAGCGCTGATAGTGT GGGCCAAAATTTT TAGAGGGAAGCCTGGAATTTTGGG CTAAGAAAATTTCTCGATCGGCCTTAATGTGTTAAA TTCGATGTCGATTGAAACTTCGACCAATCTTTCTC TCCCCTTAAGTCCATGTCCTGGATTTGAAAATTCCT TAACTTTAACTACATTTCTCAAGAAAAGCAAATTGAA TCTAATGTCCACAGAAATCCCAAACGACAAAAACAA CCGACCCTAACGAAAATTCAAAACCTGGTTAGTGGCTT ACATGACAGAGATGATGGAAGTGGACGAAGATGAGAT TGATCTGAGCGTCCCTTTGATGAATATGGTCTCGAT TCTTCTATGGCAGTTGCTTTGATCGCTGATCTAGAGG ATTGGTTACGACGAGATTTACATCGCACCCCTGATCTA CGATTATCCAACTTAGAAAAGTTGGCTAAACAGGTT AGTGAACCCCTGACATTTTATAAAGTTTGTGCTTAAA AATTTTGAGGAAGTTCTAAAATGACAAAATTATGGCAA ATCTATGCTCATTACTATGATCTAGTGGTAGGACAT AAAGGTTATAACAAAGATTACGCCACTGAAGTAGAAT TCATTCACAATTTAGTTGAGACTTACACAACCTGAAGC CAAATCTATCTATACTTGGGCTGTGGTACGGGTTAT CATGCCGCTCTTTTAGCACAGAAAGGTTATCTGTAC ATGGTGTGATCTCAGTGTGAAATGTTAGAGCAGGC TAAAACCTCGCATGGAAGATGAAACAATAGCTTCTAAT CTGAGTTTTTCTCAAGGAAATATTTGTGAAATCCGTT TAAATCGTCAGTTAATGTTGTTCTGTCTATTTC TGTGGTTAACTATCAAACGACCAATCAAATTTACTG
---

- continued

GCAACGTTTGCAACGGTTAAAAACCAATTTAAAAGCTG  
 GGGGGATTTTTATTTGTGATGTGTCTATGGGCTTA 5  
 CGTACTGGGGGAATTTAAGAGTCGGCCTACGGCATCA  
 ATATTGCGTTTAGAGGATAATTCCAATGGTAACGAAG  
 TAACCTATATTAGTGAACATAATTTTTAACCCATGA 10  
 AAATATAGTGAAGTTACTCACAATTTATGGGTAACA  
 AATCAAGAAAAACAACCTCTAGAGAATTCACGGGAAA  
 CACATCTTCAGCGCTATCTTTCAAGCCTGAAGTTGA 15  
 ATTGTTGGCTGATGCTTGTGAACATACTGTTCTTGAT  
 GCGATGCCCTGGCTTGAACAACGCTCTTTGACAAACA  
 TTCCFTGCTTCAGTTTGTGTTTATGGGCATAA 20  
 AACAAACCCATTTCAGCTTAAATTCCTGCTAAAAAAATC  
 CAACTTACCTTATTCTCTGAAACACACAAGCCATGA  
 ATACAATTCAGATGCCAAGACCGAAAATTACTCAAT 25  
 CTTAAATCAGTCAATTCGAAGACCTCTCAAACCTGAGT  
 AATATCCTATTACGATAAGATTTGCGTTCCTCTTTG  
 TTTGGAATGTGAGCAGAGGAGTCTCTATATTGGCTAG  
 AGAAATGTTTATGTCAAGAGCATCAGGGCTTCGATGT 30  
 ACAAGTTAAGTATCATCAAAAAATGCTGAAGAATATG  
 TTACGTTTGACCGATAGTTTGGATTATCTATGGCCAG  
 TTAACCGTGAAATGCGGCTCATGAAAGCTGGGGGGTC 35  
 AATTGAACGGGCGATCACCAATAACATAAAGCTTTTT  
 CTTCAATTTAAGAAACTGTAACCGTATTAAATTAGA  
 AAAACCGCAGTGAGGAATTTGAATGGAACCCATCGCA 40  
 ATTATTGGTCTTGCTTGCCGCTTCCAGGGGCTGACA  
 ATCCAGAAGCTTTCTGGCAACTCATGCGAAATGGGGT  
 GGATGCGATCGCCGATATTCCTCCTGAACGTTGGGAT 45  
 ATTGAGCGTTTCTACGATCCCACACCTGCCACTGCCA  
 AGAAGATGTATAGTCGCCAGGGCGGTTTTCTAAAAAA  
 TGTCGATCAATTTGACCCCTCAATTTTTCCGAATTTCT 50  
 CCCC TAGAAGCCACCTATCTAGATCCTCAACAAAGAC  
 TGCTACTGGAAGTCACTGGGAAGCCTTAGAAAATGC  
 TGCCATTGTGCTGAAACCTTAGCTGGTAGCCAATCA 55  
 GGGGTTTTTATTGGTATCAGTGTGATTATCATC  
 GTTTGGCTTATCAAAGTCTACTAACTTGACCCGCTA  
 TGTGGGTACAGGCAACAGCACCAGTATTGCGGCTAAC 60  
 CGTTTATCATACTGTTTGTATTGCGTGGCCCCAGTT  
 TGGCCGTAGATACCGCTTGTCTTCTTCCCTCGTCGC  
 CGTTCACCTGGCCGTGTCAGAGTTGCAAGTCAAGAA 65  
 TCGAACCTCTGCTTAGTGGGGGAGTTAATCTCATTT

- continued

TGTGCGCCAGAGACAACCGTTGTTTTTTCCCAAGCGAG  
 AATGATCGCCCCGACAGTCGTGTGAAAACCTTTGAC  
 GCGAGGGCCGATGGTTATGTGCGCTCGGAAGGCTGTG  
 GAGTAGTCGTAATAAACGCTCTTAGGGATGCCATTCA  
 GGACGGCGATCGCATTTTAGCAGTGATTGAAGGTTCC  
 GCGGTGAATCAGGATGGTTTAAAGTAATGGACTCACGG  
 CCCCTAATGGCCCTGCTCAACAGGCGGTGATTCTGTCA  
 GGCCCTGGCAAATGCCAGGTAAAACCGGCCAGATT  
 AGCTATGTCGAAGCCATGGCACGGGACAGAATTGG  
 GGGATCCGATCGAAGTTAAATCTCTGAAAGCGGTTTT  
 GGGTGAAAAGCGATCGCTCGATCAAACCTGTTGGCTC  
 GGTCTGTGAAAACCAATTTGGTCAATTTAGAAGCGG  
 CGGCGGGAATGGCGGGTCTGATTAAGTCTGTTCTCTG  
 CCTACAACACCAAGAAATTCGCCCTAATCTCCACTTT  
 CAAACCCCTTAATCCCTATATTTCCCTAGCTGACACAG  
 CTTTTGCGATTTCCACTCAGGCTCAACCTGGCGGAC  
 CAAACCCCTAAGTCTGGTGAACCGGTGTCGAACGA  
 CGTTTAGCAGGACTCAGTTCCTTTGGGTTTGGGGGA  
 CAAATCCCATGTGATTTCTCAGCGAAGCCCTGTGAC  
 CGTTAAAAACAATCAACAAAATGGGCAGAAGTTGATA  
 GAACGTCCCCTGGCATTTGCTGACTTTATCTGCCAAGA  
 ATGAAGAAGCCTTAAAAGCCTTAGTCCATTGTTATCA  
 AAAGTATTTAGCTGATCATGAAATTCCTCTCGCT  
 GATGTTTGTGTTTACGGCCAATAGTCGGCGATCGCACT  
 TTAATCATCGTTTAGGAGTAGTGGCTAGAGATCGCTT  
 AGAAATGTTGCAGAAGTTAGAGAATTTAGTAACCAA  
 GAAAGGATGAGAGAACC GAAGATTAACAAAAAAG  
 AAAAACCTAAAATGTTTTTCTATTTGCGGTCAGG  
 TTCTCAATATGTAGGTATGGGTGCTCAACTGTACGAA  
 ACCCAACCCATCTTTCCGCAACCTTGGATCGCTGTG  
 CTGAAATCCTGCGACCCATTTAGATCAACCCCTCTT  
 AGAAATCTTTATCCTGCTGACCCAGAAGCCGAAACA  
 GCGAGTTTTTACCTAGAGCAGACTGCCTATACCCAAC  
 CCCTTTTATTCGCATTCGAGTATGCCCTAGCACAGTT  
 ATGGCGTTCCTGGGGAATAGAACGGCGGCAGTAATT  
 GGTACACAGTGTGCGTGAATATGTGGCGGCCACCGTTG  
 CCGGAGCCTTAAGTCTAGAAGAAGGATTAACGCTAAT  
 TGCCAAACGGGCAAACTGATGCAGTCTCTCCCAAG  
 AATGGGACAATGATCGCGTTTTTTGCCGAGAAAGAGC  
 GGGTTAAGCTGTATTGAGCCTTATAGACTGATGT

27

-continued

AGCGATCGCTGCTGTTAATGGACCAGAAAATTTTGTT  
 ATTTTCAGGAAAAGCGCCGATTATTGCTGAGATTATCA 5  
 TTCATTTAACGGCAGCAGGAATAGAAGTTCGTCCCTCT  
 CAAAGTTTCCCATGCTTTTCACTCGCACCTGTTGGAG  
 CCAATTTTAGATTCCCTTAGAACAGGAAGCTGCTGCTA 10  
 TTTCCTACCAACCCCTGCAAAATCCCTTAGTTGCTAA  
 TTTAACGGGGGAGTTCACCAGAAGGAGCAACGATT  
 GAGGCTCGTTACTGGCGAAATCATGCACGCAACCCCTG 15  
 TACAATTTTATGGGAGTATCCAAACGCTGATCGAGCA  
 GAAATTCAGTCTTTTTTTAGAAAGTTAGCCCTAAACCG  
 ACTTTATCTCGATTGGGTCAACAATGTTGTCCAGAAA 20  
 GATCGACCACTTGGCTATTTTCCCTCGCCCTCCTCA  
 AGAAGAAGAACAAGCCCTACTAAATAGTTTGGCGATT  
 CTCTATGATTTCCAAGGAGCCGAAATAAACTGGGAAG 25  
 GGTTTAATCAAAATATATCCCCACCATTTACTGGCTCT  
 ACCGACCTATCCTTTTCAACGTCAACGCTATTGGCTT  
 GAAACCGGTAACCGACTTCTGAAGAAACAACCATGA 30  
 CGACCAATGCCACTAATGTCCAAGCTATCTCCAGCCA  
 TCAAAAACAACAGGAGATTCTAATCACATTGCAAAAC  
 CTAGTGGGAAATTTACTGCAATGTCCCCTGCTGATG 35  
 TCAATGTTTCATACACCTTTCCCTGGAGATGGGGCAGA  
 TTCCATTGTCATGTTGAGGCGGTCAGACGATTGAG  
 AATACCTATAACGTTAAAATTGCTATGCGTCAGTTAT 40  
 TTGAGGAGTTATCTACTTTAGATGCTTTAGCTACTTA  
 TTTAGCTCAAAATCCGGCTACTGATTGCCAAACTGCT  
 CAAATTAATACCGAGGTGTTTTCTGCGCCCATTCGCT 45  
 GCTCAAATAACCGATCGCCCAATGTCGTGCTGAGTTC  
 TAATACCAACGGCTTCAACGTCAACAGCTTCTCCA  
 GGTTTTTCGGGCATCGCCCCCTTCAGGAATGGGAG  
 GAGCAGGGGAAATGGGAGGAGTTGAAGTGCCTCAAGT 50  
 TTCTGTGCCACAACAGTGCAGTAAACAGCCTCAGGT  
 TCAACCGTTTCTAGTCTGCCCTGGAAAACATTATGG  
 GTCAACAGTTACAACCTGATGGCCAAACAGTTAGAAGT 55  
 CTTGCAACCGCAATTTTGCCTCCGACTCCCCGA  
 ACCACAGAAAATTTCCCATCTTCCGTGCTCAAAATA  
 GGTCAAACGGACTTACACAACAGTTAATTTCCCCCA 60  
 GCAATTAGCGGCGAACCTTAGAGCCAATAGCCAGTCGC  
 ACCCGTCAAACAGCAATCAAGCTTCTGCTCCTAAAC  
 CGACAGTAACAGCCACTCCCTGGGGCCGAAAAAAC 65  
 ACCACAGGTGGATTCACTCCCAACAACAGCAACAT

28

-continued

CTAGAGGCATTAATGTCTCGCTTTACGGAACGTACCA  
 AAACCTCTAAGCAAATTTGTGCAAAGCGATCGCCTGCG  
 TTTAGCAGATAGTCGAGCCTCGGTCCGATTCCGTATG  
 TCTATTAAGAGATGCTTTATCCATTTGTGGCCCAAC  
 GTTCTCAAGGATCAAGAATTTGGGATGTGGACGGTAA  
 TGAATATATTGATATGACGATGGGGCAAGGGGTAAACG  
 CTGTTTGGGCATCAACCAGACTTTCATTATGTCGGCCC  
 TACAAAGCCAACTCACTGAAGGCATTATCTCAATCC  
 GCGATCGCCAAATTTGGGAGAAGTGGCCGCCTTAATT  
 TGTGAACCTAACAGGAGCCGAACGAGCTTGTTTTTGCA  
 ACTCTGGAACCGAAGCCGTAATGGCCGCTATTCTGAT  
 CGCCAGGGCAACAACAGGTTCGGAGTAAATTTGCCCTC  
 TTTGAAGGCTCCTATCATGGACATGCGGACGGAACCC  
 TTTTTAGGAACCAAATTTATGATAACCAACTCCACTC  
 TTTTCCCTAGCTCTAGGCGTTCCCCCAGCCTTAGT  
 TCCGATGTGGTGGTATTGGACTATGGCAGTCCGGAAG  
 CTCTGAACTATTTACAACCCAGGGGAGGATTTAGC  
 GGCGGTCTTAGTAGAACCAATTCAAAGTGGCAATCCT  
 CTACTCCAACCCCAACATTTCTCAAAGTCTGCGAC  
 AAATACCAGTCAAATGGGCATTGCCCTGATTTTTGA  
 TGAAATGATTACGGGTTTTTCGATCGCACCCAGGGGGA  
 GCGCAAGCTTTATTTGGAGTACAGGCGGATATTGCCA  
 CCTATGGCAAAGTAGTTGCGGGAGGAATGCCCATTTGG  
 AGTTATTGCAGGTAAGGCCATTATCTGGACAGCATT  
 GACGGGGGAATGTGGCGTTATGGCGATAAATCCTATC  
 CTGGGGTGGACAGAACCTTTTTTGGGGGAACCTTTAA  
 TCAGCATCCGTTAGCAATGGTAGCGGCTAGGGCTGTC  
 CTGACCCATTTAAAGGAGCAGGGGCCAGGCTGCAAC  
 AACAATTAACCTGAACGCACTGCGGCCTTAGCCGATAC  
 ACTGAATCATTATTTTCAAGCCGAAGAAGTTCCTATT  
 AAAATCGAACAGTTTAGTTCCTTTCTCCGGTTTGCCC  
 TCTCTGGCAATTTGATTTACTTTTCTATCACATGGT  
 AGAAAAAGGTATTTATGTCTGGGAATGGCGTAAACAT  
 TTTCTTTCAACCGCCATACGGAAGCCGATCTTGCCC  
 AATTGTCCAAGCGGTTAAGGATAGCATCACAGAATT  
 GCCTCAGGGAGGTTTTATCCCCGAAAAAGCCTTCC  
 TGGCCAGTGCCAACGCCTCAAATTTGATCCCCCCTAA  
 CCCCCCTTGATAAGGGGATTGATCCCCCCTAACCCCT  
 CCTTGATAAGGGGATTGATCCCCCCTAACCCCTT  
 GATAAGGGGGAGATGTTGATGTCGCGCTTGATAAGG

- continued

GAGGAAATTCATCTCTGTTAGGGACAGTAAGTTAGG  
 GAAAGGGAGCGGGTCTCAAGACCAAAAAACGATACAG 5  
 TTTAGCCTCTACTACTTTGGTAGCTATGAAGCGGAAT  
 TTAACCCGAATAAATAAATACTTACTGTTTGAAGGAGC  
 TAAATTTGGCGATCGCGCTGGTTTTACGGCCCTTTGG 10  
 ATTCTGAACGTCATTTCCACGCTTTTGGTGGTTTTT  
 CTCCAATCCTTCGGTTTTGGCGCGGCTTTAGCACG  
 GGAAACCAACAGATTCAACTGCGATCAGGCAGTGTG 15  
 GTTTTACCGCTACATAATTCATCCGAGTCGCCGAAG  
 AATGGGCAGTGGTGGACAATCTTCCACGGGCCGCGT  
 TGGTATTGCTTTTGCATCGGGTTGGCATCCCCAGGAT 20  
 TTTGTCTTGGCTCCCAGTCCCTTGGCCAAACATCGGG  
 AATTGATGTTCCAAGAAATGAAACCGTCCAGAACT  
 TTGGCGAGGGGGAAGCGATCACCGTGCCAGACGGAAAG 25  
 GGTCAAAGGGTAGAGGTTAAAACCTATCCCCAACCGA  
 TGCAGTCCAGTTACCCAGCTGGATTACTATTGTCAA  
 TAATCCCGATACCTATATCAGAGCAGGGCGATCCGGT 30  
 GCTAATATCCTTACCAATCTGATGGGGCAAAGCGTGG  
 AAGATTTAGCCGTAATATGCGCTATATCGTCAATC  
 TTTGGCAGAGCATGGTTATGATCCCGCTCGGGAACG 35  
 GTGACAGTTCCTCGCATACTTTTGTGGCAAGGATT  
 TAGAACAAGTTCGAGAACAGGCTCGCCAACCCCTTGG  
 GCAATACCTCACCTCCTCTGTGCGACTCTTGCAGAAC 40  
 ATGGTCAAGAGCCAGGGCATGAAAGTGGATTTGAAAC  
 AATTAAGAGACGAAGATCGGGACTTTCTCCTCGCTTC  
 TGCCATAAAACGCTATACAGAAACAGTCTTTAATT  
 GGCACACCCGAATCCTGTCTCAAATATTGATCATT 45  
 TGCAGTCCATCGGTGGATGAAGTGGCTTGTTTTAT  
 TGATTTGGGGTAGATGAACAAACAGTTTTGGCCAAT  
 TTACCCTATCTCCAGTCCCTAAAAGACTTATATCAAC 50  
 CTCATCTCCCCCTTATCAAGGGGGTTAGGGGGGA  
 TCAATCCCTTATCAAGGGGGTTAGGGGGGATCAA  
 TCCCCTTATCAAGGGGGTTAGGGGGTATCAATCCC 55  
 CTTATCAAGGGGGTTAGGGGGTATCAATCCCCTTA  
 TCAAGGGGGTTAGGGGGGATCAATCCCTTATCAA  
 GGAGAGTTAGGGGGGATCAATCCCTTATCAAGGGG 60  
 GGTTAGGGGGGATCAAGTCCCTCTCACGAAGCCCA  
 ACGACAACGTGGATTTTGGCTCAATTAGGAGACAAC  
 GGCTCTGTGGCCTATAACCAATCAGTGACATTGCAAT 65  
 TAAGTGGCCATTAAATCCCGTCGCAATGAATCAAGC

- continued

TATTCAACAAATCAGCGATCGCCATGAAGCGTTACGA  
 ACCAAAATTAATGCCAGGGAGATAGTCAAGAAATCC  
 TGCCCCAGGTCGAAATTAAGTCCCTATCTTAGACTT  
 CAGTCTTGACCAAGCTTCGGCCCAACAGCAAGCAGAA  
 CAATGGTTAAAGGAAGAAAGTGA AAAACCCCTTTGATT  
 TGAGCCAGGGTTCCTCTCGTGCCTGGCATCTACTCAA  
 ATTAGAACCAGAATTACATTTGTTAGTATTAACGGCC  
 CATCACATTATCAGTGACGGTTGGTCAATGGGGGTAA  
 TCCTTCGGGAATTAGGAGAGTTATATTCAGCCAAATG  
 TCAGGGTGTACGGCTAATCTTAAAACCCAAAACAG  
 TTTTCGAGAATTGATTGAATGGCAAAGCCAGCCAAGCC  
 AAGGGGAAGAACTGAAAAACAGCAAGCCTATTGGTT  
 AGCAACCCCTGCGCATCCCCTGTTTTGAATTTACCC  
 ACTGACAAACCTCGTCCAGCTTTACCCAGTTACCAAG  
 CTAATCGTCGAAGTCTAACTTTAGATAGCCAATTTAC  
 AGAAAACTAAAGCAATTTAGTCGTAAACAGGGCTGT  
 ACCTTGCTGATGACCCCTGTTATCGGTTTATAACATTC  
 TCGTTCATCGTTTACCGGGACAGGATGATATTCTGGT  
 GGGTCTGCCAGCCTCTGGACGGGGCTTTTAGATAGT  
 GAAGGTATGGTGGGTTATTGCACCATTTTTACC  
 TTCGAGTCAATTAGCAGGTAATCCCACTTTTGTCTGA  
 ATATCTCAAACAAATGCGGGGGTTTTGTTGTCGGCT  
 TATGAACATCAGGACTATCCCTTTGCTCTTTGCTCA  
 ATCAGTTAGATTTACCGGTAATACCAGTCGCTCTCC  
 TTTAATTGATGTCAGTTCAATTTAGAACCAGTTATT  
 AACCTACCCAAAATGAAAGGATTAGAGATTAGTTTGT  
 TGCCCTCAAAGTGAAGTTTTAAGGATCGAGATTTGCA  
 TTGGAATGTGACAGAAATGGGTGGAGAAGCTCTGATT  
 GATTGTGACTACAATACAGACTTATTTAAGATGAAA  
 CGATTCAGCGTTGGTTAGGCCATTTTCAAACCTTACT  
 TGAGGCAGTTATTAATGATTCGCAACAAAATCTCGCG  
 GAATTACCCCTTATTAAGTTCTGCTGAACGACAACAGT  
 TATTAGTGGATTGGAATCAAACCAAGACCGACTATCC  
 CCAAGATCAGTGTATTATCAATATTGTAAGCGCAA  
 GTTGAACGGACTCCCGATGCGATTGCGGTGGTATTG  
 AAACCTCAACAATTAAGTACAGTGAATTAATTTGTCG  
 AGCCAATCAGTTAGCACATTATTTACAAAATTAGGA  
 GTTGGGCCAGAGGCTTAGTTCGTTATTTGGTTCGAAC  
 GTTCTTTAGAAATGATTGTGCGATTGTTAGGGATTCT  
 CAAGGCTGGGGGAGCCTATGACCTCTTGATCCTGAC

31

- continued

TATCCCCCTGAACGCTCTCAATTTATGTTAGAAGATA  
 GTCAATTTTTCTCCTCTTAACCCAACAGCATTACT 5  
 GGAACTTTTTGCTCAGTCTTCAGAAACGGCTACTCCC  
 AAGATTATTTGTTGGATAGCGACTACCAAATTATTT  
 CCCAGGCAAAGAATATTAATCCCGAAAATTCAGTCAC 10  
 AACGAGTAATCTTGCCATGTAAATTTATACCTCTGGT  
 TCGACAGGTA AACCGAAGGGCGTGATGAATAATCATG  
 TTGCTATTAGTAATAAATGTTATGGGTACAAGACAC 15  
 TTATCCTCTAACCACAGAACTGTATTTTACAAAAA  
 ACTCCCTTTAGTTTTGATGTTTCAGTGTGGAAATAT  
 TCTGGCCCTACTAAACGGAGCGCCTTTGGTTTTTGC 20  
 CAAGCCGAATGGCCATAAAGATGCCAGTTACTTAGTC  
 AATCTGATTCAAGAGCAACAAGTAACAACGCTACATT  
 TTGTGTCTTCTATGCTACAGCTTTTTCTGACAGAAAA 25  
 AGACGTAGAAAAATGTAATAGTCTTAAACGAGTCATT  
 TGTAGTGGTGAAGCCCTTTCTTTAGAGCTTCAAGAAC  
 GTTTTTTTGCTCGTTTAGTCTGTGAATTACACAATCT 30  
 TTATGGACCGACAGAAAGCCGCTATTCATGTCACATTT  
 TGGCAATGTCAATCAGATAGCAATTTGAAAACAGTAC  
 CCATTTGGTCGGCCGATCGCTAATATCCAAATTTACAT 35  
 TTTAGACTCTCATCTTCAGCCAGTACCTATTGGAGTA  
 ATCGGAGAATGACACATTTGGTGGGGTTGGTTTGGCGC  
 GGGGTTATTTAAACAGGCCTGAGTTAACGGCGGAGAA  
 ATTTATTGCAAAATCCGTTTGCCTTCCCTTGATCCCCC 40  
 CTAACCCCTTGTATAAGGGGGGAGATGAGAGCTATA  
 AAACTTTTAAAAGGGGGGAGAGCAACCATCAAGATT  
 GTATAAACCGGAGATTTAGTCTGTTATTTACCCGAT 45  
 GGCAAGATTGAGTATCTAGGGCGCATTGATAATCAGG  
 TAAAAAATTCGCGGTTTCCGGATTGAATTGGGGGAAAT  
 TGAAGCGGTTTTGCTATCCCATCCCAGGTACGAGAA 50  
 GCGGTCGTTTTGGTGAGCGAAAGCGATCGCTCTGAAA  
 ATCGGGCTTTGGTCGCTTATATTGTCCCTAATGATCC  
 TGCTTGACGACTCAATCATTACGAGAGTTTGTAAA 55  
 CGGCAGCTTCTGACTATATGATCCGACTTATTGGC  
 TGATCCTTGACAATTTACCGTTAACCCAGCAATGGCAA  
 AATTGATCGTCGGGCTTTACCGTTACCTAATCCAGAG 60  
 TTAATCGTTCGATAGACTATGTGGCTCCCAAAAATC  
 CTACCCAGGAGCGATCGCCGCTATTTTGGTCAAGT  
 TTTAAAACGGAAAAAGTGGGAATTTATGATAACTTT 65  
 TTTGAGATCGGCGTAATTTCTTTGCAAGCCACTCAAG

32

- continued

TTATTTACGCTTACGAGAAAGTTTTGCCCTAGAGTT  
 GCCCTTGCCTCGCCTGTTTGAACAACCGACTGTGGCG  
 GATTGGCTTTAGCCGTAACGGACATTCATGCCACTT  
 TACAAAAATTACAAACCCCTATTGATGATTTATCAGG  
 CGATCGCGAGGAGATTGAACTATGAAATCTATTGAAA  
 CCTTTTGTGAGATTTAGCCAATCAAGATATTAACCT  
 CTGGATGGACGGCGATCGCTCGGTTGTAATGCACCC  
 CAGGGCCTATTAACCCAGAGATTCAAACAGAACTGA  
 AAAACCGTAAAGCAGAAATCATTCACTTTCTCAATCA  
 ACTGGGTTTCAGAGGAGCAAATTAATCTAGAACGATT  
 CTTCCCATTCCTCGTGTGGCCAATTACCCCTCTCCT  
 TTGCCAGTCGCGACTCTGGTTCTTGATCAATTAGA  
 AGGAGCCACGGGAACCTATAACATGACAGGGGCTTG  
 AGTTTAAGCGGCCTCTTCAGTCTGAAGCCCTCAAAC  
 AAGCCCTAAGAATCATTCAACGCCATGAGCCATT  
 GCGTACCAGTTTCCAATCGGTTGACGGGGTTCCAGTG  
 CAGGTGATTAATCCCTATCCTGTTTGGGAATTAGCGA  
 TGGTTGATTTGACAGGAAGGAGACAGAAAGCAGAAAA  
 ATTGGCCTATCAGGAATCCCAAACCCGTTTGATTTG  
 ACCAATAGTCTTTGTTGAGGGTAACGCTCCTCAAAT  
 TACAGCCAGAAAAGCATATTTTATTAATTAATATGCA  
 CCATATTTATTTCCGATGGCTGGTCAATCGGTGTTTTT  
 GTTCGTGAATTGTCCCATCTCTATAGGGCTTTTGTGG  
 CGGGTAAAGAACCAACTTTACCGATTTTACCAATTCA  
 GTATGCGGATTTTGGCGTTTGGCAGCGAGAGTGGTTA  
 CAGGGTAAGGTTTTAGCGGCTCAATTGGAATATTGGA  
 AGCGACAATTGGCAGATGCTCCTCCTGCTGGAACCT  
 GCCCACTGATCGCCCTCGTCCGCAATCAAACCTTT  
 CAAGGCAAGACAGAAAGATTGAGCTAGATAGGAAAC  
 TGACCCAAGAATTAAGGCATTAAGTCAACAGTCGGG  
 TTGTACTTTATTTATGACTTTGTTGGCCGCTTTTGGG  
 GTGGTTTTATCCGTTATAGTGGCCAGACTGATATCG  
 TCATTGGTTCCGGGATCGCCAACCGTAATCGCCAAGA  
 CATTGAGGGGTTAATTGGCTTTTTTGTTAACACTTTG  
 GCGTTGAGGTTAGATTTATCAGAAAAACCCAGCTTTG  
 CCGCTTTTTTAAAACAAGTACAGGAAGTCACTCAGGA  
 TGCCTATGAGCATCAAGACTTGGCCCTTTGAAATGTTA  
 GTGGAAGAATTACAACCTAGAGCGCAAATAGACCGAA  
 ATCCTTTGGTACAGGTGATGTTTGGCCCTACAAAATGC  
 GGCCAATGAAACCTGGAATTTACCTGGGTTGACCATT

- continued

GAAGAAATGCTCTGGGAACCTTGAACCTGCCCGTTTTG  
 ACCTAGAGGTTCAATTTATCAGAAGTTAACGCCGGCAT 5  
 AGCTGGATTCTGTTGCTACACCATTGATCTATTTGAT  
 GATGCAACGATCGCCCGTCTATTGGAACATTTTCAGA  
 ATCTTCTCAGGGCAATTATTGTTAATCCTCAAGAATC 10  
 GGTAAGTTTATTACCCTTGTGTCAGAACAGGAAGAA  
 AAGCAACTTTTAGTTGATTGGAATCAAACCCAAGCCG  
 ATTATCCCAAGATAAGCTTGTCCATCAGTTATTTGA 15  
 AGTTCAAGCAGCCAGTCAGCCAGAAGCGATCGCTCTA  
 ATCTTTGAAAATCAGGTTTGTGACCTATGGAGAATTAA  
 ACCATCGCGCCAATCAATTAGCTCACTATCTTCAGTC 20  
 GTTAGGAGTCACCAAAGAACAAATCGTCGGGGTTTAT  
 CTGGAACGTTCCCTTGAATGGCGATCGGATTTTAG  
 GTATFCTCAAAGCAGGAGCCGCCTATCTCCCATTTGA 25  
 TCCTGAATATCCCTCAGTACGCACCCAATTTATTTCTC  
 GAAGATACCCAACCTTCGCTTCTCTTAACCTCAGGCAG  
 AACTGGCAGAAAACCTGCCAGACTCAAAACAAAAT 30  
 TATCTGTCTAGATCGGACTGGCCAGAAATTACCTCC  
 CAACCCAGACAAACCTAGACCTAAAGATAGAACCTA  
 ATAACCTAGCCTATTGCATCTATACTTCTGGTTCCAC 35  
 AGGACAACCCAAAGGAGTACTGATTTCCCATCAAGCC  
 CTACTCAACTTAATTTCTGGCATCAACAAGCGTTTG  
 AGATTGGCCCCTTACATAAAGCGACCCAAGTGGCAGG  
 CATTGCTTTCGATGCAACGGTTTGGGAATTGTGGCCC 40  
 TATCTGACCACAGGAGCCTGTATTAATCTGGTTCCCC  
 AAAATATTCTGCTCTCACCGACGATTTACGGGATTG  
 GTTGCTTAACCGAGAAATACCATGAGTTTGTGCCA 45  
 ACTCCTTTAGCTGAAAAATATTATCCTTGATTTGGC  
 CTAACCATCTTGTCTAAAAACCTGTTACTGGGAGG  
 TGACAAAACCTCATTTTTATCCTGCTGCGTCCCTTCCC 50  
 TTTCAGGTCATTAACAACCTATGCCCAACGGAAAATA  
 CAGTGGTTGCGACCTCTGACTGGTCAAATCATCTTC  
 ATCTCATCACTTTGGAACCTCGACTATTGGTCGTCCC 55  
 ATTGCCAACGTCCAAATCTATTTATTAGACCAAACC  
 TACAACCTGTCCCCATTGGTGTACCAGGAGAATTACA  
 TTTAGGTGGGGCGGGTTTAGCGCAGGGCTATCTCAAT 60  
 CGTCTGAGTTAACGGCTGAAAAATTTATTGCCAATC  
 CCTTTGATCCCCCCTAACCCCTTGATAAGGGGGG  
 AGAAGAACCCTCAAACCTCTATAAAAACGGGAGACTTA 65  
 GCCCGTTATTTACCCGATGGCAATGTAGAATTTTGG

- continued

GACGTATTGACAATCAGGTAATAAATTCGGGGTTTTCG  
 CATCGAAACTGGGGAAATCGAAGCCGTTTTAAGTCAA  
 TATTTCTATTAGCTGAAAGTGTAGTCGTTGCCAAGG  
 AAGATAAATACTGGGGATAAACGCCCTCGTGCTTATTT  
 GGTTCGCCCTTGCAAATGAGGCCCTACCAGAGCAA  
 TTAGCCCAATGGCAAAGTGAATACATCAGTGATTGGC  
 AAAGTCTCTATGAAAGAACCTATAGTCAAGGGCAAGA  
 CAGCCTAGCTGATCTCACTTTAATATCACGGGTGG  
 AATAGCAGTTATACTCGTCAACCCCTTCCTGCTTCAG  
 AAATGCGAGAGTGGGTGAAAACACTGTTAGTCGCAT  
 CTTGGCTTTCACCCAGAACCGGTTTAGAAATTGGT  
 TGTGGTACAGGTTTGTACTCTCCAGGGTAGCAAAGC  
 ATTGCTTGAATATGGGCAACGGATATTCCCAAGG  
 GCGGATCCAGTATGTTGAACGGGTTTGAATGCCGTT  
 GAAGGTTTAGAACAGGTTAAATTACGCTGTCAAATGG  
 CAGATAAATTTGAAGGTATTGCCCTACATCAATTTGA  
 TACCGTCGTCTTAAATTCGATTATTCAGTATTTTCCC  
 AGTGTGGATTATCTGTTACAGGTGCTTGAAGGGGCGA  
 TCAACGTCATTGGCGAGCGAGGTGAGATTTTGTCCG  
 GGATGTGCGGAGTTTACCCCTATTAGAGCCATATCAT  
 GCGGCTGTGCAATTAGCCCAAGCTTCTGACTCGAAAA  
 CTGTTGAACAATGGCAACAACAGGTGCGTCAAAGTGT  
 AGCAGGTGAAGAAGAACTGGTCAATGATCCACATG  
 TTCTGGCTTTAAAAACAACATTTCCGCAAAATTAGCT  
 GGGTAGAAATTC AACCGAAACGGGTGTGGCTCACAA  
 TGAGTTAACTCAATTTCTGCTATGATGCTACTCTCCAT  
 TTAGAGACTATCAATAATCAAGCATTATTGAGCGGCA  
 ATCCAACGGTAATTACCTGGTTAAATGGCAACTTGA  
 CCAACTGCTTTAACACAAATTAAGATAAATTTATTA  
 ACAGACAAAACCTGAATTGTGGGGAATTCGTGGTATTC  
 CTAATCAGCGAGTTGAAGAGGCTCTAAAAATTTGGGA  
 ATGGGTGGAAAATGCCCTGATGTTGAAACGGTTGAA  
 CAACTCAAAAACTTCTCAAACAACAAGTAGATACTG  
 GTATTAATCTGAAACAGGTTTGGCAATTAGCTGAGTC  
 TCTCGGTTACACCGCTCACCTTAGTTGGTGGGAAAGT  
 AGTCAAGACGGTTCCTTTGATGTCATTTTCAGCGGA  
 ATTCAGAAGCGGAGGACTCAAAAAATTAACCCTTTC  
 AAAACTTGCTTCTGGGATGAAAACCCCTTTAAAAATA  
 AAGCCCTGGAGTACTATAACAACCCCTCTGCGCG  
 GTAAGTTAGTCCAAAAATTAATCTCAAAGTACGAGA

35

- continued

ATTTCTGCAAGAAAACTACCCAGTTATATGGTTCCC  
 CAGGCGTTTGTGCTGCTTGATTCCCTTCCTTTGACCC 5  
 CCAATGGTAAGGTGGATCGTAAGGCGTTACCTTCTCC  
 TGATGCGGCGACTCGTGATTTAGCGAACAGTTTGTGTC  
 TTACCCCGCAATCCGATTGAAGCTCAACTGACTCAAA 10  
 TTTGGAGTGAAGTTTTGGGACTGGAACGCATTGGCGT  
 TAAGGACAACTTTTTGAATTGGGAGGACATTCCTT  
 TTGGCTACCCAGGTTTTATCAAGAATTAATTCAGCCT 15  
 TTGGACTTGATCTTCTGTGCAAAATATGTTTGAATC  
 ACCAACGATCGCGGCATTGCGGGTTATATCAAGCG  
 GTAGATTGGGTCGCCAGGATCAAGCCGATAGCTCGT 20  
 TAAATCATGAAAACTAGAGTAGTGGAGTTCTAAGT  
 TATGACGAAAAAGATTGTTGAATTTGCTGTATCTA  
 CGGGATTTAGGCATTACTTTAGAAGCTGATGAAAACC 25  
 GCTTACGCTGTGAGCTCCCGAAGGAATTTTGACCCC  
 AGCACTCCGTCAAGAAATTGGCGATCACAACTGGAA  
 TTATTACAATTTTTACAATGGGTCAAACAGTCTAAAA 30  
 GTACCGCTCATTTGCCTATTAACCTGTGCTAGAGA  
 CGGTCAATTACCCCTGTCTTTGCTCAACAACGTTTA  
 TGGTTTTACATTATCTTCCCTGATAGTCTGTTCCCT 35  
 ACAATACCCCTGAAATATTGCAAAATTGATGGGAATCT  
 CAATCTGACTGTGCTAGAGCAGAGTTGGGGGAATTA  
 ATTAACCGCCATGAAATTTTAGAACACATTCCCA 40  
 CTGTTTCAGGGGAACCGATTTCAGAAAAATGCACTTCC  
 TAGTCGTTTTAGTTAAAAGTTGATAATTATCAAGAT  
 TTAGACGAAAAATGAACAATCAGCTAAAATCAACAAG 45  
 TAGCAGAATTGGAAGCAGGACAAGCTTTTGATTTAAC  
 GGTGGGGCCACTGATTAGTTTAAAGCTATTGCAATTG  
 AGTCCCAGAAAGTCGGTGTGCTGTTGAAAATGCACC  
 ATATTATCTATGATGGCTGCTTTTGGGATTCTGAT 50  
 TCGGAATATCGGCTCTATACGAAGCATTTTAAAG  
 AACTTAGCCAATCCTCCTCCTGCTGTCTATTTCAGT  
 ATGCAGATTTTGGGTTTGGCAACGTCAATATCTCTC 55  
 AGGTGAGGTCCTAGATAAACTCAATTATTGGCAA  
 GAACAGTTAGCAACAGTCTCTCCTGTTCTTACTTTAC  
 CAACGGATAGACCCCGTCCGGCGATACAAAATTTTCA 60  
 GGGAGGAGTTGAGCGTTTTCAACTGGATCAAAATGTC  
 ACTCAAGGCTTAAAAAGTTAGGTCAAGATCAGGTTG  
 CAACCCGTGTTTATGACGTTGTTGGCCGGTTTCGGCGT 65  
 TTTGCTATCTCGTTATAGTGGTCAATCTGATCTGATG

36

- continued

GTGGGTTCTCCGATCGCTAATCGTAATCAAGCAGCGA  
 TCGAACCTTTAATGGCTTTTTTGTCTAACACTTTGGC  
 TTTAAGAATTAATTTATCAGAAAAATCCAGTTTTTTA  
 GAATTATTAGAACAAGTTAAACAGACAACTTTAGAGG  
 GTTATGCTCACCAAGACCTACCTTTGAGATGTTAGT  
 AGAAAAGCTACAACCTGACCGTGATTTGAGCAGAAAT  
 CCTTTAGTACAAGTCATGTTTGGCTACAAAATACCT  
 CTCAAGATACTTGGAACTTTCGGGTTAAGTATTGA  
 AAGTTTATCTTTATCAGTGGAGAACTGTGAGATTT  
 GATCTAGAAGTAAACTGCTGGCAAAATTCAGAAGGTT  
 TAGCAATAGATTGATTACAGCAGAGATTTATTGA  
 CACTGCAACAATTGCAAGAATGGGAGAACATTTTCAA  
 AATTTAGTTTCAGGCAATCATACTCAATCCAAAAGCTA  
 CAGTTAAAGAACTTCCTTTATTAACCCCAAGGAACG  
 TGAGCAATTATTAATATCTTGAATAATAGCAAGACT  
 GATTATCCTCAAGAGCAGTGTATTTATCAATTTATTTG  
 AAGCACAAAGTTGAACGACTCCAAGGCGATCGCAGT  
 GGTATTTGAGGAGCAATCATTAAACATACACTGAATTA  
 AACCATCGCGCTAATCAGTTAGCCCATTTATTACAAA  
 CTTTAGGCGTGGGAGCAGAAGTCTTAGTCGGTATTTTC  
 CCTAGAACGTTCTTTAGAGATGATTATCGGCTTATTA  
 GGGATTCTCAAGGTAGTGGTGTCTATCTTCTCTGTG  
 ATCCAGACTATCCCACTGAGCGTCTCAGTTGATGTT  
 AGAAGACAGTCAAGTTCCTTTTTGATTACCCACAGT  
 TCTTTATTAGCAAAATTCCTCCCTCTCAAGCAACTC  
 TGATTTGTTTAGATCATATCCAAGAGCAGATTTCTCA  
 ATATCTCCAGATAATCTCAATGTCAGTTAACTCCT  
 GCCAATTTAGCTAACGTTATTTATACCTCTGGCTCTA  
 CGGGTAAGCCTAAAGGGGTGATGGTTGAACATAAAGG  
 TTTAGTTAACTTAGCTCTGCTCAAATTCATCTTTT  
 GCAGTCAACCATAACAGTCGTGTGCTCAATTTGCTT  
 CTTTTAGTTTTGATGCTTGTATTTAGAAAATTTTGTG  
 GACCTTTGGTTCTGGAGCGACGCTTTATCTTGCACAA  
 AAAGATGCTTTATTGCCAGGTCAGCCATTAATTGAAC  
 GGTTAGTAAAGAAATGGAATTAATCATGTGACTTTGCC  
 GCCTTCAGCTTTAGTGGTTTTACCCAGGAACCGTTA  
 CGCAACTTAGAAACCTTAATTTGGCCGGGTGAGGCTT  
 GTTCTCTTGATTTAGTAAACAATGGTCAATCGATAG  
 AAACTTTTTCAATGCCTATGGGCCAACGGAAGCGAGT  
 GTTTGTGCCACTATTGGACAATGTTATCAAGATGATT

37

-continued

TAAAGGTGACGATTGGTAAGGCGATCGCCAATGTCCA  
 AATTTATATTTTAGATGCCTTTTACAGCCGGTGCCG 5  
 GTGGGAGTGTGAGGAGTTATACATTGGTGGAGTTG  
 GGGTGGCAAGGGCTATTTAAATCGTCTGAATTAAC  
 CCAAGAAAAATTTATGCTAATCCTTTTAGTAACGAC 10  
 CCAGATTCTCGGCTCTATAAACTGGCGACTTAGCGC  
 GTTATTTACCCGATGGTAATATTGAATATTTAGACG  
 CATTGACAATCAGGTAAAAATTCGCGGTTTTTCGCATT 15  
 GAGTTAGGAGAAAATGAAGCGGTTCTGAGTCAATGTC  
 CCGATGTGCAAAATACGGCGGTGATTGTCCGGAAGA  
 TACTCCTGGCGATAAGCGCTTAGTTGCCTATGTGGTT 20  
 CTTACTTCTGACTCCCAGATAACTACTAGCGAACTGC  
 GTC AATTTTGGCGAATCAATACCCGCTATCTTGT  
 TCCTAATACCTTTGTATTTTAGATGATTTGCCCTA 25  
 ACCCCAGTGGCAAATGCGATCGCCGTTCTTACCTA  
 TACCCGAAACACAAGCGTTATCAAATGACTATATTGC  
 CCCTAAATCTCCACTGAAGAAATCTGGCTCAAATA 30  
 TGGGGCAAGTTCTCAAGATAGAAAGAGTCAGCAGAG  
 AAGATAATTTCTTTGAATGGGGGGCATTCCCTTTT  
 AGCTACCCAGGTAATGTCCCGTCTCGGTGAAACTTTT 35  
 CAAGTCGAATTACCTTTGCGTAGTCTCTTACCGCTC  
 CCAC TATTGCTGAATGGCCCTAACAAATGAGCAATC  
 TCAGCAAACCATGCTGCTCCCCCATCCTAACCGA 40  
 AACGACAGTGCTAACCTCCCGTTATCTTTTGCTCAAC  
 AACGTTTATGGTTTCTGGATCAATTAGAACCTAACAG  
 CGCCTTTTATCATGTAGGGGAGCCGTAAGACTAGAA 45  
 GGAACATTAATATTAAGCTTAGAGCAAAGCTTAA  
 AAGAAATTATTAATCGTCATGAAGCTTTACGCACAAA  
 TTTTATAACGATTGATGGTCAAGCCACTCAAATTATT  
 CACCCTACTATTAATGGCGATTGCTGTGTTGTTGATT 50  
 GTC AAAATTTAACCGACTCAATCTCTGGAATTGC  
 GGAAGCTGAAAAGCCCTTTAATCTTGCTCAAGATTGC  
 TTATTTGCTGCTACTTTATTCGTGCGATCACCGCTAG 55  
 AATATCATCTACTCGTGACCATGCACCATATTGTTAG  
 CGATGGCTGGTCAATTGGAGTATTTTTCAAGAACTA  
 ACTCATCTTTACGCTGTCTATAATCAGGGTTTACCCT 60  
 CATCTTTAACGCCATTA AAAATACAATAAGCTGATTT  
 TGCGGTCTGGCAACGGAATTGGTTACAAGGTGAAATT  
 TTAAGTAATCAATTGAATTATTGGCGCGAACAAATTAG 65  
 CAAATGCTCCTGCTTTTTTACCTTTACCGACAGATAG

38

-continued

ACCTAGGCCCGCAATCCAACTTTTATTGGTTCTCAT  
 CAAGAATTTAACTTTCTCAGCCATTAAGCCAAAAAT  
 TGAATCAACTAAGTCAGAAGCATGGAGTGACTTTATT  
 TATGACTCTCTGGCTGCTTTTGTACCTTACTTTTAC  
 CGTTATACAGGACAAGCAGATATTTTAGTTGGTTCTC  
 CTATTGCTAACCGTAATCGTAAGGAAATGAGGGATT  
 AATCGGCTTTTTTGTTAATACATTAGTTCTGAGATTG  
 AGTTTAGATAATGATTTAAGTTTTCAAAATTTGCTAA  
 ACCATGTTAGAGAGGTTTCTTTAGCAGCCTACGCCCA  
 TCAAGATTTACCTTTTGAATGTTAGTAGAAGCACTA  
 CACCCTCAACGAGATCTCAGTCATACCCCTTATTTTC  
 AGGTAATGTTTGTTTTGCAAAATACACCAGTGGCTGA  
 TCTAGAACTTAAAAATGTAAAGGTTTGTCTCTACCG  
 ATGGAAAATAAGACTGCTAAATTTGATTTAACCTTAT  
 CAATGGAGAATCTAGAGGAAGGATTGATTGGGGTTTG  
 GGAATATAACACCGATCTATTTAATGGCTCAACCATT  
 GAGCGAATGAGTGGACATTTTGTCACTTTGTTAGAAG  
 ATATTGTTGCCGCTCCAACGAAGTCAGTTTTACGGTT  
 GTC TTTGCTGACGCAAGAGGAAAAATGCAATTATTG  
 ATTAAAAATCAGGGTGTTCAGTTGATTATCTCAAG  
 AGCAGTGATCCATCAATTATTGAAGCGCAAGTTGA  
 ACGGACTCCCGATGCGATTGCGGTGTTATTTGAGGAG  
 CAATCATTAACCTATGCTGAATTAATCATCAAGCTA  
 ATCAGTTAGTCCATTACTTACAACTTTAGGAATTGG  
 GCCAGAGGTCTTAGTCGCTATTTAGTAGAACGTTCT  
 TTAGAAATGATTATCGGCTTATTAGCCATTCTCAAGG  
 CGTGTGGTGCTTATCTCCCTCTGCTCCTGACTATCC  
 CACTGAGCGTCTTCAAGTTATGTTAGAAGATAGTCAA  
 GCTTCTTTTTTATTACCCACAGTCTTTATTAGAAA  
 AATTGCCTTCTCTCAAGCGACTCTAATTTGTTTAGA  
 TCACATCCAAGAGCAGATTTCTCAATATCTCCCGAT  
 AATCTTCAAAGTGAGTTAACTCCTTCCAATTTGGCTA  
 ACGTTATTTACACCTCTGGCTCTACGGGTAAGCCTAA  
 AGGGGTGATGGTTGAACATCGGGCTTAGTTAACTTA  
 GCGAGTTCTCAAATCAATCTTTGTCAGTCAAAAAATA  
 ACAGTCGTGACTGCAATTTGCTTCTTTAGTTTTGA  
 TGCTTGTATTTAGAAAATTTGATGACCTTTGGTTCT  
 GGAGCGACTCTTTATCTGCTCAAAAAAATGATTTAT  
 TGCCAGGTCAGCCATTAATGAAAGGTTAGAAAAGAA  
 TAAAATACCCATGTTACTTTACCCCTTCAGCTTTA

-continued

GCTGTTTTACCAAAAAACCGTTACCCAACCTTACAAA  
 CTTTAATTGTGGCGGTGAGGCTTGTCCTCTGGATTT 5  
 AGTCAAACAATGGTCAGTCGGTAGAAACTTTTTCAAT  
 GCCTATGGCCCGACAGAAACGAGTGTTTGTGCCACGA  
 TTGGACAATGTTATCAAGATGATTTAAAGGTCACGAT 10  
 TGGTAAGGCGATCGCTAATGTCCAATTTATATTTG  
 GATGCCCTTTTACAACAGTACCCATCGGAGTACCAG  
 GGGAATTATACATTGGTGGAGTCGGAGTTGCGAGGGG 15  
 TTATCTAAATCGTCTGAATTAACGGCGGAAAGATTT  
 ATTCCTAATCCTTTTGATCCCCCTAACCCCCCTTA  
 AAAAGGGGGGAGATAAGAGCTATGAACTTTTAAAAA 20  
 GGGGGAAGAGCAACCATCAAACTCTATAAAACGGGA  
 GATTTAGCTCGTTATTTACCCGATGGCAATATTGAAT  
 ATTTAGGACGCATTGACAATCAGTAAAAATTGCGGG 25  
 TTTTCGCATTGAGTTAGGAGAAATTGAAGCGGTTCTG  
 AGTCAATGTCCCGATGTGCAAAATACGGCGGTGATTG  
 TCCGTGAAGATACTCTGGCGATAAACGTTTAGTTGC 30  
 CTATGTGGTTCTTACTTCTGACTCCCAGATAACTACT  
 AGCGAACTGCGTCAATCTTGCTAATCAATTACCTG  
 CCTATCTCGTTCCTCAATACCTTTGTATTTTAGATGA 35  
 TTTGCCCCAACCCCCAATGGTAAATGCGATCGCCGT  
 TCCTTACCGCTTCTGATGATCAGACCAGAAAAATA  
 TTCCATAAATTGGCCCGGTAATTTAGTGAATTACA 40  
 ATTAGCTCAAATCTGGTCAGAGATTTTAGGCATTAAT  
 AATATTGGTATTTCAGGAAACTTCTTTGAATTAGGCG  
 GTCATTCTTTATTAGCAGTCAGTCTGATCAATCGTAT  
 TGAACAAAAGTTAGATAAACGTTTACCATTAACCAGT 45  
 CTTTTTCAAAATGGAACCATAGCAAGCTAGCTCAAT  
 TACTAGCGCAAGAAACAACCTCAGCCAGCCTCTTCACC  
 GTTGATTGCTATCCAGTCTCAAGGTGATAAACTCCA 50  
 TTTTTTGCTGTTCAATCCATTGGTGGTAATGTGCTAT  
 GTTATGCCGATTTAGCTCGTAATTTAGGAACGAAACA  
 GCCGTTTTATGGATTACAATCATTAGGGCTAAGTGAA 55  
 TTAGAAAAACTGTAGCCTCTATTGAAGAAATGGCGA  
 TGATTTATATTGAAGCAATACAACTGTTCAGCCCTC  
 TGGTCCCTACTATTTAGGAGGTTGGTCAATGGGAGGA 60  
 GTGATAGCTTTTGAATCGCCCAACATTTAGACCC  
 AAGGTCAAGAAGTTGCTTFACTGGCTTTAATAGATAG  
 TTATTCTCCAGTTTACTTAATTCAGTTAATAGGGAG 65  
 AAAAAATCTGCTAATTCCTGACAGAAGATTTAATG

-continued

AAGATATCAATATTGCCATTCTTTCATCAGAGACTT  
 AGCAAGTATATTTAATCAAGAAATCTCTTCTCTGGG  
 AGTGAACTTGCTCATTTTACATCAGACGAATTACTAG  
 ACAAGTTTATTACTTGGAGTCAAGAGACGAATCTTTT  
 GCCGTCAGATTTTGGGAAGCAGCAGGTTAAAACTGG  
 TTTAAAGTTTTCCAGATTAATCACCAGCTTTGAGCA  
 GCTATTCTCCCAAGCGTATCTGGGTAGAAAGTGT  
 CTTAGGAGCGGAAGACAGTTCTATTAAAACTCTGGT  
 TGGCATCAAGTAATCAATGACTTGCAATCTCAATGGA  
 TTAGCGCGATCACTACGGTTTAAATAAAAATCCAGT  
 CCTCGCTGAAAACTCAATAGCTACCTAGCCTAAAAAC  
 TTTCAAAAAGCCTGATTATTGTTTAAAAATGAATGATC  
 GTTACCGGTCAGAGGACAAGTATGACAACCCAAACA  
 GCTTCTAGTGCCAATGCCCTTGCTTCTTTAACCAAT  
 TTTTAAGGGATGTAAGGCGATCGCCCAACCCATTG  
 GTATCCACTGTATCAATAAAAAGAAGCTTTTCTGAG  
 GTTATTGTTCTCGGGAATGCTATCACTGCTTATCT  
 TTTTGATTGTGGGATTAGTCGCGCTCACGGCTTTTAA  
 TAGTTTTGTTAATCGTGTAAATTTGATGTCAATATT  
 CAAGAAAAGATGCGTCTCAATTTGCCAGTACATTAA  
 CTGTCTATGCGATCGGATTAATCTGTGTAACGCTGCT  
 GGCAGGGTTCCTAAAGATATTCGCAAAAAAATTGCC  
 CTAGATTGGTATCAATGGTTAAACACCAGATTGTAG  
 AGAAATATTTTAGTAATCGTGCTATATAAAATTA  
 CTTTCAATCTGACATTGATAACCCCGATCAACGTCTA  
 GCCCAGGAAATGAAACCGATCGCCACAAACGCCATTA  
 GTTCTCGGCCACTTTTTTGGAAAAAAGTTGGAAAT  
 GCTAACTTTTTTAGTGGTAGTTTGGTCAATTTCTCGA  
 CAGATTGCTATTCGCTAATGTTTACACGATTATCG  
 GTAATTTTATTGCCGCTATCTAAATCAAGAATTAAG  
 CAAGATCAATCAGGCACAACCTGCAATCAAAAGCAGAT  
 TATAACTATGCCTTAACCCATGTTCCGACTCATGCGG  
 AATCTATTGCTTTTTTTCGGGGAGAAAAAGGGAACA  
 AAATATTATTACAGCGAGTTTTTCAGGAAGTTATCAAT  
 GATACGAAAAATAAAATTAAGTGGGAAAAAGGGAATG  
 AAATTTTTTAGTCGGGCTATCGTTCGCTATTAGTT  
 TTTTCTTTTTTAGTCTTGGCCCTTTGTATATTA  
 GGAGAAATTGATTATGGACAAGTTGAGCAAGCTTCAT  
 TAGCTAGTTTTATGTTTGCATCGGCCCTGGGAGAAAT  
 AATTACAGAATTTGGTACTTCAGGACGTTTTTCTAGT

-continued

TATGTAGAACGTTTAAATGAATTTTCTAATGCCTTAG  
 AACTGTGACTAAACAAGCCGAGAATGTCAGCACAAT 5  
 TACAACCATAGAAGAAAATCATTTCCTTTGAACAC  
 GTCACCCCTAGAAACCCCTGACTATGAAAAGGTGATTG  
 TTGAGGATTTATCTTACTGTTCAAAAAGGTGAAGG 10  
 ATTATTGATTGTCCGGCCAGTGGTCGAGGTAAAAGT  
 TCTTTATTAAGGGCGATCGCCGGTTTATGGAATGCTG  
 GCACTGGGCGTTTAGTGCCTCCCTAGAAGAAAT 15  
 TCTCTTTTGCCTCAACGTCCTACATTATTTGGGA  
 ACCTTACGCGAACAATTGCTGTATCCTCTAACCAATA  
 GTGAGATGAGCAATACCGAACTTCAAGCAGTATTACA 20  
 ACAAGTCAATTGCAAAATGTGCTAAATCGGGTGGAT  
 GACTTTGACTCCGAAAAACCCCTGGGAAAACATTCTCT

-continued

CCCTCGGTGAACAACAACGCTAGCCTTTGCTCGATT  
 GTTAGTGAATTCCTCCGAGTTTACCATTTTAGATGAG  
 GCGACCAGTGCCTTAGATTTAACAAATGAGGGGATTT  
 TATACGAGCAATTACAACTCGCAAGACAACCTTTAT  
 TAGTGTGGGTCATCGAGAAAGTTTGTTTAATTACCAT  
 CAATGGGTTTTAGAACTTCTGCTGACTCTAGTTGGG  
 AACTCTTAAGCGTTCAGATTATCGCCTTAAAAAAGC  
 GGGAGAAATGTTTACTAATGCTTCGAGTAACAATTCC  
 ATAACACCCGATATTACTATCGATAATGGATCAGAAC  
 CAGAAATAGTCTATTCTCTTGAAGGATTTCCCATCA  
 GGAAATGAAACTATTAACAGACCTATCACTCTAGC  
 ATTCCGAGTAAAGCCAGTCGAGGGAAGGTGATTACAG  
 CCAAGGATGGTTTTACCTACCTTTATGACAAAAATCC  
 TCAGATATTAAGTGGCTCAGAACTTAA

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 61

<210> SEQ ID NO 1

<211> LENGTH: 580

<212> TYPE: PRT

<213> ORGANISM: *Microcystis aeruginosa*

<400> SEQUENCE: 1

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

-continued

---

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

&lt;210&gt; SEQ ID NO 2

&lt;211&gt; LENGTH: 88

-continued

---

```

<212> TYPE: PRT
<213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 2

Met Ser Thr Glu Ile Pro Asn Asp Lys Lys Gln Pro Thr Leu Thr Lys
1          5          10          15

Ile Gln Asn Trp Leu Val Ala Tyr Met Thr Glu Met Met Glu Val Asp
20          25          30

Glu Asp Glu Ile Asp Leu Ser Val Pro Phe Asp Glu Tyr Gly Leu Asp
35          40          45

Ser Ser Met Ala Val Ala Leu Ile Ala Asp Leu Glu Asp Trp Leu Arg
50          55          60

Arg Asp Leu His Arg Thr Leu Ile Tyr Asp Tyr Pro Thr Leu Glu Lys
65          70          75          80

Leu Ala Lys Gln Val Ser Glu Pro
85

<210> SEQ ID NO 3
<211> LENGTH: 431
<212> TYPE: PRT
<213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 3

Met Glu Pro Ile Ala Ile Ile Gly Leu Ala Cys Arg Phe Pro Gly Ala
1          5          10          15

Asp Asn Pro Glu Ala Phe Trp Gln Leu Met Arg Asn Gly Val Asp Ala
20          25          30

Ile Ala Asp Ile Pro Pro Glu Arg Trp Asp Ile Glu Arg Phe Tyr Asp
35          40          45

Pro Thr Pro Ala Thr Ala Lys Lys Met Tyr Ser Arg Gln Gly Gly Phe
50          55          60

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

Leu Glu Ala Thr Tyr Leu Asp Pro Gln Gln Arg Leu Leu Leu Glu Val
85          90          95

Thr Trp Glu Ala Leu Glu Asn Ala Ala Ile Val Pro Glu Thr Leu Ala
100         105         110

Gly Ser Gln Ser Gly Val Phe Ile Gly Ile Ser Asp Val Asp Tyr His
115         120         125

Arg Leu Ala Tyr Gln Ser Pro Thr Asn Leu Thr Ala Tyr Val Gly Thr
130         135         140

Gly Asn Ser Thr Ser Ile Ala Ala Asn Arg Leu Ser Tyr Leu Phe Asp
145         150         155         160

Leu Arg Gly Pro Ser Leu Ala Val Asp Thr Ala Cys Ser Ser Ser Leu
165         170         175

Val Ala Val His Leu Ala Cys Gln Ser Leu Gln Ser Gln Glu Ser Asn
180         185         190

Leu Cys Leu Val Gly Gly Val Asn Leu Ile Leu Ser Pro Glu Thr Thr
195         200         205

Val Val Phe Ser Gln Ala Arg Met Ile Ala Pro Asp Ser Arg Cys Lys
210         215         220

Thr Phe Asp Ala Arg Ala Asp Gly Tyr Val Arg Ser Glu Gly Cys Gly
225         230         235         240

Val Val Val Leu Lys Arg Leu Arg Asp Ala Ile Gln Asp Gly Asp Arg
245         250         255

```



-continued

---

Asp Ser Leu Glu Gln Glu Ala Ala Ala Ile Ser Tyr Gln Pro Leu Gln  
 210 215 220  
 Ile Pro Leu Val Ala Asn Leu Thr Gly Glu Val Leu Pro Glu Gly Ala  
 225 230 235 240  
 Thr Ile Glu Ala Arg Tyr Trp Arg Asn His Ala Arg Asn Pro Val Gln  
 245 250 255  
 Phe Tyr Gly Ser Ile Gln Thr Leu Ile Glu Gln Lys Phe Ser Leu Phe  
 260 265 270  
 Leu Glu Val Ser Pro Lys Pro Thr Leu Ser Arg Leu Gly Gln Gln Cys  
 275 280 285  
 Cys Pro Glu Arg Ser Thr Thr Trp Leu Phe Ser Leu Ala Pro Pro Gln  
 290 295 300  
 Glu Glu Glu Gln Ser Leu Leu Asn Ser Leu Ala Ile Leu Tyr Asp Ser  
 305 310 315 320  
 Gln Gly Ala Glu

<210> SEQ ID NO 5  
 <211> LENGTH: 68  
 <212> TYPE: PRT  
 <213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 5

Ile Thr Leu Gln Thr Leu Val Gly Asn Leu Leu Gln Leu Ser Pro Ala  
 1 5 10 15  
 Asp Val Asn Val His Thr Pro Phe Leu Glu Met Gly Ala Asp Ser Ile  
 20 25 30  
 Val Met Val Glu Ala Val Arg Arg Ile Glu Asn Thr Tyr Asn Val Lys  
 35 40 45  
 Ile Ala Met Arg Gln Leu Phe Glu Glu Leu Ser Thr Leu Asp Ala Leu  
 50 55 60  
 Ala Thr Tyr Leu  
 65

<210> SEQ ID NO 6  
 <211> LENGTH: 394  
 <212> TYPE: PRT  
 <213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 6

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



-continued

Gly Ser Leu Val Arg Trp His Leu Leu Lys Leu Glu Pro Glu Leu His  
 115 120 125  
 Leu Leu Val Leu Thr Ala His His Ile Ile Ser Asp Gly Trp Ser Met  
 130 135 140  
 Gly Val Ile Leu Arg Glu Leu Gly Glu Leu Tyr Ser Ala Lys Cys Gln  
 145 150 155 160  
 Gly Val Thr Ala Asn Leu Lys Thr Pro Lys Gln Phe Arg Glu Leu Ile  
 165 170 175  
 Glu Trp Gln Ser Gln Pro Ser Gln Gly Glu Glu Leu Lys Lys Gln Gln  
 180 185 190  
 Ala Tyr Trp Leu Ala Thr Leu Ala Asp Pro Pro Val Leu Asn Leu Pro  
 195 200 205  
 Thr Asp Lys Pro Arg Pro Ala Leu Pro Ser Tyr Gln Ala Asn Arg Arg  
 210 215 220  
 Ser Leu Thr Leu Asp Ser Gln Phe Thr Glu Lys Leu Lys Gln Phe Ser  
 225 230 235 240  
 Arg Lys Gln Gly Cys Thr Leu Leu Met Thr Leu Leu Ser Val Tyr Asn  
 245 250 255  
 Ile Leu Val His Arg Leu Thr Gly Gln Asp Asp Ile Leu Val Gly Leu  
 260 265 270  
 Pro Ala Ser Gly Arg Gly Leu Leu Asp Ser Glu Gly Met Val Gly Tyr  
 275 280 285  
 Cys Thr His Phe Leu Pro Ile Arg Ser Gln Leu Ala  
 290 295 300

<210> SEQ ID NO 8  
 <211> LENGTH: 430  
 <212> TYPE: PRT  
 <213> ORGANISM: *Microcystis aeruginosa*

<400> SEQUENCE: 8

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

-continued

180					185					190					
Ala	Lys	Pro	Asn	Gly	His	Lys	Asp	Ala	Ser	Tyr	Leu	Val	Asn	Leu	Ile
			195				200					205			
Gln	Glu	Gln	Gln	Val	Thr	Thr	Leu	His	Phe	Val	Ser	Ser	Met	Leu	Gln
	210					215					220				
Leu	Phe	Leu	Thr	Glu	Lys	Asp	Val	Glu	Lys	Cys	Asn	Ser	Leu	Lys	Arg
	225					230					235				240
Val	Ile	Cys	Ser	Gly	Glu	Ala	Leu	Ser	Leu	Glu	Leu	Gln	Glu	Arg	Phe
				245							250				255
Phe	Ala	Arg	Leu	Val	Cys	Glu	Leu	His	Asn	Leu	Tyr	Gly	Pro	Thr	Glu
			260						265					270	
Ala	Ala	Ile	His	Val	Thr	Phe	Trp	Gln	Cys	Gln	Ser	Asp	Ser	Asn	Leu
		275					280					285			
Lys	Thr	Val	Pro	Ile	Gly	Arg	Pro	Ile	Ala	Asn	Ile	Gln	Ile	Tyr	Ile
	290					295					300				
Leu	Asp	Ser	His	Leu	Gln	Pro	Val	Pro	Ile	Gly	Val	Ile	Gly	Glu	Leu
	305					310					315				320
His	Ile	Gly	Gly	Val	Gly	Leu	Ala	Arg	Gly	Tyr	Leu	Asn	Arg	Pro	Glu
				325					330					335	
Leu	Thr	Ala	Glu	Lys	Phe	Ile	Ala	Asn	Pro	Phe	Ala	Ser	Leu	Asp	Pro
			340					345					350		
Pro	Leu	Thr	Pro	Leu	Asp	Lys	Gly	Gly	Asp	Glu	Ser	Tyr	Lys	Thr	Phe
		355					360					365			
Lys	Lys	Gly	Gly	Glu	Gln	Pro	Ser	Arg	Leu	Tyr	Lys	Thr	Gly	Asp	Leu
	370					375					380				
Ala	Arg	Tyr	Leu	Pro	Asp	Gly	Lys	Ile	Glu	Tyr	Leu	Gly	Arg	Ile	Asp
	385					390					395				400
Asn	Gln	Val	Lys	Ile	Arg	Gly	Phe	Arg	Ile	Glu	Leu	Gly	Glu	Ile	Glu
			405						410					415	
Ala	Val	Leu	Leu	Ser	His	Pro	Gln	Val	Arg	Glu	Ala	Val	Val		
		420					425						430		

<210> SEQ ID NO 9  
 <211> LENGTH: 65  
 <212> TYPE: PRT  
 <213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 9

Glu	Ala	Ile	Ala	Ala	Ile	Phe	Gly	Gln	Val	Leu	Lys	Leu	Glu	Lys	Val
1			5					10						15	
Gly	Ile	Tyr	Asp	Asn	Phe	Phe	Glu	Ile	Gly	Gly	Asn	Ser	Leu	Gln	Ala
		20					25						30		
Thr	Gln	Val	Ile	Ser	Arg	Leu	Arg	Glu	Ser	Phe	Ala	Leu	Glu	Leu	Pro
		35					40					45			
Leu	Arg	Arg	Leu	Phe	Glu	Gln	Pro	Thr	Val	Ala	Asp	Leu	Ala	Leu	Ala
	50					55					60				
Val															
65															

<210> SEQ ID NO 10  
 <211> LENGTH: 300  
 <212> TYPE: PRT  
 <213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 10

Pro Arg Asp Gly Gln Leu Pro Leu Ser Phe Ala Gln Ser Arg Leu Trp

-continued

1	5	10	15
Phe Leu Tyr	Gln Leu Glu Gly Ala Thr Gly Thr Tyr Asn Met Thr Gly		
	20	25	30
Ala Leu Ser	Leu Ser Gly Pro Leu Gln Val Glu Ala Leu Lys Gln Ala		
	35	40	45
Leu Arg Thr	Ile Ile Gln Arg His Glu Pro Leu Arg Thr Ser Phe Gln		
	50	55	60
Ser Val Asp	Gly Val Pro Val Gln Val Ile Asn Pro Tyr Pro Val Trp		
	65	70	80
Glu Leu Ala Met	Val Asp Leu Thr Gly Lys Glu Thr Glu Ala Glu Lys		
	85	90	95
Leu Ala Tyr	Gln Glu Ser Gln Thr Pro Phe Asp Leu Thr Asn Ser Pro		
	100	105	110
Leu Leu Arg	Val Thr Leu Leu Lys Leu Gln Pro Glu Lys His Ile Leu		
	115	120	125
Leu Ile Asn Met	His His Ile Ile Ser Asp Gly Trp Ser Ile Gly Val		
	130	135	140
Phe Val Arg	Glu Leu Ser His Leu Tyr Arg Ala Phe Val Ala Gly Lys		
	145	150	155
Glu Pro Thr	Leu Pro Ile Leu Pro Ile Gln Tyr Ala Asp Phe Ala Val		
	165	170	175
Trp Gln Arg	Glu Trp Leu Gln Gly Lys Val Leu Ala Ala Gln Leu Glu		
	180	185	190
Tyr Trp Lys	Arg Gln Leu Ala Asp Ala Pro Pro Leu Leu Glu Leu Pro		
	195	200	205
Thr Asp Arg	Pro Arg Pro Ala Ile Gln Thr Phe Gln Gly Lys Thr Glu		
	210	215	220
Arg Phe Glu	Leu Asp Arg Lys Leu Thr Gln Glu Leu Lys Ala Leu Ser		
	225	230	235
Gln Gln Ser	Gly Cys Thr Leu Phe Met Thr Leu Leu Ala Ala Phe Gly		
	245	250	255
Val Val Leu	Ser Arg Tyr Ser Gly Gln Thr Asp Ile Val Ile Gly Ser		
	260	265	270
Ala Ile Ala	Asn Arg Asn Arg Gln Asp Ile Glu Gly Leu Ile Gly Phe		
	275	280	285
Phe Val Asn	Thr Leu Ala Leu Arg Leu Asp Leu Ser		
	290	295	300

&lt;210&gt; SEQ ID NO 11

&lt;211&gt; LENGTH: 409

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Microcystis aeruginosa

&lt;400&gt; SEQUENCE: 11

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

-continued

---

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

&lt;210&gt; SEQ ID NO 12

&lt;211&gt; LENGTH: 65

&lt;212&gt; TYPE: PRT

<213> ORGANISM: *Microcystis aeruginosa*

&lt;400&gt; SEQUENCE: 12

Ala Gln Leu Thr Gln Ile Trp Ser Glu Val Leu Gly Leu Glu Arg Ile  
 1                  5                                  10                                  15  
 Gly Val Lys Asp Asn Phe Phe Glu Leu Gly Gly His Ser Leu Leu Ala  
                   20                                  25                                  30  
 Thr Gln Val Leu Ser Arg Ile Asn Ser Ala Phe Gly Leu Asp Leu Ser  
                   35                                  40                                  45

-continued

---

Val Gln Ile Met Phe Glu Ser Pro Thr Ile Ala Gly Ile Ala Gly Tyr  
50 55 60

Ile  
65

<210> SEQ ID NO 13  
<211> LENGTH: 305  
<212> TYPE: PRT  
<213> ORGANISM: *Microcystis aeruginosa*  
  
<400> SEQUENCE: 13

Ala Arg Asp Gly His Leu Pro Leu Ser Phe Ala Gln Gln Arg Leu Trp  
1 5 10 15  
Phe Leu His Tyr Leu Ser Pro Asp Ser Arg Ser Tyr Asn Thr Leu Glu  
20 25 30  
Ile Leu Gln Ile Asp Gly Asn Leu Asn Leu Thr Val Leu Glu Gln Ser  
35 40 45  
Leu Gly Glu Leu Ile Asn Arg His Glu Ile Phe Arg Thr Thr Phe Pro  
50 55 60  
Thr Val Ser Gly Glu Pro Ile Gln Lys Ile Ala Leu Pro Ser Arg Phe  
65 70 75 80  
Gln Leu Lys Val Asp Asn Tyr Gln Asp Leu Asp Glu Asn Glu Gln Ser  
85 90 95  
Ala Lys Ile Gln Gln Val Ala Glu Leu Glu Ala Gly Gln Ala Phe Asp  
100 105 110  
Leu Thr Val Gly Pro Leu Ile Gln Phe Lys Leu Leu Gln Leu Ser Pro  
115 120 125  
Gln Lys Ser Val Leu Leu Leu Lys Met His His Ile Ile Tyr Asp Gly  
130 135 140  
Trp Ser Phe Gly Ile Leu Ile Arg Glu Leu Ser Ala Leu Tyr Glu Ala  
145 150 155 160  
Phe Leu Lys Asn Leu Ala Asn Pro Leu Pro Ala Leu Ser Ile Gln Tyr  
165 170 175  
Ala Asp Phe Ala Val Trp Gln Arg Gln Tyr Leu Ser Gly Glu Val Leu  
180 185 190  
Asp Lys Gln Leu Asn Tyr Trp Gln Glu Gln Leu Ala Thr Val Ser Pro  
195 200 205  
Val Leu Thr Leu Pro Thr Asp Arg Pro Arg Pro Ala Ile Gln Thr Phe  
210 215 220  
Gln Gly Gly Val Glu Arg Phe Gln Leu Asp Gln Asn Val Thr Gln Gly  
225 230 235 240  
Leu Lys Lys Leu Gly Gln Asp Gln Val Ala Thr Leu Phe Met Thr Leu  
245 250 255  
Leu Ala Gly Phe Gly Val Leu Leu Ser Arg Tyr Ser Gly Gln Ser Asp  
260 265 270  
Leu Met Val Gly Ser Pro Ile Ala Asn Arg Asn Gln Ala Ala Ile Glu  
275 280 285  
Pro Leu Ile Gly Phe Phe Ala Asn Thr Leu Ala Leu Arg Ile Asn Leu  
290 295 300

Ser  
305

<210> SEQ ID NO 14  
<211> LENGTH: 395  
<212> TYPE: PRT

-continued

---

<213> ORGANISM: *Microcystis aeruginosa*

<400> SEQUENCE: 14

Thr Tyr Thr Glu Leu Asn His Arg Ala Asn Gln Leu Ala His Tyr Leu  
 1 5 10 15  
 Gln Thr Leu Gly Val Gly Ala Glu Val Leu Val Gly Ile Ser Leu Glu  
 20 25 30  
 Arg Ser Leu Glu Met Ile Ile Gly Leu Leu Gly Ile Leu Lys Val Gly  
 35 40 45  
 Gly Ala Tyr Leu Pro Leu Asp Pro Asp Tyr Pro Thr Glu Arg Leu Gln  
 50 55 60  
 Leu Met Leu Glu Asp Ser Gln Val Pro Phe Leu Ile Thr His Ser Ser  
 65 70 75 80  
 Leu Leu Ala Lys Leu Pro Pro Ser Gln Ala Thr Leu Ile Cys Leu Asp  
 85 90 95  
 His Ile Gln Glu Gln Ile Ser Gln Tyr Ser Pro Asp Asn Leu Gln Cys  
 100 105 110  
 Gln Leu Thr Pro Ala Asn Leu Ala Asn Val Ile Tyr Thr Ser Gly Ser  
 115 120 125  
 Thr Gly Lys Pro Lys Gly Val Met Val Glu His Lys Gly Leu Val Asn  
 130 135 140  
 Leu Ala Leu Ala Gln Ile Gln Ser Phe Ala Val Asn His Asn Ser Arg  
 145 150 155 160  
 Val Leu Gln Phe Ala Ser Phe Ser Phe Asp Ala Cys Ile Ser Glu Ile  
 165 170 175  
 Leu Met Thr Phe Gly Ser Gly Ala Thr Leu Tyr Leu Ala Gln Lys Asp  
 180 185 190  
 Ala Leu Leu Pro Gly Gln Pro Leu Ile Glu Arg Leu Val Lys Asn Gly  
 195 200 205  
 Ile Thr His Val Thr Leu Pro Pro Ser Ala Leu Val Val Leu Pro Gln  
 210 215 220  
 Glu Pro Leu Arg Asn Leu Glu Thr Leu Ile Val Ala Gly Glu Ala Cys  
 225 230 235 240  
 Ser Leu Asp Leu Val Lys Gln Trp Ser Ile Asp Arg Asn Phe Phe Asn  
 245 250 255  
 Ala Tyr Gly Pro Thr Glu Ala Ser Val Cys Ala Thr Ile Gly Gln Cys  
 260 265 270  
 Tyr Gln Asp Asp Leu Lys Val Thr Ile Gly Lys Ala Ile Ala Asn Val  
 275 280 285  
 Gln Ile Tyr Ile Leu Asp Ala Phe Leu Gln Pro Val Pro Val Gly Val  
 290 295 300  
 Ser Gly Glu Leu Tyr Ile Gly Gly Val Gly Val Ala Arg Gly Tyr Leu  
 305 310 315 320  
 Asn Arg Pro Glu Leu Thr Gln Glu Lys Phe Ile Ala Asn Pro Phe Ser  
 325 330 335  
 Asn Asp Pro Asp Ser Arg Leu Tyr Lys Thr Gly Asp Leu Ala Arg Tyr  
 340 345 350  
 Leu Pro Asp Gly Asn Ile Glu Tyr Leu Gly Arg Ile Asp Asn Gln Val  
 355 360 365  
 Lys Ile Arg Gly Phe Arg Ile Glu Leu Gly Glu Ile Glu Ala Val Leu  
 370 375 380  
 Ser Gln Cys Pro Asp Val Gln Asn Thr Ala Val  
 385 390 395

-continued

---

```

<210> SEQ ID NO 15
<211> LENGTH: 65
<212> TYPE: PRT
<213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 15

Glu Ile Leu Ala Gln Ile Trp Gly Gln Val Leu Lys Ile Glu Arg Val
1           5           10           15
Ser Arg Glu Asp Asn Phe Phe Glu Leu Gly Gly His Ser Leu Leu Ala
20           25           30
Thr Gln Val Met Ser Arg Leu Arg Glu Thr Phe Gln Val Glu Leu Pro
35           40           45
Leu Arg Ser Leu Phe Thr Ala Pro Thr Ile Ala Glu Leu Ala Leu Thr
50           55           60

Ile
65

```

```

<210> SEQ ID NO 16
<211> LENGTH: 299
<212> TYPE: PRT
<213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 16

Asn Asp Ser Ala Asn Leu Pro Leu Ser Phe Ala Gln Gln Arg Leu Trp
1           5           10           15
Phe Leu Asp Gln Leu Glu Pro Asn Ser Ala Phe Tyr His Val Gly Gly
20           25           30
Ala Val Arg Leu Glu Gly Thr Leu Asn Ile Thr Ala Leu Glu Gln Ser
35           40           45
Leu Lys Glu Ile Ile Asn Arg His Glu Ala Leu Arg Thr Asn Phe Ile
50           55           60
Thr Ile Asp Gly Gln Ala Thr Gln Ile Ile His Pro Thr Ile Asn Trp
65           70           75           80
Arg Leu Ser Val Val Asp Cys Gln Asn Leu Thr Asp Thr Gln Ser Leu
85           90           95
Glu Ile Ala Glu Ala Glu Lys Pro Phe Asn Leu Ala Gln Asp Cys Leu
100          105          110
Phe Arg Ala Thr Leu Phe Val Arg Ser Pro Leu Glu Tyr His Leu Leu
115          120          125
Val Thr Met His His Ile Val Ser Asp Gly Trp Ser Ile Gly Val Phe
130          135          140
Phe Gln Glu Leu Thr His Leu Tyr Ala Val Tyr Asn Gln Gly Leu Pro
145          150          155          160
Ser Ser Leu Thr Pro Ile Lys Ile Gln Tyr Ala Asp Phe Ala Val Trp
165          170          175
Gln Arg Asn Trp Leu Gln Gly Glu Ile Leu Ser Asn Gln Leu Asn Tyr
180          185          190
Trp Arg Glu Gln Leu Ala Asn Ala Pro Ala Phe Leu Pro Leu Pro Thr
195          200          205
Asp Arg Pro Arg Pro Ala Ile Gln Thr Phe Ile Gly Ser His Gln Glu
210          215          220
Phe Lys Leu Ser Gln Pro Leu Ser Gln Lys Leu Asn Gln Leu Ser Gln
225          230          235          240
Lys His Gly Val Thr Leu Phe Met Thr Leu Ala Ala Phe Ala Thr
245          250          255

```

-continued

---

Leu Leu Tyr Arg Tyr Thr Gly Gln Ala Asp Ile Leu Val Gly Ser Pro  
                   260                                  265                                  270

Ile Ala Asn Arg Asn Arg Lys Glu Ile Glu Gly Leu Ile Gly Phe Phe  
                   275                                  280                                  285

Val Asn Thr Leu Val Leu Arg Leu Ser Leu Asp  
           290                                  295

<210> SEQ ID NO 17  
 <211> LENGTH: 415  
 <212> TYPE: PRT  
 <213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 17

Thr Tyr Ala Glu Leu Asn His Gln Ala Asn Gln Leu Val His Tyr Leu  
 1                  5                                  10                                  15

Gln Thr Leu Gly Ile Gly Pro Glu Val Leu Val Ala Ile Ser Val Glu  
                   20                                  25                                  30

Arg Ser Leu Glu Met Ile Ile Gly Leu Leu Ala Ile Leu Lys Ala Cys  
           35                                  40                                  45

Gly Ala Tyr Leu Pro Leu Ala Pro Asp Tyr Pro Thr Glu Arg Leu Gln  
           50                                  55                                  60

Phe Met Leu Glu Asp Ser Gln Ala Ser Phe Leu Ile Thr His Ser Ser  
 65                  70                                  75                                  80

Leu Leu Glu Lys Leu Pro Ser Ser Gln Ala Thr Leu Ile Cys Leu Asp  
                   85                                  90                                  95

His Ile Gln Glu Gln Ile Ser Gln Tyr Ser Pro Asp Asn Leu Gln Ser  
                   100                                  105                                  110

Glu Leu Thr Pro Ser Asn Leu Ala Asn Val Ile Tyr Thr Ser Gly Ser  
           115                                  120                                  125

Thr Gly Lys Pro Lys Gly Val Met Val Glu His Arg Gly Leu Val Asn  
 130                  135                                  140

Leu Ala Ser Ser Gln Ile Gln Ser Phe Ala Val Lys Asn Asn Ser Arg  
 145                  150                                  155                                  160

Val Leu Gln Phe Ala Ser Phe Ser Phe Asp Ala Cys Ile Ser Glu Ile  
                   165                                  170                                  175

Leu Met Thr Phe Gly Ser Gly Ala Thr Leu Tyr Leu Ala Gln Lys Asn  
           180                                  185                                  190

Asp Leu Leu Pro Gly Gln Pro Leu Met Glu Arg Leu Glu Lys Asn Lys  
           195                                  200                                  205

Ile Thr His Val Thr Leu Pro Pro Ser Ala Leu Ala Val Leu Pro Lys  
 210                  215                                  220

Lys Pro Leu Pro Asn Leu Gln Thr Leu Ile Val Ala Gly Glu Ala Cys  
 225                  230                                  235                                  240

Pro Leu Asp Leu Val Lys Gln Trp Ser Val Gly Arg Asn Phe Phe Asn  
           245                                  250                                  255

Ala Tyr Gly Pro Thr Glu Thr Ser Val Cys Ala Thr Ile Gly Gln Cys  
           260                                  265                                  270

Tyr Gln Asp Asp Leu Lys Val Thr Ile Gly Lys Ala Ile Ala Asn Val  
           275                                  280                                  285

Gln Ile Tyr Ile Leu Asp Ala Phe Leu Gln Pro Val Pro Ile Gly Val  
 290                  295                                  300

Pro Gly Glu Leu Tyr Ile Gly Gly Val Gly Val Ala Arg Gly Tyr Leu  
 305                  310                                  315                                  320

Asn Arg Pro Glu Leu Thr Ala Glu Arg Phe Ile Pro Asn Pro Phe Asp  
           325                                  330                                  335

-continued

---

Pro Pro Leu Thr Pro Leu Lys Lys Gly Gly Asp Lys Ser Tyr Glu Thr  
 340 345 350

Phe Lys Lys Gly Glu Glu Gln Pro Ser Lys Leu Tyr Lys Thr Gly Asp  
 355 360 365

Leu Ala Arg Tyr Leu Pro Asp Gly Asn Ile Glu Tyr Leu Gly Arg Ile  
 370 375 380

Asp Asn Gln Val Lys Ile Arg Gly Phe Arg Ile Glu Leu Gly Glu Ile  
 385 390 395 400

Glu Ala Val Leu Ser Gln Cys Pro Asp Val Gln Asn Thr Ala Val  
 405 410 415

<210> SEQ ID NO 18  
 <211> LENGTH: 65  
 <212> TYPE: PRT  
 <213> ORGANISM: *Microcystis aeruginosa*

<400> SEQUENCE: 18

Leu Gln Leu Ala Gln Ile Trp Ser Glu Ile Leu Gly Ile Asn Asn Ile  
 1 5 10 15

Gly Ile Gln Glu Asn Phe Phe Glu Leu Gly Gly His Ser Leu Leu Ala  
 20 25 30

Val Ser Leu Ile Asn Arg Ile Glu Gln Lys Leu Asp Lys Arg Leu Pro  
 35 40 45

Leu Thr Ser Leu Phe Gln Asn Gly Thr Ile Ala Ser Leu Ala Gln Leu  
 50 55 60

Leu  
 65

<210> SEQ ID NO 19  
 <211> LENGTH: 227  
 <212> TYPE: PRT  
 <213> ORGANISM: *Microcystis aeruginosa*

<400> SEQUENCE: 19

Thr Pro Phe Phe Ala Val His Pro Ile Gly Gly Asn Val Leu Cys Tyr  
 1 5 10 15

Ala Asp Leu Ala Arg Asn Leu Gly Thr Lys Gln Pro Phe Tyr Gly Leu  
 20 25 30

Gln Ser Leu Gly Leu Ser Glu Leu Glu Lys Thr Val Ala Ser Ile Glu  
 35 40 45

Glu Met Ala Met Ile Tyr Ile Glu Ala Ile Gln Thr Val Gln Ala Ser  
 50 55 60

Gly Pro Tyr Tyr Leu Gly Gly Trp Ser Met Gly Gly Val Ile Ala Phe  
 65 70 75 80

Glu Ile Ala Gln Gln Leu Leu Thr Gln Gly Gln Glu Val Ala Leu Leu  
 85 90 95

Ala Leu Ile Asp Ser Tyr Ser Pro Ser Leu Leu Asn Ser Val Asn Arg  
 100 105 110

Glu Lys Asn Ser Ala Asn Ser Leu Thr Glu Glu Phe Asn Glu Asp Ile  
 115 120 125

Asn Ile Ala Tyr Ser Phe Ile Arg Asp Leu Ala Ser Ile Phe Asn Gln  
 130 135 140

Glu Ile Ser Phe Ser Gly Ser Glu Leu Ala His Phe Thr Ser Asp Glu  
 145 150 155 160

Leu Leu Asp Lys Phe Ile Thr Trp Ser Gln Glu Thr Asn Leu Leu Pro  
 165 170 175

-continued

---

Ser Asp Phe Gly Lys Gln Gln Val Lys Thr Trp Phe Lys Val Phe Gln  
 180 185 190

Ile Asn His Gln Ala Leu Ser Ser Tyr Ser Pro Lys Thr Tyr Leu Gly  
 195 200 205

Arg Ser Val Phe Leu Gly Ala Glu Asp Ser Ser Ile Lys Asn Pro Gly  
 210 215 220

Trp His Gln  
 225

<210> SEQ ID NO 20  
 <211> LENGTH: 345  
 <212> TYPE: PRT  
 <213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 20

Phe Ser Leu Tyr Tyr Phe Gly Ser Tyr Glu Ala Glu Phe Asn Pro Asn  
 1 5 10 15

Lys Tyr Asn Leu Leu Phe Glu Gly Ala Lys Phe Gly Asp Arg Ala Gly  
 20 25 30

Phe Thr Ala Leu Trp Ile Pro Glu Arg His Phe His Ala Phe Gly Gly  
 35 40 45

Phe Ser Pro Asn Pro Ser Val Leu Ala Ala Ala Leu Ala Arg Glu Thr  
 50 55 60

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

Ser Ile Arg Val Ala Glu Glu Trp Ala Val Val Asp Asn Leu Ser Gln  
 85 90 95

Gly Arg Val Gly Ile Ala Phe Ala Ser Gly Trp His Pro Gln Asp Phe  
 100 105 110

Val Leu Ala Pro Gln Ser Phe Gly Gln His Arg Glu Leu Met Phe Gln  
 115 120 125

Glu Ile Glu Thr Val Gln Lys Leu Trp Arg Gly Glu Ala Ile Thr Val  
 130 135 140

Pro Asp Gly Lys Gly Gln Arg Val Glu Val Lys Thr Tyr Pro Gln Pro  
 145 150 155 160

Met Gln Ser Gln Leu Pro Ser Trp Ile Thr Ile Val Asn Asn Pro Asp  
 165 170 175

Thr Tyr Ile Arg Ala Gly Ala Ile Gly Ala Asn Ile Leu Thr Asn Leu  
 180 185 190

Met Gly Gln Ser Val Glu Asp Leu Ala Arg Asn Ile Ala Leu Tyr Arg  
 195 200 205

Gln Ser Leu Ala Glu His Gly Tyr Asp Pro Ala Ser Gly Thr Val Thr  
 210 215 220

Val Leu Leu His Thr Phe Val Gly Lys Asp Leu Glu Gln Val Arg Glu  
 225 230 235 240

Gln Ala Arg Gln Pro Phe Gly Gln Tyr Leu Thr Ser Ser Val Gly Leu  
 245 250 255

Leu Gln Asn Met Val Lys Ser Gln Gly Met Lys Val Asp Phe Glu Gln  
 260 265 270

Leu Arg Asp Glu Asp Arg Asp Phe Leu Leu Ala Ser Ala Tyr Lys Arg  
 275 280 285

Tyr Thr Glu Thr Ser Ala Leu Ile Gly Thr Pro Glu Ser Cys Arg Gln  
 290 295 300

Ile Ile Asp His Leu Gln Ser Ile Gly Val Asp Glu Val Ala Cys Phe

-continued

---

```

305                310                315                320
Ile Asp Phe Gly Val Asp Glu Gln Thr Val Leu Ala Asn Leu Pro Tyr
                325                330                335
Leu Gln Ser Leu Lys Asp Leu Tyr Gln
                340                345

```

```

<210> SEQ ID NO 21
<211> LENGTH: 33
<212> TYPE: PRT
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 21

```

```

Ile Asp Pro Pro Leu Thr Pro Leu Asp Lys Gly Ile Asp Pro Pro Leu
1                5                10                15
Thr Pro Leu Asp Lys Gly Ile Asp Pro Pro Leu Thr Pro Leu Asp Lys
                20                25                30

```

```

Gly

```

```

<210> SEQ ID NO 22
<211> LENGTH: 77
<212> TYPE: PRT
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 22

```

```

Pro Tyr Gln Gly Gly Leu Gly Gly Asp Gln Ser Pro Tyr Gln Gly Gly
1                5                10                15
Leu Gly Gly Asp Gln Ser Pro Tyr Gln Gly Gly Leu Gly Gly Asp Gln
                20                25                30
Ser Pro Tyr Gln Gly Gly Leu Gly Gly Asp Gln Ser Pro Tyr Gln Gly
                35                40                45
Gly Leu Gly Gly Asp Gln Ser Pro Tyr Gln Gly Glu Leu Gly Gly Asp
                50                55                60
Gln Ser Pro Tyr Gln Gly Gly Leu Gly Gly Asp Gln Val
65                70                75

```

```

<210> SEQ ID NO 23
<211> LENGTH: 382
<212> TYPE: PRT
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 23

```

```

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

```

-continued

---

Val Arg Ser Leu Pro Leu Leu Glu Pro Tyr His Ala Ala Val Gln Leu  
 130 135 140  
 Ala Gln Ala Ser Asp Ser Lys Thr Val Glu Gln Trp Gln Gln Gln Val  
 145 150 155 160  
 Arg Gln Ser Val Ala Gly Glu Glu Glu Leu Val Ile Asp Pro Thr Leu  
 165 170 175  
 Phe Leu Ala Leu Lys Gln His Phe Pro Gln Ile Ser Trp Val Glu Ile  
 180 185 190  
 Gln Pro Lys Arg Gly Val Ala His Asn Glu Leu Thr Gln Phe Arg Tyr  
 195 200 205  
 Asp Val Thr Leu His Leu Glu Thr Ile Asn Asn Gln Ala Leu Leu Ser  
 210 215 220  
 Gly Asn Pro Thr Val Ile Thr Trp Leu Asn Trp Gln Leu Asp Gln Leu  
 225 230 235 240  
 Ser Leu Thr Gln Ile Lys Asp Lys Leu Leu Thr Asp Lys Pro Glu Leu  
 245 250 255  
 Trp Gly Ile Arg Gly Ile Pro Asn Gln Arg Val Glu Glu Ala Leu Lys  
 260 265 270  
 Ile Trp Glu Trp Val Glu Asn Ala Pro Asp Val Glu Thr Val Glu Gln  
 275 280 285  
 Leu Lys Lys Leu Leu Lys Gln Gln Val Asp Thr Gly Ile Asn Pro Glu  
 290 295 300  
 Gln Val Trp Gln Leu Ala Glu Ser Leu Gly Tyr Thr Ala His Leu Ser  
 305 310 315 320  
 Trp Trp Glu Ser Ser Gln Asp Gly Ser Phe Asp Val Ile Phe Gln Arg  
 325 330 335  
 Asn Ser Glu Ala Glu Asp Ser Lys Lys Leu Thr Leu Ser Lys Leu Ala  
 340 345 350  
 Phe Trp Asp Glu Lys Pro Phe Lys Ile Lys Pro Trp Ser Asp Tyr Thr  
 355 360 365  
 Asn Asn Pro Leu Arg Gly Lys Leu Val Gln Lys Leu Ile Pro  
 370 375 380

&lt;210&gt; SEQ ID NO 24

&lt;211&gt; LENGTH: 270

&lt;212&gt; TYPE: PRT

<213> ORGANISM: *Microcystis aeruginosa*

&lt;400&gt; SEQUENCE: 24

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

-continued

Leu Ala Thr Phe Ala Thr Val Lys Asn His Leu Lys Ala Gly Gly Ile  
 130 135 140  
 Phe Ile Cys Asp Val Ser Tyr Gly Ser Tyr Val Leu Gly Glu Phe Lys  
 145 150 155 160  
 Ser Arg Pro Thr Ala Ser Ile Leu Arg Leu Glu Asp Asn Ser Asn Gly  
 165 170 175  
 Asn Glu Val Thr Tyr Ile Ser Glu Leu Asn Phe Leu Thr His Glu Asn  
 180 185 190  
 Ile Val Glu Val Thr His Asn Leu Trp Val Thr Asn Gln Glu Asn Gln  
 195 200 205  
 Leu Leu Glu Asn Ser Arg Glu Thr His Leu Gln Arg Tyr Leu Phe Lys  
 210 215 220  
 Pro Glu Val Glu Leu Leu Ala Asp Ala Cys Glu Leu Thr Val Leu Asp  
 225 230 235 240  
 Ala Met Pro Trp Leu Glu Gln Arg Pro Leu Thr Asn Ile Pro Cys Pro  
 245 250 255  
 Ser Val Cys Phe Val Ile Gly His Lys Thr Thr His Ser Ala  
 260 265 270

&lt;210&gt; SEQ ID NO 25

&lt;211&gt; LENGTH: 1743

&lt;212&gt; TYPE: DNA

<213> ORGANISM: *Microcystis aeruginosa*

&lt;400&gt; SEQUENCE: 25

atgactatta actatggtga tctgcaagaa ccctttaata aattctcaac cctagtgtgaa 60  
 ttactccggt atcgggcaag cagtcaaccg gaacgcctcg cctatatattt tctgcgagac 120  
 ggagaaatcg aagaagctcg ttaaacctat ggggaactgg atcaaaaggc tagggcgatc 180  
 gccgcttate tacaatcctt agaagccgag ggcgaaaggg gtttactgct ctatccccc 240  
 ggactagatt ttatttcage tttttttggt tgtttatatg cgggagtcgt tgccattccc 300  
 gcctatccac ccgacggaa tcaaaacctt ttgcgtttac aggcgattat tgccgattct 360  
 caagcccgat ttaccttcaac caatgccgct ctatttccca gtttaaaaaa ccaatgggct 420  
 aaagaccctg aattaggagc aatggaatgg attgttaccg atgaaattga ccatcacctc 480  
 agggaggatt ggctagaacc aaccctcgaa aaaaacagtc tcgcttttct acaatacacc 540  
 tctggttcaa cggaactcc aaagggagta atggtcagtc accataattt gttgattaat 600  
 tcagccgatt tagatcgtgg ttggggccat gatcaagata gcgtaatggt cacttggtca 660  
 ccgaccttcc atgatatggg tctgatttat ggggttattc agcctttgta caaaggattt 720  
 ctttgttaca tgatgtcccc tgccagcttt atggaacgac cgttacgttg gttacaggcc 780  
 ctttctgata aaaaagcaac ccatagtcgc gcccccaact ttgcctacga tctttgtgtg 840  
 cggaaaatc cccctgaaaa acgggctacg ttagacttaa gccattgggtg catggcctta 900  
 aatggggccc aaccctcag agcggaggtta cttaaaaagt ttgcggaggc tttcaagtt 960  
 tctggtttca aagccacagc cctttgtcct ggctacgggt tagcagaagc caccctgaaa 1020  
 gttacggcgg ttagttatga cagtccccct tacttttacc ccgttcaggc taatgcttta 1080  
 gaaaaaata agattgtggg agcactgaa accgatacca atgtgcagac cctcgtgggc 1140  
 tgccgctgga caacgattga tactcaaacc gtcattgtca atcctgaaac cctgaaacct 1200  
 tgctcccctg aaattgtcgg cgaaatttgg gtatcaggtt caacaatcgc ccaaggctat 1260  
 tggggaaaac ctcaagagac tcaggaaacc tttcaagctt atttggcaga tacaggagcc 1320

-continued

---

```

gggccttttc tgcgaacagg agacttgggc ttcattaaag atggtgaatt gtttatcaca 1380
ggtcggctca aggaaattat tctgattcga ggacgcaata attatcccca ggatattgaa 1440
ttaaccgtcc aaaatagtca tcccgtctg cgtcccagtt gtggggctgc tttaccgtt 1500
gaaaataagg gcgaagaaaa gctcgtggtc gttcaggaag tggagcgcac ctggctccgt 1560
aaggtagata tagatgaggt aaaaagagcc attcgtaaag ctggtgtcca ggaatatgat 1620
ttacaggttt atgcatcgc gctgatcagg actggcagtt taccaaaaac ctctagcgg 1680
aaaaattcagc gtcgtagctg tcgggccaaa ttttagagg gaagcctgga aattttgggc 1740
taa 1743

```

```

<210> SEQ ID NO 26
<211> LENGTH: 267
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 26

```

```

atgtccacag aaatcccaaa cgacaaaaaa caaccgacc taacgaaaat taaaaactgg 60
ttagtggcct acatgacaga gatgatggaa gtggacgaag atgagattga tctgagcgtt 120
ccctttgatg aatatggtct cgattcttct atggcagttg ctttgatcgc tgatctagag 180
gattggttac gacgagattt acatcgcacc ctgatctacg attatccaac tctagaaaag 240
ttggctaaac aggttagtga accctga 267

```

```

<210> SEQ ID NO 27
<211> LENGTH: 1293
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 27

```

```

atggaaccca tcgcaattat tggctttgct tgccgctttc caggggctga caatccagaa 60
gctttctggc aactcatgcg aaatgggggtg gatgcatcgc ccgatattcc tctgaacgt 120
tgggatattg agcgtttcta cgatcccaca cctgccactg ccaagaagat gtatagtcgc 180
cagggcgggt ttctaaaaaa tgcgatcaa tttgacctc aatttttccg aatttctccc 240
ctagaagcca cctatctaga tcctcaaaa agactgctac tggaaagcac ctgggaagcc 300
ttagaaaatg ctgccattgt gcctgaaacc ttagctggta gccaatcagg ggtttttatt 360
ggatcagtg atgtggatta tcctgctttg gcttatcaaa gtccactaa cttgacggcc 420
tatgtgggta caggcaacag caccagtatt gcggctaacc gtttatcata tctgttgat 480
ttgctgggcc ccagtttggc cgtagatacc gcttgctctt ctccctcgt cgcggtcac 540
ttggcctgtc agagtttgca aagtcaagaa tcgaacctct gcttagtggg gggagttaat 600
ctcattttgt cgcagagac aaccgttgtt ttttccaag cgagaatgat cgcgcccgac 660
agtcgttgta aaacctttga cgcgagggcc gatggttatg tgcgctcggg aggctgtgga 720
gtagtcgtac ttaaacgtct tagggatgcc attcaggacg gcgatcgcatt ttagcagtg 780
attgaaggtt ccgcggtgaa tcaggatggt ttaagtaatg gactcacggc ccctaattgg 840
cctgctcaac aggcggtgat tcgtcaggcc ctggcaaatg cccaggtaaa accggcccag 900
attagctatg tcgaagccca tggcacgggg acagaattgg gggatccgat cgaagttaa 960
tctctgaaa cggttttggg tgaaaagcga tcgctcgcac aaacctgttg gctcggttct 1020
gtgaaaacca acattggtca tttagaagcg gcggcgggaa tggcgggtct gattaaagtc 1080

```

-continued

---

```

gttctctgcc tacaacacca agaaattccc cctaattctcc actttcaaac ccttaatccc 1140
tatatttccc tagctgacac agcttttgcg attcccactc aggctcaacc ctggcggacc 1200
aaacccccta agtctgggta aaacgggtgc gaacgacggt tagcaggact cagttccttt 1260
gggtttgggg ggacaaatc ccatgtgatt etc 1293

```

```

<210> SEQ ID NO 28
<211> LENGTH: 972
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 28

```

```

gtttttctat ttgccggtca aggttctcaa tatgtaggta tgggtcgtca actgtacgaa 60
acccaaccca tctttcgcca aacctggat cgctgtgctg aaactctgcg accccattta 120
gatcaacccc tcttagaaat tctttatcct gctgaccag aagccgaaac agcgagtttt 180
tacctagagc agactgccta tacccaaccc actttattcg cattcgagta tgccttagca 240
cagttatggc gttcctgggg aatagaaccg gcggcagtaa ttggtcacag tgtcggtgaa 300
tatgtggcgg ccaccgttgc cggagcctta agtctagaag aaggattaac gctaattgcc 360
aaacgggcaa aactgatgca gtctctcccc aagaatggga caatgatcgc cgtttttgcc 420
gcagaagagc gggtaaagc tgttattgag ccttatagga ctgatgtagc gatcgctgct 480
gttaatggac cagaaaatth tgttatttca ggaaaagcgc cgattattgc tgagattatc 540
attcatttaa cggcagcagg aatagaagtt cgtcctctca aagtttccca tgettttcac 600
tcgcacctgt tggagccaat tttagattcc ttagaacagg aagctgctgc tatttcctac 660
caacccctgc aaattccctt agttgctaat ttaacggggg aagttctacc agaaggagca 720
acgattgagg ctcgttactg gcgaaatcat gcacgcaacc ctgtacaatt ttatgggagt 780
atccaaacgc tgatcgagca gaaattcagt ctttttttag aagttagccc taaaccgact 840
ttatctcgat tgggtcaaca atggtgtcca gaaagatcga ccacttgctt attttccctc 900
gcccctctc aagaagaaga acaaagccta ctaaatagtt tggcgattct ctatgattec 960
caaggagccg aa 972

```

```

<210> SEQ ID NO 29
<211> LENGTH: 204
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 29

```

```

atcacattgc aaacctagt gggaaattta ctgcaattgt ccctgctga tgtcaatggt 60
catacacctt tcttgagat gggggcagat tccattgtca tggttgaggc ggtcagacgg 120
attgagaata cctataacgt taaaattgct atgcgtcagt tatttgagga gttatctact 180
ttagatgctt tagctactta tttta 204

```

```

<210> SEQ ID NO 30
<211> LENGTH: 1182
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 30

```

```

aaagagatgc ttatcccat tgtggcccaa cgttctcaag gatcaagaat ttgggatgtg 60
gacggtaatg aatataattga tatgacgatg gggcaagggg taacgctggt tgggcatcaa 120
ccagacttca ttatgtcggc cctacaaagc caactcactg aaggcattca tctcaatccg 180

```

-continued

---

```

cgatcgccaa ttgtgggaga agtggccgcc ttaatttggtg aactaacagg agccgaacga 240
gcttggtttt gcaactctgg aaccgaagcc gtaatggccg ctattcgtat cgcagggca 300
acaacaggtc ggagtaaaat tgcctctttt gaaggctcct atcatggaca tgcggacgga 360
acccttttta ggaaccaaat tattgataac caactccact cttttccctt agctctaggc 420
gttcccccca gccttagttc cgatgtgggtg gtattggact atggcagtgc ggaagctctg 480
aactatttac aaaccagggg gcaggattta gcggcggtct tagtagaacc aattcaaagt 540
ggcaatcctc tactccaacc ccaacaattt ctccaaagtc tgcgacaaat taccagtcaa 600
atgggcattg ccctgatttt tgatgaaatg attacgggtt ttcgatcgca cccaggggga 660
gcgcaagcct tatttgagat acaggcggat attgccacct atggcaaagt agttgcggga 720
ggaaatgccc ttggagttat tgcaggtaag gccattatc tggacagcat tgacggggga 780
atgtggcgtt atggcgataa atcctatcct ggggtggaca gaaccttttt tgggggaacc 840
tttaatcagc atccgttagc aatggtagcg gctagggctg tctgaccca ttaaaggag 900
caggggccag gtctgcaaca acaattaact gaacgcactg cggccttagc cgatacactg 960
aatcattatt ttcaagccga agaagttcct attaaaatcg aacagtttag ttctttcttc 1020
cggtttgccc tctctggcaa tttgattta cttttctatc acatggtaga aaaaggtatt 1080
tatgtctggg aatggcgtaa acattttctt tcaaccgcc atacggaagc cgatcttggc 1140
caatttgtcc aagcgggtaa ggatagcatc acagaattgc gt 1182

```

```

<210> SEQ ID NO 31
<211> LENGTH: 900
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 31

```

```

gggggggatc aagtcctctt caccgaagcc caacgacaac tgtggatttt ggctcaatta 60
ggagacaacg gctctgtggc ctataacca tcaagtacat tgcaattaag tggccatta 120
aatcccgtcg caatgaatca agctattcaa caaatcagcg atcgccatga agcgttacga 180
accaaaatta atgccagggg agatagtcaa gaaatcctgc cccaggtcga aattaactgc 240
cctatcttag acttcagtct tgaccaagct tggcccaac agcaagcaga acaatggtta 300
aaggaagaaa gtgaaaaacc ctttgatttg agccagggtt ctctcgtgcg ttggcatcta 360
ctcaaattag aaccagaatt acatttgta gtattaacgg occatcacat taccagtgc 420
ggttggtcaa tgggggtaat ccttcgggaa ttaggagagt tatattcagc caaatgtcag 480
ggtgttacgg ctaatcttaa aaccccaaaa cagtttcgag aattgattga atggcaaagc 540
cagccaagcc aaggggaaga actgaaaaaa cagcaagcct attggttagc aacccttggc 600
gatccccctg ttttgaattt acccaactgac aaacctcgtc cagctttacc cagttaccaa 660
gctaactcgc gaagtctaac tttagatagc caatttacag aaaaactaaa gcaatttagt 720
cgtaaacagg gctgtacctt gctgatgacc ctgttatcgg tttataacat tctcgttcat 780
cgtttgacgg gacaggatga tattctggtg ggtctgccag cctctggacg ggggctttta 840
gatagtgaag gtatgggtgg ttattgcacc cattttttac caattcgcag tcaattagca 900

```

```

<210> SEQ ID NO 32
<211> LENGTH: 1290
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

-continued

&lt;400&gt; SEQUENCE: 32

```

acttacagtg aattaaattg tcgagccaat cagtttagcac attatttaca aaaattagga      60
gttgggccag aggtcttagt cggtatcttg gtcgaacggt ctttagaaat gattgtcggg      120
ttgttaggga ttctcaaggc tgggggagcc tatgtacctc ttgatcctga ctatccccct      180
gaacgtcttc aatttatgtt agaagatagt caattttttc tcctcttaac ccaacagcat      240
ttactggaat cttttgctca gtcttcagaa acggttactc ccaagattat ttgtttggat      300
agcgactacc aaattatttc ccaggcaaag aatattaatc ccgaaaattc agtcacaacg      360
agtaatcttg cctatgtaat ttataacctc ggttcgacag gtaaacgaa gggcgtgatg      420
aataatcatg ttgctattag taataaattg ttatgggtac aagacactta tcctctaacc      480
acagaagact gtattttaca aaaaactccc tttagttttg atgtttcagt gtgggaatta      540
ttctggcccc tactaaacgg agcgcggttg gtttttgcca agccgaatgg ccataaagat      600
gccagttact tagtcaatct gattcaagag caacaagtaa caacgctaca ttttgtgtct      660
tctatgctac agctttttct gacagaaaaa gacgtagaaa aatgtaatag tcttaaacga      720
gtcatttcta gtggtgaagc cctttcttta gagcttcaag aacgtttttt tgctcgttta      780
gtctgtgaat tacacaatct ttatggaccg acagaagccg ctattcatgt cacattttgg      840
caatgtcaat cagatagcaa tttgaaaaca gtaccattg gtcggccgat cgctaataac      900
caaatttaca ttttagactc tcattctcag ccagtaccta ttggagtaat cggagaattg      960
cacattgggtg gggttgggtt ggcgcggggt tatttaaaaca ggctgagtt aacggcggag      1020
aaatttattg caaatccggt tgcttccctt gatccccccc taacccccct tgataagggg      1080
ggagatgaga gctataaaac ttttaaaaag gggggagagc aaccatcaag attgtataaa      1140
acgggagatt tagctcgta tttaccgat ggcaagattg agtatctagg ggcattgat      1200
aatcaggtaa aaattcgcgg tttccgatt gaattggggg aaattgaagc ggttttgcta      1260
tcccatcccc aggtacgaga agcggtcggt      1290

```

&lt;210&gt; SEQ ID NO 33

&lt;211&gt; LENGTH: 195

&lt;212&gt; TYPE: DNA

<213> ORGANISM: *Microcystis aeruginosa*

&lt;400&gt; SEQUENCE: 33

```

gaggcgatcg ccgctatttt tggcgaagt ttaaaactgg aaaaagtggg aatttatgat      60
aacttttttg agatcggcgg taattctttg caagccactc aagttatttc acgcttacga      120
gaaagttttg ccctagagtt gcccttgcgt cgcctgtttg aacaaccgac tgtggcggat      180
ttggctttag ccgta      195

```

&lt;210&gt; SEQ ID NO 34

&lt;211&gt; LENGTH: 900

&lt;212&gt; TYPE: DNA

<213> ORGANISM: *Microcystis aeruginosa*

&lt;400&gt; SEQUENCE: 34

```

cctcgtgatg gccaaattacc cctctccttt gcccagtcgc gactctgggt cttgtatcaa      60
ttagaaggag ccacgggaac ctataacatg acaggggcct tgagttaaag cgggcctctt      120
caggtcgaag ccctcaaaaca agccctaaga actatcattc aacccatga gccattgcgt      180
accagtttcc aatcggttga cggggttcca gtgcagggtga ttaatcccta tcctgtttgg      240
gaattagcga tgggtgattt gacaggaaaag gagacagaag cagaaaaatt ggcctatcag      300

```

-continued

---

```

gaatcccaaa ccccgtttga tttgaccaat agtcctttgt tgagggtaac gctcctcaaa 360
ttacagccag aaaagcatal tttattaatt aatatgcacc atattatttc cgatggctgg 420
tcaatcggtg tttttgttcg tgaattgtcc catctctata gggcttttgt ggcgggtaaa 480
gaaccaactt taccgatttt accaattcag tatgcggatt ttgccgtttg gcagcgagag 540
tggttacagg gtaaggtttt agcggctcaa ttggaatatt ggaagcgaca attggcagat 600
gctcctcctc tgctggaact gcccactgat cgcctcctgc ccgcaatcca aacctttcaa 660
ggcaagacag aaagatttga gctagatagg aaactgaccc aagaatataa ggcattaagt 720
caacagtcgg gttgtacttt atttatgact ttgttggccg cttttggggt ggttttatcc 780
cgttatagtg gccagactga tatcgtcatt ggttcggcga tcgccaaaccg taatcgccaa 840
gacattgagg ggtaattgg cttttttgtt aacactttgg cgttgagggt agatttatca 900

```

&lt;210&gt; SEQ ID NO 35

&lt;211&gt; LENGTH: 1227

&lt;212&gt; TYPE: DNA

<213> ORGANISM: *Microcystis aeruginosa*

&lt;400&gt; SEQUENCE: 35

```

acctatggag aattaacca tcgcgccaat caattagctc actatcttca gtcgtagga 60
gtcaccaaaag aacaaatcgt cggggtttat ctggaacggt cccttgaaat ggcgatcgga 120
tttttaggta ttctcaaagc aggagccgcc tatctcccca ttgatcctga atatccctca 180
gtacgcaccc aatttattct cgaagatacc caactttcgc ttctctaac tcaggcagaa 240
ctggcagaaa aactgcccc gactcaaaac aaaattatct gtctagatcg ggactggcca 300
gaaattacct cccaaccca gacaaaccta gacctaaaga tagaacctaa taacctagcc 360
tattgcatct atacttcttg ttccacagga caacccaag gagtactgat ttccatcaa 420
gccctactca acttaatttt ctggcatcaa caagcgtttg agattggccc cttacataaa 480
gcgaccacaag tggcaggcat tgctttcgat gcaacggttt gggaaattgtg gccctatctg 540
accacaggag cctgtattaa tctggttccc caaaatattc tgctctcacc gacggattta 600
cgggattggt tgcttaaccg agaaattacc atgagttttg tgccaactcc tttagctgaa 660
aaattattat ccttggattg gcctaaccat tcttgtctaa aaacctggt actgggaggt 720
gacaaaactc atttttatcc tgctcgtcc cttcccttcc aggtcattaa caactatggc 780
ccaacggaaa atacagtgtg tcgcacctct ggactggtea aatcatcttc atctcatcac 840
tttgaacte cgactattgg tcgtccatt gccaacgtcc aaatctattt attagaccaa 900
aacctacaac ctgtccocat tgggttacca ggagaattac atttaggtgg ggcgggttta 960
gcgcagggct atctcaatcg tctgagtta acggetgaaa aatttattgc caatcccttt 1020
gatccccccc taacccccct tgataagggg ggagaagaac cctcaaaact ctataaaacg 1080
ggagacttag cccgttattt acccgatggc aatgtagaat ttttgggacg tattgacaat 1140
caggtaaaaa ttccgggttt tcgcatcgaa actggggaaa tcgaagccgt ttttaagtcaa 1200
tatttcctat tagctgaaag tgtagtc 1227

```

&lt;210&gt; SEQ ID NO 36

&lt;211&gt; LENGTH: 195

&lt;212&gt; TYPE: DNA

<213> ORGANISM: *Microcystis aeruginosa*

&lt;400&gt; SEQUENCE: 36

-continued

---

```

gctcaactga ctcaaatttg gagtgaagtt ttgggactgg aacgcattgg cgtaagggac 60
aacttttttg aattgggagg acattctctt ttggctaccc aggttttacc aagaattaat 120
tcagcctttg gacttgatct ttctgtgcaa attatgtttg aatcaccaac gatcgcggggc 180
attgcggggtt atatt 195

```

```

<210> SEQ ID NO 37
<211> LENGTH: 915
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 37

```

```

gctagagacg gtcatttacc cctgtctttt gctcaacaac gtttatgggt tttacattat 60
ctttcccttg atagtcgctc ctacaatacc ctggaaatat tgcaaattga tgggaatctc 120
aatctgactg tgctagagca gagtttgggg gaattaatta accgccatga aatttttaga 180
acaacattcc cactgttttc aggggaaccg attcagaaaa ttgcacttcc tagtcgtttt 240
cagttaaaag ttgataatta tcaagattta gacgaaaatg aacaatcagc taaaattcaa 300
caagtagcag aattggaagc aggacaagct tttgatttaa cgggtggggc actgattcag 360
tttaagctat tgcaattgag tcccagaagc tcggtgctgc tgttgaaaat gcaccatatt 420
atctatgatg gctggtcttt tgggattctg attcgggaat tatcggctct atacgaagca 480
tttttaaga acttagccaa tcctctccct gcgttgctta ttcagtatgc agattttgag 540
gtttggcaac gtcaatatct ctccaggtgag gtcttagata aacaactcaa ttattggcaa 600
gaacagttag caacagtctc tcctgttctt actttaccaa cggatagacc ccgtccggcg 660
atacaactt ttcagggagg agttgagcgt tttcaactgg atcaaatgt cactcaaggt 720
cttaaaaagt taggtcaaga tcaggttgca accctgttta tgacgttggg ggcgggttcc 780
ggcgttttgc tatctcgcta tagtggtcaa tctgatctga tgggtgggttc tccgatcgtc 840
aatcgtaatc aagcagcgat cgaaccttta attggctttt ttgctaacac tttggcttta 900
agaattaatt tatca 915

```

```

<210> SEQ ID NO 38
<211> LENGTH: 1185
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 38

```

```

acatacactg aattaacca tcgcgctaata cagttagccc attatttaca aacttttagc 60
gtgggagcag aagtcttagt cggtatcttc ctagaacggt ctttagagat gattatcggc 120
ttattagga ttctcaaggt aggtggtgct tatcttctc ttgatccaga ctatcccact 180
gagcgtcttc agttgatgtt agaagacagt caagttcctt ttttgattac ccacagttct 240
ttattagcaa aattgcctcc ctctcaagca actctgattt gtttagatca tatccaagag 300
cagatttctc aatattctcc agataatctt caatgtcagt taactcctgc caatttagct 360
aacgttattt atacctctgg ctctacgggt aagcctaaag ggggtgatggg tgaacataaa 420
ggtttagtta acttagctct tgctcaaat caatcttttg cagtcaacca taacagtcgt 480
gtgtgcaat ttgtctctt tagttttgat gottgtattt cagaaatttt gatgacctt 540
ggttctggag cgacgttcta tcttgcaaaa aaagatgctt tattgccagg tcagccatta 600
attgaaagggt tagtaaaaga tgggaattact catgtgactt tgccgccttc agcttttagt 660
gttttaccac aggaaccggt acgcaactta gaaaccttaa ttgtggcggg tgaggcttgt 720

```

-continued

---

```

tctcttgatt tagtgaaca atggcaatc gatagaaact tttcaatgc ctatgggcca 780
acggaagcga gtgtttgtgc cactattgga caatgttacc aagatgattt aaaggtgacg 840
attggaagc cgatcgcaa tgcctaaatt tatattttag atgccttttt acagccggtg 900
ccggtgggag tgtcaggaga gttatacatt ggtggagtgt ggggtggcaag gggctattta 960
aatcgtcctg aattaacca agaaaaattt attgctaacc cttttagtaa cgaccagat 1020
tctcggctct ataaaactgg cgacttagcg cgttatttac ccgatggtaa tattgaatat 1080
ttaggacgca ttgacaatca ggtaaaaatt cgcgggtttc gcattgagtt aggagaaatt 1140
gaagcgggtc tgagtcaatg tcccgatgtg caaaatacgg cgggtg 1185

```

```

<210> SEQ ID NO 39
<211> LENGTH: 195
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 39

```

```

gaaattctgg ctcaaatatg ggggcaagtt ctcaagatag aaagagtcag cagagaagat 60
aattttctttg aattgggggg gcattccctt tttagctacc aggtaatgtc ccgtctgcgt 120
gaaacttttc aagtcgaatt acctttgcgt agtctcttta ccgctccac tattgctgaa 180
ttggccctaa caatt 195

```

```

<210> SEQ ID NO 40
<211> LENGTH: 897
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 40

```

```

aacgacagtg ctaacctccc gttatctttt gctcaacaac gtttatggtt tctggatcaa 60
ttagaacctc acagcgcctt ttatcatgta gggggagccg taagactaga aggaacatta 120
aatattactg ccttagagca aagcttaaaa gaaattatta atcgtcatga agctttacgc 180
acaaatttta taacgattga tggcgaagcc actcaaatc ttcaccctac tattaattgg 240
cgattgtctg ttgttgattg tcaaaattta accgacctc aatctctgga aattgctgaa 300
gctgaaaagc cctttaatct tgctcaagat tgcttatttc gtgctacttt atcgtgcca 360
tcaccgctag aatatcatct actcgtgacc atgcaccata ttgttagcga tggctggcca 420
attggagtat tttttcaaga actaactcat ctttacgctg tctataatca gggtttacc 480
tcatctttta cgctattaa aatacaatat gctgattttg cggctctggca acggaattgg 540
ttacaaggty aaattttaag taatcaattg aattattggc gcgaacaatt agcaaatgct 600
cctgcttttt tacctttacc gacagataga cctaggcccg caatccaaac ttttattggt 660
tctcatcaag aatttaaact ttctcagcca ttaagccaaa aattgaatca actaagtcag 720
aagcatggag tgactttatt tatgactctc ctggctgctt ttgctacctt actttaccgt 780
tatacaggac aagcagatat tttagttggt tctcctattg ctaaccgtaa tcgtaaggaa 840
attgagggat taatcgctt ttttgtaat acattagttc tgagattgag tttagat 897

```

```

<210> SEQ ID NO 41
<211> LENGTH: 1245
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 41

```

-continued

---

```

acctatgctg aattaaatca tcaagcta atcagctagtc attacttaca aacttttagga 60
attgggccag aggtcttagt cgctatttca gtagaacgtt ctttagaaat gattatcggc 120
ttattagcca ttctcaaggc gtgtgggtgc tatctccctc ttgctcctga ctatcccact 180
gagcgtcttc agttcatggt agaagatagt caagcttctt ttttgattac ccacagtctt 240
ttattagaaa aattgccttc ttctcaagcg actctaattt gtttagatca catccaagag 300
cagatttctc aatattctcc cgataatctt caaagtgagt taactccttc caatttggct 360
aacgttattt acacctctgg ctctacgggt aagcctaaag ggggtgatgg tgaacatcgg 420
ggcttagtta acttagcgag ttctcaaatt caatcttttg cagcaaaaa taacagtcgt 480
gtactgcaat ttgcttcctt tagttttgat gcttgattt cagaaattt gatgacctt 540
ggttctggag cgactcttta tcttgctcaa aaaaatgatt tattgccagg tcagccatta 600
atggaaaggt tagaaaagaa taaaattacc catgttactt taccctctc agcttttagct 660
gttttaccaa aaaaaccgtt acccaactta caaacttta ttgtggcggg tgaggcttgt 720
cctctggatt tagtcaaaca atggctcagtc ggtagaaact ttttcaatgc ctatggcccg 780
acagaaacga gtgtttgtgc cacgattgga caatgttatc aagatgattt aaaggtcacg 840
attggaagg cgatcgctaa tgtccaaatt tatattttgg atgctttttt acaaccagta 900
cccacggag taccagggga attatacatt ggtggagtcg gagttgcgag gggttatcta 960
aatcgtcctg aattaacgce ggaaagattt attcctaac cttttgatcc cccctaacc 1020
cccctaaaa aggggggaga taagagctat gaaactttta aaaaggggga agagcaacca 1080
tcaaaactct ataaaacggg agatttagct cgttatttac ccgatggcaa tattgaatat 1140
ttaggacgca ttgacaatca ggtaaaaatt cgcggttttc gcattgagtt aggagaaatt 1200
gaagcgggtc tgagtcaatg tcccgatgtg caaatacgg cggtg 1245

```

```

<210> SEQ ID NO 42
<211> LENGTH: 196
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 42

```

```

ttacaattag ctcaaatctg gtcagagatt ttaggcatta ataatttggt tattcaggaa 60
aacttctttg aattagggcg tcattcttta ttagcagtcg gtctgatcaa tcgtattgaa 120
caaaagttag ataaacgttt accattaacc agtctttttc aaaatggaac catagcaagt 180
ctagctcaat tactag 196

```

```

<210> SEQ ID NO 43
<211> LENGTH: 681
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 43

```

```

actccatttt ttgctgttca tcccattggt ggtaatgtgc tatgttatgc cgatttagct 60
cgtaatttag gaacgaaaca gccgttttat ggattacaat cattagggct aagtgaatta 120
gaaaaaactg tagcctctat tgaagaaatg gcgatgattt atattgaagc aatacaaact 180
gttcaagcct ctggctccta ctatttagga ggttggtcaa tgggaggagt gatagctttt 240
gaaatcgccc aacaattatt gaccaagggt caagaagtgg ctttactggc tttaatagat 300
agttattctc ccagtttact taattcagtt aatagggaga aaaattctgc taattcctcg 360
acagaagaat ttaatgaaga tatcaatatt goctattctt tcatcagaga cttagcaagt 420

```

-continued

---

```

atatttaatc aagaaatctc tttctctggg agtgaacttg ctcattttac atcagacgaa 480
ttactagaca agtttattac ttggagtcaa gagacgaatc ttttgccgtc agattttggg 540
aagcagcagc ttaaacctg gtttaaagt ttccagatta atcaccaagc tttgagcagc 600
tatttcccca agacgtatct gggtagaagt gttttcttag gagcgggaaga cagttctatt 660
aaaaatcctg gttggcatca a 681

```

```

<210> SEQ ID NO 44
<211> LENGTH: 1065
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 44

```

```

agegggtctc aagacaaaa aacgatacag tttagcctct actactttgg tagctatgaa 60
gcggaattta acccgaataa atataactta ctgtttgaag gagctaaatt tggcgatcgc 120
gctggtttta cggccctttg gattcctgaa cgtcatttcc acgcttttgg tggtttttct 180
cccaatcctt cggttttggc ggcggcttta gcacgggaaa ccaaacagat tcaactgcca 240
tcaggcagtg tggttttacc gctacataat tccatccgag tcgccgaaga atgggcagtg 300
gtggacaatc tttcccaggg ccgcgcttgg attgcttttg catcgggttg gcatccccag 360
gattttgtct tggctcccca gtcctttggc caacatcggg aattgatggt ccaagaaatt 420
gaaaccgtcc agaaactttg gcgaggggaa gcgatcaccg tgccagacgg aaagggtcaa 480
agggtagagg ttaaaccta tccccaccg atgcagtecc agttaccag ctggattact 540
attgtcaata atcccagatc ctatatcaga gcaggggcca tcggtgctaa tacccttacc 600
aatctgatgg ggcaaacgt ggaagattta gcccgtaata ttgcgctata tcgtcaatct 660
ttggcagagc atggttatga tcccgcgtcg ggaacggtga cagttctcct gcatactttt 720
gttggaagg atttagaaca agttcgagaa caggctcgcc aaccttttgg gcaataacct 780
acctcctctg tcggactcct gcagaacatg gtcaagagcc agggcatgaa agtggatttt 840
gaacaattaa gagacgaaga tcgggacttt ctccctcgtt ctgcctataa acgctataca 900
gaaaccagtg ctttaattgg cacacccgaa tcctgtcgtc aaattattga tcatttgcag 960
tccatcggtg tggatgaagt ggcttgtttt attgattttg gggtagatga acaaacagtt 1020
ttggccaatt taccctatct ccagtcoccta aaagacttat atcaa 1065

```

```

<210> SEQ ID NO 45
<211> LENGTH: 99
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 45

```

```

attgatcccc ccctaacccc ccttgataag gggattgatc cccccctaac ccccttgat 60
aaggggattg atccccccct aacccccctt gataagggg 99

```

```

<210> SEQ ID NO 46
<211> LENGTH: 231
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 46

```

```

ccttatcaag gggggttagg ggggatcaa tccccttacc aaggggggtt aggggggat 60
caatccccct atcaaggggg gttaggggtt gatcaatccc cttatcaagg gggggttagg 120

```

-continued

---

```

gggtgatcaat ccccttatca agggggggtta ggggggggac aatcccctta tcaaggagag 180
ttaggggggg atcaatcccc ttatcaaggg gggtagggg gggatcaagt c 231

```

```

<210> SEQ ID NO 47
<211> LENGTH: 1146
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 47

```

```

cctgtctcag aaatgcgaga gtgggtcgaa aacactgtta gtcgcatctt ggctttccaa 60
ccagaacgcg gtttagaaat tggttgtggg acaggtttgt tactctccag ggtagcaaag 120
cattgtcttg aatattgggc aacggattat tccaaggggg cgatccagta tgttgaacgg 180
gtttgcaatg ccgttgaagg tttagaacag gttaaattac gctgtcaaat ggcagataat 240
tttgaaggta ttgcctaca tcaatttgat accgtcgtct taaattcgat tattcagtat 300
tttcccagtg tggattatct gttacaggtg cttgaagggg cgatcaacgt cattggcgag 360
cgaggtcaga tttttgtcgg ggatgtgcgg agtttaccct tattagagcc atatcatgcg 420
gctgtgcaat tagcccaagc ttctgactcg aaaactgttg aacaatggca acaacaggtg 480
cgtcaaatgt tagcagggtg agaagaactg gtcattgatc ccacattgtt cctggcttta 540
aaacaacatt ttccgcaaat tagctgggta gaaattcaac cgaaacgggg tgtggctcac 600
aatgagttaa ctcaatttcg ctatgatgac actctccatt tagagactat caataatcaa 660
gcattattga gcggaatcc aacgtaatt acctgggtaa attggcaact tgaccaactg 720
tctttaacac aaattaaaga taaattatta acagacaaac ctgaattgtg gggaaattcg 780
ggatttccta atcagcgagt tgaagaggct ctaaaaattt gggaaatggg ggaatggcc 840
cctgatgttg aaacggttga acaactcaaa aaacttctca aacaacaagt agatactggg 900
attaatcctg aacaggtttg gcaattagct gactctctcg gttacaccgc tcaccttagt 960
tgggtgggaaa gtagtcaaga cggttccttt gatgtcattt ttcagcggaa ttcagaagcg 1020
gaggactcaa aaaaattaac cttttcaaaa cttgctttct gggatgaaaa accctttaa 1080
ataaagccct ggagtgacta tactaacaac cctctgcgcg gtaagttagt ccaaaaatta 1140
attcct 1146

```

```

<210> SEQ ID NO 48
<211> LENGTH: 813
<212> TYPE: DNA
<213> ORGANISM: Microcystis aeruginosa

```

```

<400> SEQUENCE: 48

```

```

atgacaaatt atggcaaatc tatgtctcat tactatgatc tagtggtagg acataaagg 60
tataacaaag attacgccac tgaagtagaa ttcattcaca atttagttga gacttacaca 120
actgaagcca aatctatcct atacttgggc tgtggtacgg gttatcatgc cgetctttta 180
gcacagaaaag ggtattctgt acatgggtgt gatctcagtg ctgaaatgtt agagcaggct 240
aaaactcgca ttgaagatga aacaatagct tctaacttga gtttttctca aggaaatatt 300
tgtgaaatcc gtttaaatcg tcagtttaat gttgttcttg ctctatttca tgtggttaac 360
tatcaaacga ccaatcaaaa tttactggca acgtttgcaa cggttaaaaa ccatttaaaa 420
gctgggggga tttttatttg tgatgtgtcc tatgggtctt acgtactggg ggaatttaag 480
agtggccta cggcatcaat attgcgttta gaggataatt ccaatggtaa cgaagtaacc 540
tatattagtg aactaaattt ttaaoccat gaaaatatag tggaagttac tcacaattta 600

```

-continued

---

```

tgggtaacaa atcaagaaaa tcaacttcta gagaattcac gggaaacaca tcttcagcgc 660
tatcttttca agcctgaagt tgaattgttg gctgatgett gtgaactaac tgttcttgat 720
gcgatgccct ggcttgaaca acgtcctttg acaaacattc cttgtccttc agtttgtttt 780
gttattgggc ataaaacaac ccattcagct taa 813

```

```

<210> SEQ ID NO 49
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer with homology to
Microcystis aeruginosa

```

```

<400> SEQUENCE: 49

```

```

ccgacctgtg ataaacaatt c 21

```

```

<210> SEQ ID NO 50
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer with homology to
Microcystis aeruginosa

```

```

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (3)..(3)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (15)..(15)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 50

```

```

cknccdgtda traanarytc 20

```

```

<210> SEQ ID NO 51
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer with homology to
Microcystis aeruginosa

```

```

<400> SEQUENCE: 51

```

```

ttcaatatcc tggggata 18

```

```

<210> SEQ ID NO 52
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer with homology to
Microcystis aeruginosa

```

```

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (13)..(13)
<223> OTHER INFORMATION: n is a, c, g, or t

```

```

<400> SEQUENCE: 52

```

```

ytcdatrtcy tgnngrta 18

```

```

<210> SEQ ID NO 53
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:

```

-continued

---

<223> OTHER INFORMATION: oligonucleotide primer with homology to  
Microcystis aeruginosa

<400> SEQUENCE: 53

cgttggttac aggccctttc t

21

<210> SEQ ID NO 54

<211> LENGTH: 20

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: oligonucleotide primer with homology to  
Microcystis aeruginosa

<220> FEATURE:

<221> NAME/KEY: misc\_feature

<222> LOCATION: (3)..(3)

<223> OTHER INFORMATION: n is a, c, g, or t

<220> FEATURE:

<221> NAME/KEY: misc\_feature

<222> LOCATION: (9)..(9)

<223> OTHER INFORMATION: n is a, c, g, or t

<220> FEATURE:

<221> NAME/KEY: misc\_feature

<222> LOCATION: (15)..(15)

<223> OTHER INFORMATION: n is a, c, g, or t

<220> FEATURE:

<221> NAME/KEY: misc\_feature

<222> LOCATION: (18)..(18)

<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 54

mgntggytnc argcnytnws

20

<210> SEQ ID NO 55

<211> LENGTH: 18

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: oligonucleotide primer with homology to  
Microcystis aeruginosa

<400> SEQUENCE: 55

ttagacttaa gccattgg

18

<210> SEQ ID NO 56

<211> LENGTH: 18

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: oligonucleotide primer with homology to  
Microcystis aeruginosa

<220> FEATURE:

<221> NAME/KEY: misc\_feature

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

<223> OTHER INFORMATION: y is t/u or c

<220> FEATURE:

<221> NAME/KEY: misc\_feature

<222> LOCATION: (3)..(3)

<223> OTHER INFORMATION: n is a, c, g, or t

<220> FEATURE:

<221> NAME/KEY: misc\_feature

<222> LOCATION: (6)..(6)

<223> OTHER INFORMATION: y is t/u or c

<220> FEATURE:

<221> NAME/KEY: misc\_feature

<222> LOCATION: (7)..(7)

<223> OTHER INFORMATION: y is t/u or c

<220> FEATURE:

<221> NAME/KEY: misc\_feature

<222> LOCATION: (9)..(9)

<223> OTHER INFORMATION: n is a, c, g, or t

<220> FEATURE:

<221> NAME/KEY: misc\_feature

-continued

---

```

<222> LOCATION: (10)..(10)
<223> OTHER INFORMATION: w is a or t/u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (11)..(11)
<223> OTHER INFORMATION: s is g or c
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (12)..(12)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (15)..(15)
<223> OTHER INFORMATION: y is t/u or c

<400> SEQUENCE: 56

ytngayytnw sncaytgg                                     18

<210> SEQ ID NO 57
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer with homology to
Microcystis aeruginosa

<400> SEQUENCE: 57

catagaagaa tcgagacat attc                               24

<210> SEQ ID NO 58
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: oligonucleotide primer with homology to
Microcystis aeruginosa
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (5)..(5)
<223> OTHER INFORMATION: s is g or c
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (6)..(6)
<223> OTHER INFORMATION: w is a or t/u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (7)..(7)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (8)..(8)
<223> OTHER INFORMATION: s is g or c
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (9)..(9)
<223> OTHER INFORMATION: w is a or t/u
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (10)..(10)
<223> OTHER INFORMATION: r is g or a
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (13)..(13)
<223> OTHER INFORMATION: n is a, c, g, or t
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (15)..(15)
<223> OTHER INFORMATION: r is g or a
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (16)..(16)

```

-continued

<223> OTHER INFORMATION: n is a, c, g, or t  
 <220> FEATURE:  
 <221> NAME/KEY: misc\_feature  
 <222> LOCATION: (19)..(19)  
 <223> OTHER INFORMATION: r is g or a  
 <220> FEATURE:  
 <221> NAME/KEY: misc\_feature  
 <222> LOCATION: (22)..(22)  
 <223> OTHER INFORMATION: y is t/u or c

<400> SEQUENCE: 58

catnswnswr tcnarnccrt aytc

24

<210> SEQ ID NO 59  
 <211> LENGTH: 666  
 <212> TYPE: PRT  
 <213> ORGANISM: Microcystis aeruginosa

<400> SEQUENCE: 59

Met Thr Thr Gln Thr Ala Ser Ser Ala Asn Ala Leu Ala Ser Phe Asn  
 1 5 10 15  
 Gln Phe Leu Arg Asp Val Lys Ala Ile Ala Gln Pro Tyr Trp Tyr Pro  
 20 25 30  
 Thr Val Ser Asn Lys Arg Ser Phe Ser Glu Val Ile Arg Ser Trp Gly  
 35 40 45  
 Met Leu Ser Leu Leu Ile Phe Leu Ile Val Gly Leu Val Ala Val Thr  
 50 55 60  
 Ala Phe Asn Ser Phe Val Asn Arg Arg Leu Ile Asp Val Ile Ile Gln  
 65 70 75 80  
 Glu Lys Asp Ala Ser Gln Phe Ala Ser Thr Leu Thr Val Tyr Ala Ile  
 85 90 95  
 Gly Leu Ile Cys Val Thr Leu Leu Ala Gly Phe Thr Lys Asp Ile Arg  
 100 105 110  
 Lys Lys Ile Ala Leu Asp Trp Tyr Gln Trp Leu Asn Thr Gln Ile Val  
 115 120 125  
 Glu Lys Tyr Phe Ser Asn Arg Ala Tyr Tyr Lys Ile Asn Phe Gln Ser  
 130 135 140  
 Asp Ile Asp Asn Pro Asp Gln Arg Leu Ala Gln Glu Ile Glu Pro Ile  
 145 150 155 160  
 Ala Thr Asn Ala Ile Ser Phe Ser Ala Thr Phe Leu Glu Lys Ser Leu  
 165 170 175  
 Glu Met Leu Thr Phe Leu Val Val Val Trp Ser Ile Ser Arg Gln Ile  
 180 185 190  
 Ala Ile Pro Leu Met Phe Tyr Thr Ile Ile Gly Asn Phe Ile Ala Ala  
 195 200 205  
 Tyr Leu Asn Gln Glu Leu Ser Lys Ile Asn Gln Ala Gln Leu Gln Ser  
 210 215 220  
 Lys Ala Asp Tyr Asn Tyr Ala Leu Thr His Val Arg Thr His Ala Glu  
 225 230 235 240  
 Ser Ile Ala Phe Phe Arg Gly Glu Lys Glu Gln Asn Ile Ile Gln  
 245 250 255  
 Arg Arg Phe Gln Glu Val Ile Asn Asp Thr Lys Asn Lys Ile Asn Trp  
 260 265 270  
 Glu Lys Gly Asn Glu Ile Phe Ser Arg Gly Tyr Arg Ser Val Ile Gln  
 275 280 285  
 Phe Phe Pro Phe Leu Val Leu Gly Pro Leu Tyr Ile Lys Gly Glu Ile  
 290 295 300

-continued

---

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

&lt;210&gt; SEQ ID NO 60

&lt;211&gt; LENGTH: 2004

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Microcystis aeruginosa

&lt;400&gt; SEQUENCE: 60

atgacaaccc aaacagcttc tagtgccaat gcccttgctt cctttaacca atttttaagg 60

-continued

---

gatgtaaagg cgatcgccca accctattgg tatcccactg tatcaaataa aagaagcttt	120
tctgaggtta ttcgttcctg gggaaatgcta tcaactgctta tctttttgat tgtgggatta	180
gtcgccgtca cggtctttaa tagttttggt aatcgctggt taattgatgt cattattcaa	240
gaaaaagatg cgtctcaatt tgccagtaca ttaactgtct atgcgatcgg attaactctgt	300
gtaacgctgc tggcaggggt cactaaagat attcgcaaaa aaattgcctt agattgggat	360
caatggttaa acaccagat tgtagagaaa tatttttagta atcgctgcta ttataaaatt	420
aactttcaat ctgacattga taaccccgat caacgtctag cccaggaaat tgaaccgatc	480
gccacaaacg ccattagttt ctcggccact tttttggaaa aaagtttgga aatgctaact	540
tttttagtgg tagtttggtc aatttctcga cagattgcta ttccgctaatt gttttacacg	600
attatcggta attttattgc cgcctatcta aatcaagaat taagcaagat caatcaggca	660
caactgcaat caaaagcaga ttataactat gccttaacc atgttcggac tcatgcgga	720
tctattgctt tttttcgggg agaaaaagag gaacaaaata ttattcagcg acgttttcag	780
gaagttatca atgatacga aaataaaatt aactgggaaa aagggatga aatttttagt	840
cggggtatc gttccgtcat tcagttttt ccttttttag tccttgccc tttgtatatt	900
aaaggagaaa ttgattatgg acaagttgag caagcttcat tagctagttt tatgtttga	960
tcggccctgg gagaattaat tacagaattt ggtacttcag gacgttttc tagttatgta	1020
gaacgtttaa atgaattttc taatgcctta gaaactgtga ctaacaagc cgagaatgtc	1080
agcacaatta caaccataga agaaaatcat tttgccttg aacacgtcac cctagaaacc	1140
cctgactatg aaaaggtgat tgttgaggat ttatctctta ctgttcaaaa aggtgaagga	1200
ttattgattg tcgggcccag tggctcaggt aaaagtctt tattaagggc gatcgccggt	1260
ttatggaatg ctggcactgg gcgtttagtg cgtcctccc tagaagaaat tctcttttg	1320
ccccaacgc cctacattat tttgggaacc ttacggaac aattgctgta tctctaac	1380
aatagtgaga tgagcaatac cgaacttcaa gcagtattac aacaagtaa tttgcaaat	1440
gtgctaaatc ggggtgatga ctttgactcc gaaaaacct gggaaaacat tctctcctc	1500
gggtaacaac aacgcctagc ctttgctcga ttgttagtga attctccgag tttaccatt	1560
ttagatgagg cgaccagtgc cttagattta acaaatgagg ggattttata cgagcaatta	1620
caaacctgca agacaacctt tattagtgtg ggtcatcgag aaagtgtgt taattaccat	1680
caatgggttt tagaacttct tgcctgactc agttgggaac tcttaagcgt tcaagattat	1740
cgcttaaaa aagcgggaga aatgtttact aatgcttoga gtaacaattc cataacacc	1800
gatattacta tcgataatgg atcagaacca gaaatagtct attctctga aggattttc	1860
catcaggaaa tgaaactatt aacagacctt tcaactctta gcattcggag taaagccagt	1920
cgaggaagtg taccatcgc caaggtggt tttacctacc tttatgacaa aaatcctcag	1980
atattaaagt ggctcagaac ttaa	2004

&lt;210&gt; SEQ ID NO 61

&lt;211&gt; LENGTH: 27260

&lt;212&gt; TYPE: DNA

<213> ORGANISM: *Microcystis aeruginosa*

&lt;400&gt; SEQUENCE: 61

atgactatta actatgggta tctgcaagaa cccttaata aattctcaac cctagttgaa	60
ttactcgtt atcgggcaag cagtcacccg gaacgcctcg cctatatttt tctgagagac	120
ggagaaatcg aagaagctcg tttaaactat ggggaactgg atcaaaggc tagggcgatc	180

-continued

---

gccgcttata	tacaatcctt	agaagccgag	ggcgaaaggg	gtttactgct	ctatccccc	240
ggactagatt	ttatttcagc	tttttttggg	tgtttatatg	cgggagtcgt	tgccattccc	300
gcctatccac	cccgacggaa	tcaaacctt	ttgcgtttac	aggcgattat	tgccgattct	360
caagcccgat	ttaccttcac	caatgccgct	ctatttccca	gtttaaaaaa	ccaatgggct	420
aaagaccctg	aattaggagc	aatggaatgg	attgttaccg	atgaaattga	ccatcacctc	480
agggaggatt	ggctagaacc	aaccctcgaa	aaaaacagtc	tcgcttttct	acaatacacc	540
tctggttcaa	cgggaactcc	aaagggagta	atggtcagtc	accataattt	gttgattaat	600
tcagccgatt	tagatcgtgg	ttggggccat	gatcaagata	gcgtaatggt	cacttggtta	660
ccgaccttcc	atgatatggg	tctgatttat	ggggttattc	agcctttgta	caaaggattt	720
ctttgttaca	tgatgtcccc	tgccagcttt	atggaacgac	cgttacgttg	gttacaggcc	780
ctttctgata	aaaaagcaac	ccatagtgcg	gcccccaact	ttgcctacga	tctttgtgtg	840
cggaaaattc	cccctgaaaa	acgggctacg	ttagacttaa	gccattgggtg	catggcctta	900
aatggggccg	aaccctgcag	agcggaggtg	cttaaaaagt	ttcgggaggc	ttttcaagtt	960
tctggtttca	aagccacagc	cctttgtcct	ggctacgggt	tagcagaagc	caccctgaaa	1020
gttacggcgg	ttagttatga	cagtccccct	tacttttata	ccgttcaggc	taatgcttta	1080
gaaaaaata	agattgtggg	agccactgaa	accgatacca	atgtgcagac	cctcgtgggc	1140
tcgcgctgga	caacgattga	tactcaaate	gtcattgtca	atcctgaaac	cctgaaacct	1200
tgctccccctg	aaattgtcgg	cgaattttgg	gtatcagggt	caacaatcgc	ccaaggetat	1260
tggggaaaa	ctcaagagac	tcaggaaaac	tttcaagctt	atttggcaga	tacaggagcc	1320
gggccttttc	tcgcaacagg	agacttgggc	ttcattaaag	atggtgaatt	gtttatcaca	1380
ggtcggctca	aggaaattat	tctgattcga	ggacgcaata	attatcccca	ggatattgaa	1440
ttaaccgtcc	aaaatagtca	tcccgtctcg	cgccccagtt	gtggggctgc	ttttaccggt	1500
gaaaataaag	gcgaagaaaa	gctcgtggtc	gttcaggaag	tggagcgcac	ctggctccgt	1560
aaagtagata	tagatgaggt	aaaaagagcc	atcgttaaag	ctggtgtcca	ggaatatgat	1620
ttacaggttt	atgcgatcgc	gctgatcagg	actggcagtt	taccaaaaac	ctctagcggg	1680
aaaattcagc	gtcgtagctg	tcgggccaaa	tttttagagg	gaagcctgga	aattttgggc	1740
taagaaaatt	tctcgtatcg	cacttaattg	gttaaattcg	tatgtcgatt	gaaacttcca	1800
ccaattcttt	ctctccccct	aagtcocatg	ctctggattt	gaaaattoct	taaactttaa	1860
ctacatttct	caagaaagca	aattgaaatc	aatgtccaca	gaaatcccaa	acgacaaaaa	1920
acaaccgacc	ctaacgaaaa	ttcaaaaactg	gttagtggct	tacatgacag	agatgatgga	1980
agtggacgaa	gatgagattg	atctgagcgt	tccctttgat	gaatatggtc	tcgattcttc	2040
tatggcagtt	gctttgatcg	ctgatctaga	ggattgggta	cgacgagatt	tacatcgcac	2100
cctgatctac	gattatccaa	ctctagaaaa	gttggtctaaa	cagggttagtg	aaccctgaca	2160
tttttataaa	gtttgtgctt	aaaaattttg	aggaagttct	aaaatgacaa	attatggcaa	2220
atctatgtct	cattactatg	atctagtggg	aggacataaa	ggttataaca	aagattacgc	2280
cactgaagta	gaattcatte	acaatttagt	tgagacttac	acaactgaag	ccaatcttat	2340
cctatacttg	ggctgtggta	cgggttatca	tgccgctctt	ttagcacaga	aagggtatcc	2400
tgtacatggt	gttgatctca	gtgctgaaat	gttagagcag	gctaaaactc	gcattgaaga	2460
tgaacaata	gcttctaate	tgagtttttc	tcaaggaaat	atttgtgaaa	tccgtttaaa	2520

-continued

tcgtcagttt aatgttggtc ttgctctatt tcatgtgggt aactatcaaa cgaccaatca	2580
aaatttactg gaaacgtttg caacgggtta aaaccattta aaagctgggg ggatttttat	2640
ttgtgatgty tccatgggtt cttacgtact gggggaattt aagagtcggc ctacggcatc	2700
aatattgcgt ttagaggata attccaatgg taacgaagta acctatatta gtgaactaaa	2760
tttttaacc catgaaaata tagtgaagt tactcacaat ttatgggtaa caaatcaaga	2820
aaatcaactt ctagagaatt cacgggaaac acatcttcag cgctatcttt tcaagcctga	2880
agttgaattg ttggctgatg cttgtgaact aactgttctt gatcggatgc cctggcttga	2940
acaacgtcct ttgacaaaca ttcctgtcc ttcagtttgt tttgttattg ggcaaaaaac	3000
aaccattca gcttaaatc tgctaaaaaa aatccaactt acctattct ctgaaaccac	3060
acaagccatg aatacaatc aagatgcca gaccgaaat tactcaatct taaatcagtc	3120
aattccaaga cctctcaaac tgagtaatat cctattacga taagatttg cgttctcctt	3180
tgtttggat gtcagcagag gagtctctat attggctaga gaaatgtta tgcgaagagc	3240
atcagggctt cgatgtaca gttaagtac atcaaaaaat gctgaagaat atgttacgtt	3300
tgaccgatag tttgattat ctatggccag ttaaccgtga aatgcggctc atgaaagctg	3360
gggggtcaat tgaacggggc atcaccaata acattaaagc ttttctcaa tttaaagaaa	3420
ctgtaaccgt attaaattg aaaaaccgca gtgaggaatt tgaatggaac ccatcgcaat	3480
tattgttctt gcttgccgtt tccaggggc tgacaatcca gaagcttctt ggcaactcat	3540
gcgaaatggg gtggatgcca tcgccgatat tccctctgaa cgttgggata ttgagcgttt	3600
ctacgatccc acacctgcca ctgccaagaa gatgtatagt cgcacggggc gttttctaaa	3660
aaatgtcgat caatttgacc ctcaattttt ccgaatttct ccctagaag ccacctatct	3720
agatcctcaa caaagactgc tactggaagt cacctgggaa gccttagaaa atgctgccat	3780
tggtcctgaa accttagctg gttagccaac aggggttttt attggtatca gtgatgtgga	3840
ttatcatcgt ttggcttacc aaagtctac taacttgacc gcctatgtgg gtacaggcaa	3900
cagcaccagt attgccccta accgtttatc atatctgttt gatttgcgtg gccccagttt	3960
ggccgtagat accgcttctt cttcttcctt cgtcgcctt cacttggcct gtcagagttt	4020
gcaaagtcaa gaatcgaacc tctgcttagt ggggggagtt aatctcattt tgcgccaga	4080
gacaaccggt gttttttccc aagcgagaat gatcgcccc gacagtcggt gtaaaacctt	4140
tgacgcgagg gccgatggtt atgtgcgctc ggaagcctgt ggagtagtcg tacttaaacg	4200
tcttagggat gccattcagg acggcgatcg cattttagca gtgattgaag gttccgcggt	4260
gaatcaggat ggtttaagta atggactcac ggcccctaat ggccctgctc aacaggcggg	4320
gattcgtcag gccctggcaa atgccaggt aaaaaccggc cagattagct atgtcgaagc	4380
ccatggcacg gggacagaat tgggggatcc gatcgaagtt aaatctctga aagcggtttt	4440
gggtgaaaaa cgatcgctcg atcaaacctg ttgctcgggt tctgtgaaaa ccaacattgg	4500
tcatttagaa gggcgccggg gaatggcggg tctgattaaa gtcgttctct gcctacaaca	4560
ccaagaaatt ccccctaate tccaatttca aacccttaac ccctatattt ccctagctga	4620
cacagctttt gcgattccca ctacggctca accctggcgg accaaacccc ctaagtctgg	4680
tgaaaacggt gtcgaacgac gtttagcagg actcagttcc tttgggttg gggggacaaa	4740
ttcccatgty attctcagc aagcccctgt caccgttaaa aacaatcaac aaaatgggca	4800
gaagttgata gaacgtccct ggcatattgt gactttatct gccaaagatg aagaagcctt	4860
aaaagcctta gtccattggt atcaaaagta tttagctgat catcatgaaa ttcctctcgc	4920

-continued

---

tgatgtttgt	tttacggcca	atagtcggcg	atcgcacttt	aatcatcggt	taggagtagt	4980
ggctagagat	cgcttagaaa	tgttgcagaa	gtagagaaac	tttagtaacc	aagaaaggat	5040
gagagaaccg	aagagtatta	acaaaaagaa	aaaacctaaa	attgtttttc	tatttgccgg	5100
tcaaggttct	caatatgtag	gtatgggtcg	tcaactgtac	gaaacccaac	ccatctttcg	5160
ccaacacctg	gatcgctgtg	ctgaaatcct	gcgaccccat	ttagatcaac	ccctcttaga	5220
aattctttat	cctgctgacc	cagaagccga	aacagcgagt	ttttacctag	agcagactgc	5280
ctatacccaa	cccactttat	tcgcattcga	gtatgcccta	gcacagtat	ggcgttcctg	5340
gggaatagaa	ccggcggcag	taattggtca	cagtgtcggg	gaatatgtgg	cggccaccgt	5400
tgccggagcc	ttaagtctag	aagaaggatt	aacgctaatt	gccaacggg	caaaactgat	5460
gcagtctctc	ccaagaatg	ggacaatgat	cgccgttttt	gccgcagaag	agcgggttaa	5520
agctgttatt	gagccttata	ggactgatgt	agcgatcgct	gctgttaatg	gaccagaaaa	5580
ttttgttatt	tcaggaaaag	cgccgattat	tgctgagatt	atcattcatt	taacggcagc	5640
aggaatagaa	gttcgctctc	tcaaagtttc	ccatgctttt	cactcgcacc	tgttgagacc	5700
aattttagat	tccttagaac	aggaagctgc	tgctatttcc	taccaacccc	tgcaaattcc	5760
cttagttgct	aatttaacgg	gggaagtctc	accagaagga	gcaacgattg	aggctcgtta	5820
ctggcgaaat	catgcacgca	accctgtaca	attttatggg	agtatccaaa	cgctgatcga	5880
gcagaaatc	agtctttttt	tagaagttag	ccctaaaccg	actttatctc	gattgggtca	5940
acaatgttgt	ccagaaagat	cgaccacttg	gctattttcc	ctcgcctctc	ctcaagaaga	6000
agaacaaagc	ctactaaata	gtttggcgat	tctctatgat	tccaaggag	ccgaaataaa	6060
ctgggaaggg	tttaatcaaa	attatcccca	ccatttactg	gctctaccga	cctatccttt	6120
tcaacgtcaa	cgctattggc	ttgaaaccgg	taaaccgact	tctgaagaaa	caaccatgac	6180
gaccaatgcc	actaatgtcc	aagctatctc	cagccatcaa	aaacaacagg	agattctaata	6240
cacattgcaa	accctagtgg	gaaatttact	gcaattgtcc	cctgctgatg	tcaatgttca	6300
tacaccttcc	ctggagatgg	gggcagattc	cattgtcatg	gttgaggcgg	tcagacggat	6360
tgagaatacc	tataacgtta	aaattgctat	gcgtcagtta	tttgaggagt	tatctacttt	6420
agatgcttta	gctacttatt	tagctcaaaa	tccggctact	gattgccaaa	ctgctcaaat	6480
taataccgag	gtgttttctg	cgcccattgc	ctgctcaaat	aaccgatcgc	ccaatgtcgt	6540
gctgagttct	aataccaacg	gctttcaacg	tcaaacagct	tctccagggt	tttcggcgat	6600
cgccccctt	gcaggaatgg	gaggagcagg	ggaaatggga	ggagttgaag	tgctcaagt	6660
ttctgtgcca	caaaccagtg	cggtaacagc	ctcaggttca	accgtttcta	gttctgccct	6720
ggaaaacatt	atgggtcaac	agttacaact	gatggccaaa	cagttagaag	tcttgcaaac	6780
ggccaatttt	gccccagcga	ctccccgaac	cacagaaaat	tccccatctt	ccgtcagtca	6840
aaataggtca	aacggactta	cacaacagtt	aattcccccc	cagcaattag	cgcggaacct	6900
agagccaata	gccagtcgca	cccgtaaac	cagcaatcaa	gcttctgctc	ctaaaccgac	6960
agtaacagcc	actccctggg	ggccgaaaaa	accaccacaca	ggtggattca	ctccccacaca	7020
acagcaacat	ctagaggcat	taattgctcg	ctttacggaa	cgtacaaaaa	cctctaagca	7080
aattgtgcaa	agcgatcgcc	tcgcttttagc	agatagtcga	gcctcggctc	gattccgtat	7140
gtctatataa	gagatgcttt	atcccattgt	ggcccaacgt	tctcaaggat	caagaatttg	7200
ggatgtggac	ggtaatgaat	atattgatat	gacgatgggg	caaggggtaa	cgctgtttgg	7260

-continued

---

gcatcaacca gacttcatta tgtcggccct acaaagccaa ctcaactgaag gcattcatct	7320
caatccgcga tcgccaattg tgggagaagt ggccgcctta atttgtgaac taacaggagc	7380
cgaacgagct tgtttttgca actctggaac cgaagccgta atggccgcta ttcgtatcgc	7440
cagggaaca acaggtcggg gtaaaattgc cctctttgaa ggctcctatc atggacatgc	7500
ggacggaacc ctttttagga accaaattat tgataaccaa ctccactctt tccccctagc	7560
tctaggcgtt cccccagcc ttagtccga tgtggtggta ttggactatg gcagtgcgga	7620
agctctgaac tatttaciaa cccaggggca ggatttagcg gcggtcttag tagaaccaat	7680
tcaaagtggc aatcctctac tccaacccca acaatttctc caaagtctgc gacaaattac	7740
cagtcaaatg gccattgccc tgatttttga tgaatgatt acgggttttc gatcgcaccc	7800
agggggagcg caagctttat ttggagtaca ggccgatatt gccacctatg gcaaagtagt	7860
tgccggagga atgcccattg gagttattgc aggtaaggcc cattatctgg acagcattga	7920
cgggggaatg tggcgttatg gcgataaatc ctatcctggg gtggacagaa ccttttttgg	7980
gggaaccttt aatcagcacc cgtagcaat ggtagcggct agggctgtcc tgaccattt	8040
aaaggagcag gggccaggtc tgcaacaaca attaactgaa cgcactgcgg ccttagccga	8100
tacactgaat cattattttc aagccgaaga agttcctatt aaaatcgaa agtttagttc	8160
tttcttcggg tttgcccctc ctggcaattt ggatttactt ttctatcaca tggtagaaaa	8220
aggattttat gtctgggaat ggcgtaaaaca ttttctttca accgcccata cggaagccga	8280
tcttgcccaa tttgtccaag cgggtaagga tagcatcaca gaattgcgtc agggaggttt	8340
tatccccgca aaaaagcctt cctggccagt gccaacgcct caaattgatc cccccctaac	8400
ccccctgat aaggggattg atccccccct aacccccctt gataagggga ttgatcccc	8460
cctaaccccc cttgataaag ggggagatgt tgatgtcgcg cttgataaag gaggaaattc	8520
tcattctgtt agggacagta agttagggaa agggagcggg tctcaagacc aaaaaacgat	8580
acagtttagc ctctactact ttggtagcta tgaagcggaa ttaaccoga ataaatataa	8640
cttactgttt gaaggagcta aatttggcga tcgcgctggt tttacggccc tttggattcc	8700
tgaacgctat ttccacgctt ttggtggttt ttctcccaat ccttcggttt tggcggcggc	8760
tttagcacgg gaaaccaaac agattcaact gcgatcaggc agtgtggttt taccgctaca	8820
taattccate cgagtcgcg aagaatgggc agtggtgac aatctttccc agggccgcgt	8880
tggtattgct tttgcatcgg gttggcatcc ccaggatttt gtcttggtc cccagtcctt	8940
tgccaacat cgggaattga tgttccaaga aattgaaacc gtccagaaac tttggcgagg	9000
ggaagcgate accgtgccag acggaagggt tcaaagggtg gaggttaaaa cctatcccca	9060
accgatgcag tcccagttac ccagctggat tactattgtc aataatcccg ataactatat	9120
cagagcaggg gcgatcgggt ctaatatcct taccaatctg atggggcaaa gcgtggaaga	9180
tttagccctg aatattgcgc tatatcgtca atctttggca gagcatggtt atgatcccgc	9240
gtcgggaacg gtgacagttc tcctgcatac ttttgttggc aaggatttag aacaagtctg	9300
agaacaggct cgccaacctt ttgggcaata cctcacctcc tctgtcggac tcttcagaa	9360
catggtcaag agccagggca tgaaagtgga ttttgaacaa ttaagagacg aagatcggga	9420
ctttctcctc gcttctgcct ataaaogcta tacagaaacc agtgctttaa ttggcacacc	9480
cgaatcctgt cgtcaaatga ttgatcattt gcagtcctac ggtgtggatg aagtggcttg	9540
ttttattgat tttgggttag atgaacaaac agttttggcc aatttaccct atctccagtc	9600
cctaaaagac ttatatcaac ctcatctccc ccttatcaa ggggggttag ggggggatca	9660

-continued

---

atccccttat	caaggggggt	taggggggga	tcaatccctt	tatcaagggg	ggttaggggg	9720
tgatcaatcc	ccttatcaag	gggggttagg	gggtgatcaa	tccccttate	aaggggggtt	9780
agggggggat	caatcccctt	atcaaggaga	gttagggggg	gatcaatccc	cttatcaagg	9840
gggggttagg	gggatcaag	tccctctcac	cgaagcccaa	cgacaactgt	ggattttggc	9900
tcaattagga	gacaacggct	ctgtggccta	taaccaatca	gtgacattgc	aattaagtgg	9960
cccattaaat	cccgtcgcaa	tgaatcaagc	tattcaacaa	atcagcgatc	gccatgaagc	10020
gttacgaacc	aaaattaatg	cccagggaga	tagtcaagaa	atcctgcccc	aggtcgaaat	10080
taactgccct	atcttagact	tcagtcctga	ccaagcttcg	gcccacacgc	aagcagaaca	10140
atggttaaag	gaagaaagtg	aaaaaccctt	tgatttgagc	cagggttctc	tcgtgcgttg	10200
gcatctactc	aaattagaac	cagaattaca	tttgttagta	ttaacggccc	atcacattat	10260
cagtgcgggt	tggtcaatgg	gggtaatcct	tcgggaatta	ggagagtat	attcagccaa	10320
atgtcagggg	gttacggcta	atcttaaaac	cccaaaacag	tttcgagaat	tgattgaatg	10380
gcaaagccag	ccaagccaag	gggaagaact	gaaaaaacag	caagcctatt	ggttagcaac	10440
ccttgccgat	ccccctgttt	tgaatttacc	cactgacaaa	cctcgtccag	ctttaccag	10500
ttaccaagct	aatcgtcgaa	gtctaacttt	agatagccaa	tttacagaaa	aactaaagca	10560
atthagtcgt	aaacagggct	gtaccttgct	gatgacctg	ttatcggttt	ataacattct	10620
cgttcatcgt	ttgacgggac	aggatgatat	tctggtgggt	ctgccagcct	ctggacgggg	10680
gcttttagat	agtgaaggta	tggtgggta	ttgcacccat	ttttaccaa	ttcgcagtca	10740
attagcaggt	aatcccactt	ttgctgaata	tctcaacaa	atgcgggggg	ttttgtgtc	10800
ggcttatgaa	catcaggact	atccccttgc	tcttttgctc	aatcagttag	atttaccgag	10860
taataaccagt	cgctctcctt	taattgatgt	cagtttcaat	ttagaaccag	ttattaacct	10920
accctaaatg	aaaggattag	agattagttt	gttgccctca	agtgtaagtt	ttaaggatcg	10980
agatttgcat	tggaatgtga	cagaaatggg	tggagaagct	ctgattgatt	gtgactacaa	11040
tacagactta	tttaaagatg	aaacgattca	gcgttggtta	ggccattttc	aaaccttact	11100
tgaggcagtt	attaatgatt	cgcaacaaaa	tctgcgggaa	ttacccttat	taagttctgc	11160
tgaacgacaa	cagttattag	tgatttgtaa	tcaaaccaag	accgactatc	cccaagatca	11220
gtgtattcat	caattatttg	aagcgcaagt	tgaacggact	cccgatgoga	ttgoggtggg	11280
atgtgaaact	caacaattaa	cttacagtga	attaattgt	cgagccaatc	agttagcaca	11340
ttatttacia	aaattaggag	ttgggccaga	ggtcttagtc	ggtattttgg	tcgaacgttc	11400
tttagaaatg	attgtcggat	tgtagggat	tctcaaggct	gggggagcct	atgtacctct	11460
tgatcctgac	tatccccctg	aacgtcttca	atztatgtta	gaagatagtc	aattttttct	11520
cctottaacc	caacagcatt	tactggaatc	ttttgctcag	tcttcagaaa	cggtactctc	11580
caagattatt	tgtttgata	gcgactacca	aattatttcc	caggcaaaga	atattaatcc	11640
cgaaaattca	gtcacacaag	gtaatcttgc	ctatgtaatt	tatacctctg	gttcgacagg	11700
taaacccaag	ggcgtgatga	ataatcatgt	tgctattagt	aataaattgt	tatgggtaca	11760
agacacttat	cctctaacca	cagaagactg	tattttacia	aaaactcctt	ttagttttga	11820
tgtttcagtg	tggaattat	tctggccctt	actaaacgga	gcgcggttgg	tttttgccaa	11880
gccgaatggc	cataaagatg	ccagttactt	agtcaatctg	attcaagagc	aacaagtaac	11940
aacgctacat	tttgtgtctt	ctatgctaca	gctttttctg	acagaaaaag	acgtagaaaa	12000

-continued

---

atgtaatagt	cttaaacgag	tcattttag	tggtgaagcc	ctttctttag	agcttcaaga	12060
acgtttttt	gctcgttag	tctgtgaatt	acacaatctt	tatggaccga	cagaagccgc	12120
tattcatgtc	acattttgac	aatgtcaatc	agatagcaat	tgaaaacag	taccattgg	12180
tcggccgac	gctaataatc	aaatttacat	tttagactct	catcttcagc	cagtacctat	12240
tgagtaaatc	ggagaattgc	acattgggtg	ggttgggttg	gcgcggggtt	atttaaacag	12300
gcctgagtta	acggcggaga	aaattattgc	aaatccgttt	gcttcccttg	atccccctt	12360
aaacccccct	gataaggggg	gagatgagag	ctataaaaact	tttaaaaagg	gggggagagca	12420
accatcaaga	ttgtataaaa	cgggagattt	agctcgttat	ttaccgatg	gcaagattga	12480
gtatctaggg	cgattgata	atcaggtaaa	aattcgcggg	ttccggattg	aattggggga	12540
aattgaagcg	gttttctat	cccatcccca	ggtacgagaa	gcggtcgttt	tggtgagcga	12600
aagcgatcgc	tctgaaaatc	gggcttggg	cgcttatatt	gtccctaag	atcctgcttg	12660
tacgactcaa	tcattacgag	agtttgttaa	acggcagctt	cctgactata	tgatcccagc	12720
ttattggctg	atccttgaca	atttaccgtt	aaccagcaat	ggcaaaattg	atcgtcgggc	12780
ttaccgtta	cctaaccag	agttaaatcg	ttcgatagac	tatgtggctc	ccaaaaatcc	12840
taccaggag	gcgatgcgc	ctatttttgg	tcaagtttta	aaactggaaa	aagtgggaat	12900
ttatgataac	tttttgaga	tcggcggtaa	ttctttgcaa	gccactcaag	ttatttcacg	12960
cttacgagaa	agttttgccc	tagagtggcc	cttgcgtcgc	ctgtttgaac	aaccgactgt	13020
ggggattttg	gcttttagcc	taacggacat	tcatgccact	ttacaaaaat	tacaaacccc	13080
tattgatgat	ttatcaggcg	atcgcgagga	gattgaaacta	tgaaatctat	tgaaaccttt	13140
ttgtcagatt	tagccaatca	agatattaaa	ctctggatgg	acggcgatcg	cctgcgttgt	13200
aatgcacccc	agggcctatt	aacccagag	attcaaacag	aactgaaaaa	ccgtaaagca	13260
gaaatcattc	actttctcaa	tcaactgggt	tcagaggagc	aaattaatcc	tagaacgatt	13320
cttccattc	ctcgtgatgg	ccaattaccc	ctctcctttg	cccagtcgcg	actctggttc	13380
ttgtatcaat	tagaaggagc	cacgggaacc	tataacatga	caggggcctt	gagtttaagc	13440
gggctcttc	aggtcgaagc	cctcaaacaa	gccctaagaa	ctatcattca	acgccatgag	13500
ccattgcgta	ccagtttcca	atcgggtgac	ggggttccag	tcaggtgat	taatccctat	13560
cctgtttggg	aattagcgat	ggttgatttg	acaggaaagg	agacagaagc	agaaaaattg	13620
gcctatcagg	aatcccaaac	cccgtttgat	tgaccaata	gtcctttggt	gagggtaacg	13680
ctctcaaat	tacagccaga	aaagcatatt	ttattaatta	atatgcacca	tattatttcc	13740
gatggctggt	caatcgggtg	ttttgttcgt	gaattgtccc	atctctatag	ggcttttggt	13800
gcgggtaaa	aaccaacttt	accgatttta	ccaattcagt	atcggatttt	tgccgtttgg	13860
cagcgagagt	ggttacaggg	taaggtttta	gcggtcaat	tggaatattg	gaagcgacaa	13920
ttggcagatg	ctctcctct	gctggaactg	cccactgac	gcctctgtcc	cgcaatccaa	13980
acctttcaag	gcaagacaga	aagatttgag	ctagatagga	aactgaccca	agaattaaag	14040
gcattaagtc	aacagtcggg	ttgtacttta	tttatgactt	tgttggccgc	ttttggggtg	14100
gttttatccc	gttatagtgg	ccagactgat	atcgtcattg	gttcggcgat	cgccaaccgt	14160
aatcgccaag	acattgaggg	gttaattggc	ttttttgta	acactttggc	ggtgaggtta	14220
gatttatcag	aaaaaccag	ctttgccgct	tttttaaac	aagtacagga	agtcactcag	14280
gatgcctatg	agcatcaaga	cttgcccttt	gaaatgtag	tggagaatt	acaactagag	14340
cgcaaatag	accgaaatcc	tttggtacag	gtgatgtttg	ccctacaaaa	tgccggcaat	14400

-continued

---

gaaacctgga atttacctgg gttgaccatt gaagaaatgt cttgggaact tgaacctgcc 14460  
cgttttgacc tagaggttca tttatcagaa gttaacgccg gcatagctgg attctgttgc 14520  
tacaccattg atctatttga tgatgcaacg atcgcccgtc tattggaaca ttttcagaat 14580  
cttctcaggg caattattgt taatcctcaa gaatcggtaa gtttattacc cttgttgtca 14640  
gaacaggaag aaaagcaact tttagtgtat tggaatcaaa cccaagccga ttaatcccaa 14700  
gataagcttg tccatcagtt atttgaagtt caagcagcca gtcagccaga agcgatcgct 14760  
ctaactcttg aaaatcaggt tttgacctat ggagaattaa accatcgcgc caatcaatta 14820  
gctcactatc ttcagctggt aggagtcacc aaagaacaaa tcgtcggggg tttctctgga 14880  
cgttcccttg aaatggcgat cggattttta ggtatttca aagcaggagc cgctatctc 14940  
cccattgatc ctgaatatcc ctcagtacgc acccaattta ttctcgaaga tacccaactt 15000  
tcgtctctct taactcaggc agaactggca gaaaaactgc cccagactca aaacaaaatt 15060  
atctgtctag atcgggactg gccagaaatt acctcccaac cccagacaaa cctagactca 15120  
aagatagaac ctaataacct agcctattgc atctatactt ctggttcac aggacaacc 15180  
aaaggagtac tgatttcca tcaagcccta ctcaacttaa tttctggca tcaacaagcg 15240  
tttgagattg gcccttaca taaagcgacc caagtggcag gcattgcttt cgatgcaacg 15300  
gtttgggaat tgtggcccta tctgaccaca ggagcctgta ttaatctggt tccccaaat 15360  
attctgctct caccgacgga tttacgggat tggttgctta accgagaaat taccatgagt 15420  
tttggtccaa ctcttttagc tgaaaaatta ttatccttgg attggcctaa ccattcttgt 15480  
ctaaaaacc tgttactggg aggtgacaaa cttcattttt atcctgctgc gtccttccc 15540  
ttcagggtca ttaacaacta tggcccaacg gaaaatacag tggttgcgac ctctggactg 15600  
gtcaaatcat cttcatctca tcaacttggga actccgacta ttggtcgtcc cattgccaac 15660  
gtccaaatct atttattaga ccaaaccta caacctgtcc ccattggtgt accaggagaa 15720  
ttacatttag gtggggcggg tttagcgcag ggctatctca atcgtcctga gttaacggct 15780  
gaaaaattta ttgccaatcc ctttgatccc cccctaacc cccttgataa ggggggagaa 15840  
gaacctcaa aactctataa aacgggagac ttagccggtt atttaccga tggcaatgta 15900  
gaatttttgg gacgtattga caatcaggt aaaaatcggg gttttcgc atcgaactggg 15960  
gaaatcgaag ccgttttaag tcaatattc ctattagctg aaagtgtagt cgttgccaag 16020  
gaagataata ctggggataa acgctcgtg gcttatttgg ttcccgcctt gcaaatgag 16080  
gccctaccag agcaattagc ccaatggcaa agtgaataca tcagtgattg gcaaatctc 16140  
tatgaaagaa cctatagtca agggcaagac agcctagctg atctcacttt taatatcag 16200  
ggttgaata gcagttatc tcgtcaacc cttcctgctt cagaaatgcg agagtgggtc 16260  
gaaaacactg ttagtcgcat cttggcttc caaccagaac gcggtttaga aattggttgt 16320  
ggtacaggtt tgttactctc cagggtagca aagcattgtc ttgaatattg ggcaacggat 16380  
tattcccaag gggcgatcca gtatgttga cgggtttgca atgccgttga aggttttaga 16440  
caggttaaat tacgctgtca aatggcagat aatgttgaag gtattgcoct acatcaattt 16500  
gataccgtcg tcttaaatc gattattcag tattttccca gtgtggatta tctgttacag 16560  
gtgcttgaag gggcgatcaa cgtcattggc gagcaggtc agatttttgt cggggatgtg 16620  
cggagtttac ccctattaga gccatcatc gggctgtgc aattagcca agcttctgac 16680  
tcgaaaactg ttgaacaatg gcaacaacag gtgcgtcaaa gtgtagcagg tgaagaagaa 16740

-continued

---

ctggtcattg atcccacatt gttcctggct ttaaaacaac attttccgca aattagctgg	16800
gtagaaattc aaccgaaacg ggggtggtc cacaatgagt taactcaatt tcgctatgat	16860
gtcactctcc atttagagac tatcaataat caagcattat tgagcggcaa tccaacggta	16920
attacctggt taaattggca acttgaccaa ctgtctttaa cacaaattaa agataaatta	16980
ttaacagaca aacctgaatt gtggggaatt cgtggatttc ctaatcagcg agtgaagag	17040
gctctaaaaa ttgggaatg ggtggaaaat gccctgatg ttgaaacggg tgaacaactc	17100
aaaaaacttc tcaacaaca agtagatact ggtattaatc ctgaacaggt ttggcaatta	17160
gctgagtctc tcggttacac cgctcacctt agttggtggg aaagtagtca agacggttcc	17220
tttgatgtca tttttcagcg gaattcagaa cgggaggact caaaaaaatt aaccctttca	17280
aaacttgctt tctgggatga aaaacccctt aaaataaagc cctggagtga ctatactaac	17340
aaccctctgc gcggaagtt agtccaaaaa ttaattccta aagtagcaga atttctgcaa	17400
gaaaaactac ccagttatat ggttccccag cgtttgtgc tgcttgattc ccttcctttg	17460
acccccaatg gtaagggtga tcgtaaggcg ttaccttctc ctgatgcggc gactcgtgat	17520
ttagcgaaca gttttgtctt accccgcaat ccgattgaag ctcaactgac tcaaatttgg	17580
agtgaagttt tgggactgga acgcattggc gttaaggaca acttttttga attgggagga	17640
cattctcttt tggctacca ggttttatca agaattaatt cagcctttgg acttgatctt	17700
tctgtgcaaa ttatgtttga atcaccaacg atcgcgggca tgcgggtta tattcaagcg	17760
gtagattggg tcgccagga tcaagccgat agctcgtaa atcatgaaa tactgaggta	17820
gtggagtctt aagttatgac gaaaaagatt gttgaatttg tctgttatct acgggattta	17880
ggcattactt tagaagctga tgaaaaccgc ttacgctgtc aggctccga aggaattttg	17940
acccagcac tccgtcaaga aattggcgat cacaaactgg aattattaca atttttaca	18000
tgggtcaaac agtctaaaag taccgctcat ttgcctatta aacctgtcgc tagagacggg	18060
catttacccc tgtcttttgc tcaacaacgt ttatggtttt tacattatct tccccgat	18120
agtcgttctt acaataacct ggaaatattg caaattgatg ggaatctcaa tctgactgtg	18180
ctagagcaga gtttggggga attaattaac cgccatgaaa tttttagaac aacattcccc	18240
actgtttcag gggaaccgat tcagaaaatt gcacttccta gtcgttttca gttaaaagtt	18300
gataattatc aagatttaga cgaaaatgaa caatcagcta aaattcaaca agtagcagaa	18360
ttggaagcag gacaagcttt tgatttaacg gtggggccac tgattcagtt taagctattg	18420
caattgagtc ccagaagtc ggtgctgctg ttgaaaatgc accatattat ctatgatggc	18480
tggtcttttg ggattctgat tcgggaatta tcggtctat acgaagcatt tttaaagAAC	18540
ttagccaatc ctctcccctg gttgtctatt cagtatgcag attttgcggg ttggcaacgt	18600
caatatctct caggtagagg cttagataaa caactcaatt attggcaaga acagttagca	18660
acagtctctc ctgttcttac tttaccaacg gatagacccc gtcggcgat acaaaacttt	18720
cagggaggag ttgagcgttt tcaactggat caaatgtca ctcaaggtct taaaaagtta	18780
ggccaagatc aggttgcaac cctgtttatg acgttgttgg ccggtttcgg cgttttgcta	18840
tctcgttata gtggtaacatc tgatctgatg gtgggttctc cgatcgctaa tcgtaatcaa	18900
gcagcgatcg aaccttaaat tggtcttttt gctaacactt tggctttaag aattaattta	18960
tcagaaaatc ccagtttttt agaattatta gaacaagtta aacagacaac tttagagggt	19020
tatgctcacc aagacctacc ctttgagatg ttagtagaaa agctacaact tgaccgtgat	19080
ttgagcagaa atccttttagt acaagtcatg tttgcgctac aaaatacctc tcaagatact	19140

-continued

---

tggaatcttt	cggtttaaag	tattgaaagt	ttatctttat	cagtggaaga	aactgtcaga	19200
tttgatctag	aagtaaaactg	ctggcaaaaat	tcagaaggtt	tagcaataga	ttggatttac	19260
agcagagatt	tatttgacac	tgcaacaatt	gcaagaatgg	gagaacattt	tcaaaattta	19320
gttcaggcaa	tcatactcaa	tccaaaagct	acagttaaag	aacttccttt	attaacaccc	19380
aaggaacgtg	agcaattatt	aatatcttgg	aataatagca	agactgatta	tcctcaagag	19440
cagtgtattt	atcaattatt	tgaagcacia	gttgaacgga	ctccaaaggc	gatcgcagtg	19500
gtatttgagg	agcaatcatt	aacatacact	gaattaaacc	atcgcgctaa	tcagttagcc	19560
cattatttac	aaacttttagg	cgtggggagca	gaagtcttag	tcggtatttc	cctagaacgt	19620
tctttagaga	tgattatcgg	cttattaggg	attctcaagg	taggtggtgc	ttatcttcct	19680
cttgatccag	actatcccac	tgagcgtctt	cagttgatgt	tagaagacag	tcaagttcct	19740
tttttgatta	cccacagttc	tttattagca	aaattgcctc	cctctcaagc	aactctgatt	19800
tgtttagatc	atatccaaga	gcagatttct	caatattctc	cagataatct	tcaatgtcag	19860
ttaactcctg	ccaatttagc	taacgttatt	tatacctctg	gctctacggg	taagcctaaa	19920
gggggtgag	ttgaacataa	aggttttagt	aacttagctc	ttgctcaaat	tcaatctttt	19980
gcagtcaacc	ataacagtcg	tgtgctgcaa	tttgcctctt	ttagttttga	tgettgtatt	20040
tcagaaaatt	tgatgacctt	tggttctgga	gcgacgcttt	atcttgcaaa	aaaagatgct	20100
ttattgccag	gtcagccatt	aattgaacgg	ttagtaaaga	atggaattac	tcattgtgact	20160
ttgcccctt	cagcttttagt	ggttttacc	caggaaccgt	tacgcaactt	agaaacctta	20220
attgtggcgg	gtgagccttg	ttctcttgat	ttagtgaaac	aatggccaat	cgatagaaac	20280
tttttcaatg	cctatggggc	aacggaagcg	agtgtttgtg	ccactattgg	acaatgttat	20340
caagatgatt	taaagggtgac	gattggtaag	gcatcgcca	atgtccaaat	ttatatttta	20400
gatgcctttt	tacagccggt	gccggtggga	gtgtcaggag	agttatacat	tggtggagtt	20460
gggggtggca	ggggctattt	aaatcgtcct	gaattaaccc	aagaaaaatt	tattgctaatt	20520
ccttttagta	acgacccaga	ttctcggctc	tataaaactg	gcgacttagc	gcgttattta	20580
cccgatggta	atattgaata	tttaggacgc	attgacaatc	aggtaaaaat	tcgcggtttt	20640
cgcattgagt	taggagaaat	tgaagcgggt	ctgagtcaat	gtcccgatgt	gcaaaatagc	20700
gcggtgattg	tcgcgaaga	tactcctggc	gataagcgc	tagttgccta	tgtggttctt	20760
acttctgact	cccagataac	tactagcgaa	ctgctgcaat	ttttggcgaa	tcaattaccc	20820
gcctatcttg	ttcctaatac	ctttgttatt	ttagatgatt	tgcccctaac	cccagtgge	20880
aaatcgcgac	gccgttcctt	acctataccc	gaaacacaag	cgttatcaaa	tgactatatt	20940
gcccctaatt	ctcccactga	agaaattctg	gctcaaatat	gggggcaagt	tctcaagata	21000
gaaagagtca	gcagagaaga	taatttcttt	gaattggggg	ggcattccct	tttagctacc	21060
caggtaatgt	cccgtctgcg	tgaactttt	caagtccaat	tacctttgcg	tagtctcttt	21120
accgtcccca	ctattgctga	attggcccta	acaattgagc	aatctcagca	aaccattgct	21180
gctcccccca	tcctaaccag	aaacgacagt	gctaacctcc	cgttatcttt	tgctcaacaa	21240
cgtttatggt	ttctggatca	attagaacct	aacagcgcct	tttatcatgt	agggggagcc	21300
gtaagactag	aaggaacatt	aaatattact	gccttagagc	aaagcttaaa	agaaattatt	21360
aatcgtcatg	aagcttttacg	cacaaatfff	ataacgattg	atggtcaagc	caactcaatt	21420
atcaccccta	ctattaattg	gcgattgtct	gttgttgatt	gtcaaaatff	aaccgacact	21480

-continued

---

caatctctgg aaattgcgga agctgaaaag ccccttaatc ttgctcaaga ttgcttattt	21540
cgtgctactt tattcgtgcy atcacogeta gaatatcatc tactcgtgac catgcaccat	21600
attgtagcgy atggctggtc aattggagta ttttttcaag aactaactca tctttacgct	21660
gtctataate agggtttacc ctcatcttta acgcctatta aaatacaata tgctgatttt	21720
gcggctctggc aacggaattg gttacaaggt gaaattttaa gtaatcaatt gaattattgg	21780
cgcgaaacat tagcaaatgc tctgctttt ttacctttac cgacagatag acctaggccc	21840
gcaatccaaa cttttattgg ttctcatcaa gaatttaaac tttctcagcc attaagccaa	21900
aaattgaatc aactaagtca gaagcatgga gtgactttat ttatgactct cctggctgct	21960
tttgctacct tactttaccg ttatacagga caagcagata ttttagttgg ttctcctatt	22020
gctaaccgta atcgtaaagga aattgagggg ttaatcggct tttttgtaa tacattagtt	22080
ctgagattga gtttagataa tgatttaagt tttcaaaatt tgctaaacca tgtagagag	22140
gtttctttag cagcctacgc ccatcaagat ttaccttttg aaatgtagt agaagcacta	22200
caccctcaac gagatctcag tcatacccct ttatttcagg taatgtttgt tttgcaaat	22260
acaccagtgg ctgatctaga acttaaaaat gtaaaggttt gtcctctacc gatggaaaat	22320
aagactgcta aatttgattt aaccttatca atggagaatc tagaggaagg attgattggg	22380
gtttgggaat ataacaccga tctatttaat ggctcaacca ttgagcgaat gaggggacat	22440
ttgtcactt tgtagaaga tattgttgc cctccaacga agtcagtttt acggttgtct	22500
ttgctgacgc aagaggaaaa actgcaatta ttgattaaaa atcagggtgt tcaagttgat	22560
tattctcaag agcagtgcat ccatcaatta tttgaagcgc aagtgaacg gactcccgat	22620
gcgattgcyg tggattttga ggagcaatca ttaacctatg ctgaattaaa tcatcaagct	22680
aatcagttag tccattactt acaaaactta ggaattgggc cagaggtctt agtcgctatt	22740
tcagtagaac gttctttaga aatgattatc ggcttattag ccattctcaa ggcgtgtggt	22800
gcttatctcc ctcttgctcc tgactatccc actgagcgtc ttcagttcat gttagaagat	22860
agtcaagctt cttttttgat taccacagt tctttattag aaaaattgcc tcttctcaa	22920
gcgactctaa tttgtttaga tcacatccaa gagcagattt ctcaatattc tcccgataat	22980
cttcaaagty agttaaactc ttccaatttg gctaacgtta tttacacctc tggctctacg	23040
ggtaagccta aaggggtgat ggttgaacat cggggcttag ttaacttagc gaggttctca	23100
attcaatctt ttgcagtcaa aaataacagt cgtgtactgc aatttgcttc ctttagtttt	23160
gatgcttgta tttcagaaat tttgatgacc tttggttctg gagcagctct ttatcttgc	23220
caaaaaaatg atttattgcc aggtcagcca ttaatggaaa ggttagaaaa gaataaaatt	23280
acctatgta ctttaccctt ttcagcttta gctgttttac caaaaaaac gttaccaaac	23340
ttacaaactt taattgtgce ggggtgagct tgcctctggy atttagtcaa acaatggtca	23400
gtcggtagaa actttttcaa tgcctatggc ccgacagaaa cgagtgtttg tgccacgatt	23460
ggacaatggt atcaagatga tttaaaggtc acgattggtg aggcgatcgc taatgtccaa	23520
atztatattt tggatgcctt tttacaacca gtaccatcgy gagtaccagg ggaattatac	23580
attggtggag tgggagttgc gaggggttat ctaaatcgtc ctgaattaac ggcggaaaga	23640
tttattccta atccttttga tcccccccta acccccctta aaaagggggg agataagagc	23700
tatgaaactt ttaaaaaggg ggaagagcaa ccatcaaac tctataaac gggagattta	23760
gctcgttatt taccgatgg caatattgaa tatttaggac gcattgacaa tcaggtaaaa	23820
attcgcggtt ttcgattga gtttaggaaa attgaagcgy tctcgtgta atgtcccgat	23880

-continued

---

gtgcaaaata	cggcggatgat	tgtccgtgaa	gatactcctg	gcgataaacg	tttagttgcc	23940
tatgtgggttc	ttactttctga	ctcccagata	actactagcg	aactgcgta	attcttggtc	24000
aatcaattac	ctgcctatct	cgttcccaat	acctttgtta	tttagatga	ttgccccta	24060
acccccaatg	gtaaatgcga	tcgcccgttc	ttaccgcttc	ctgatgatca	gaccagaaaa	24120
aatattccta	aaattggccc	gcgtaattta	gtggaattac	aattagctca	aatctgggtca	24180
gagattttag	gcattaataa	tattgggtatt	caggaaaact	tctttgaatt	aggcgggtcat	24240
tctttattag	cagtcagctc	gatcaatcgt	attgaacaaa	agttagataa	acgtttacca	24300
taaaccagtc	tttttcaaaa	tggaaccata	gcaagtctag	ctcaattact	agcgcgaagaa	24360
acaactcagc	cagcctcttc	accggtgatt	gctatccagt	ctcaagggtga	taaaactcca	24420
ttttttgctg	ttcatcccat	tgggtggtaat	gtgctatggt	atgccgattt	agctcgtaat	24480
ttaggaacga	aacagccgtt	ttatgggatta	caatcattag	ggctaagtga	attagaaaaa	24540
actgtagcct	ctattgaaga	aatggcgatg	atztatattg	aagcaataca	aactgttcaa	24600
gcctctggtc	cctactatct	aggaggttgg	tcaatgggag	gagtgatagc	ttttgaaatc	24660
gcccacaat	tattgaccca	aggtaagaa	gttgctttac	tggctttaat	agatagttat	24720
tctcccagtt	tacttaatte	agttaatag	gagaaaaatt	ctgctaattc	cctgacagaa	24780
gaatttaatg	aagatatcaa	tattgcctat	tctttcatca	gagacttagc	aagtatattt	24840
aatcaagaaa	tctctttctc	tgggagtga	cttgctcatt	ttacatcaga	cgaattacta	24900
gacaagttta	ttacttgtag	tcaagagacg	aatcttttgc	cgtagattt	tgggaagcag	24960
cagggttaaaa	cctgggttaa	agttttccag	attaatcacc	aagctttgag	cagctattct	25020
cccagagcgt	atctgggtag	aagtgttttc	ttaggagcgg	aagacagttc	tattaaaat	25080
cctgggtggc	atcaagtaat	caatgacttg	caatctcaat	ggattagcgg	cgatcactac	25140
ggtttaatta	aaaatccagt	cctcgctgaa	aaactcaata	gctacctagc	ctaaaacttt	25200
caaaaagcct	gattattggt	taaaatgaat	gatcgttcac	cggtcagagg	acaagtatga	25260
caacccaaac	agcttctagt	gccaatgccc	ttgcttcctt	taaccaattt	ttaagggatg	25320
taaaggcgat	cgcccaacc	tattgggtatc	ccactgtatc	aaataaaga	agcttttctg	25380
aggttattcg	ttcctgggga	atgctatcac	tgcttatctt	tttgattgtg	ggattagtcg	25440
cggtcacggc	ttttaatagt	tttgtaatc	gtcgtttaat	tgatgtcatt	attcaagaaa	25500
aagatgctgc	tcaatttgcc	agtacattaa	ctgtctatgc	gatcggatta	atctgtgtaa	25560
cgctgctggc	agggttcaat	aaagatattc	gcaaaaaaat	tgccctagat	tggtatcaat	25620
ggttaaacac	ccagattgta	gagaaatatt	ttagtaatcg	tgccatttat	aaaattaact	25680
ttcaatctga	cattgataac	cccgatcaac	gtctagccca	ggaaattgaa	ccgatcgcca	25740
caaacgccat	tagtttctcg	gccacttttt	tggaaaaaag	tttgaaaatg	ctaacttttt	25800
tagtggtagt	ttggcaatt	tctcgacaga	ttgctattcc	gctaagtgtt	tacacgatta	25860
tcggtaattt	tattgccc	tatctaaatc	aagaattaag	caagatcaat	caggcacaac	25920
tgcaatcaaa	agcagattat	aactatgcct	taacccatgt	tcggactcat	gcggaatcta	25980
ttgctttttt	tcggggagaa	aaagaggaac	aaaatattat	tcagcgacgt	tttcagggaag	26040
ttatcaatga	tacgaaaaat	aaaattaact	gggaaaaagg	gaatgaaatt	tttagtcggg	26100
gctatcgttc	cgctattcag	ttttttcctt	ttttagtcct	tggccctttg	tatattaaag	26160
gagaaattga	ttatggacaa	gttgagcaag	cttcattagc	tagttttatg	ttgcatcgg	26220

-continued

---

```

ccttgggaga attaattaca gaatttggta cttcaggacg tttttctagt tatgtagaac 26280
gtttaaata ga attttcta at gccttagaaa ctgtgactaa acaagccgag aatgtcagca 26340
caattacaac catagaagaa aatcattttg cctttgaaca cgtcaccta gaaaccctg 26400
actatgaaaa ggtgattgtt gaggatttat ctcttactgt tcaaaaagg gaaggattat 26460
tgattgtcgg gccagtggt cgaggtaaaa gttctttatt aagggcgatc gccggtttat 26520
ggaatgctgg cactgggctg ttagtgcgct cccccctaga agaaattctc tttttgcccc 26580
aacgtcccta cattattttg ggaaccttac gcgaacaatt gctgtatcct ctaaccaata 26640
gtgagatgag caataccgaa cttcaagcag tattacaaca agtcaatttg caaatgtgc 26700
taaatcgggt gtagacttt gactccgaaa aacctggga aaacattctc tcctcgggtg 26760
aacaacaacg cctagccttt gctcgattgt tagtgaattc tccgagtttt accattttag 26820
atgaggcgac cagtgcctta gatttaacaa atgaggggat tttatacgag caattacaaa 26880
ctcgcaagac aacctttatt agtgtgggtc atcgagaaag tttgtttaat taccatcaat 26940
gggttttaga actttctgct gactctagtt gggaaactct aagcgttcaa gattatcgcc 27000
ttaaaaaagc gggagaaatg tttactaatg cttcgagtaa caattccata acaccgata 27060
ttactatcga taatgatca gaaccagaaa tagtctatc tcttgaagga ttttccatc 27120
aggaaatgaa actattaaca gacctatcac tctctagcat tcggagtaaa gccagtcgag 27180
ggaagtgat tacagccaag gatggttta cctacctta tgacaaaaat cctcagatat 27240
taaagtggct cagaacttaa 27260

```

---

The invention claimed is:

1. An isolated nucleic acid sequence encoding a peptide comprising a sequence with at least 95% identity to SEQ ID NO. 1.
2. An isolated nucleic acid encoding a microginin synthetase enzyme complex, wherein the isolated nucleic acid encodes the following activities:
  - a) adenylation domain (A\*) wherein, the adenylation domain comprises the nucleic acid sequence of claim 1
  - b) acyl carrier protein (ACP)
  - c) elongation module (EM) of polyketide synthases (PKS) comprising the following activities:
    - i. ketoacylsynthase (KS)
    - ii. acyl transferase (AT)
    - iii. acyl carrier protein (ACP 2)
  - d) aminotransferase (AMT)
  - e) three to five elongation modules (EM) of non-ribosomal peptide synthetases (NRPS) comprising the following activities:
    - i. condensation domain (C)
    - ii. adenylation domain (A)
    - iii. thiolation domain (T)
  - f) thioesterase (TE).
3. The isolated nucleic acid according to claim 2, additionally comprising sequences encoding the following activities or domains:
  - a) a monooxygenase (MO)
  - b) an integrated N-methyltransferase domain (MT) within one or more elongation modules (EM) of NRPS;
  - c) a non-integrated N-methyltransferase (MT),
  - d) a modifying activity (MA) wherein, said MA is selected from the group comprising the following activities:
    - halogenase, sulfatase, glycosylase, racemase, O-methyltransferase and C-methyltransferase
  - e) two or more peptide repeat spacer sequences (SP) consisting of one or more repeats of being either glycine rich or proline and leucine rich, located adjacently upstream and downstream of the MO or the MA, or both.
4. The isolated nucleic acid according to claim 2, further comprising at least one nucleic acid sequence encoding at least one protein sequence as follows:
  - a. adenylation domain (A\*) according to SEQ ID NO. 1
  - b. acyl carrier protein (ACP) according to SEQ ID NO. 2
  - c. elongation module of polyketide synthases:
    - i. ketoacylsynthase domain (KS) according to SEQ ID NO. 3
    - ii. acyl transferase domain (AT) according to SEQ ID NO. 4
    - iii. acyl carrier protein domain (ACP 2) according to SEQ ID NO. 5
  - d. aminotransferase (AMT) according to SEQ ID NO. 6
  - e. elongation modules of non-ribosomal peptide synthetases:
    - i. condensation domain (C) according to SEQ ID NO. 7
    - ii. adenylation domain (A) according to SEQ ID NO. 8
    - iii. thiolation domains (T) according to SEQ ID NO. 9
  - f. elongation modules of non-ribosomal peptide synthetases responsible for the activation and condensation of leucine:
    - i. condensation domain (C 2) according to SEQ ID NO. 10
    - ii. adenylation domain (A 2) according to SEQ ID NO. 11
    - iii. thiolation domain (T 2) according to SEQ ID NO. 12

## 135

- g. elongation modules of non-ribosomal peptide synthetases responsible for the activation and condensation of tyrosine 1 :
  - i. condensation domain (C 3) according to SEQ ID NO. 13
  - ii. adenylation domain (A 3) according to SEQ ID NO. 14
  - iii. thiolation domain (T 3) according to SEQ ID NO. 15
- h. elongation modules of non-ribosomal peptide synthetases responsible for the activation and condensation of tyrosine 2:
  - i. condensation domain (C 4) according to SEQ ID NO. 16
  - ii. adenylation domain (A 4) according to SEQ ID NO. 17
  - iii. thiolation domain (T 4) according to SEQ ID NO. 18
- i. thioesterase (TE) according to SEQ ID NO. 19
- j. two or more peptide repeat spacer sequences (SP1/SP2) according to SEQ ID NO. 21 and 22
- l. an integrated N-methyltransferase domain (MT) within the elongation module (EM) of the NRPS responsible for the activation and condensation of leucin according to SEQ ID 23 and
  - m.) a non-integrated N-methyltransferase (MT 2) according to SEQ ID NO. 24.
- 5. The isolated nucleic acid according to claim 2, further comprising at least one nucleic acid sequence as follows:
  - a) an adenylation domain (A\*) according to SEQ ID NO. 25,
  - b) acyl carrier protein (ACP) according to SEQ ID NO. 26,
  - c) elongation module of polyketide synthases encoding for the condensation of acetate:
    - i. ketoacylsynthase domain (KS) according to SEQ ID NO. 27
    - ii. acyl transferase domain (AT) according to SEQ ID NO. 28
    - iii. acyl carrier protein domain (ACP 2) according to SEQ ID NO. 29
  - d) aminotransferase (AMT) according to SEQ ID NO. 30,
  - e) elongation modules of non-ribosomal peptide synthetases encoding for the activation and condensation of alanin:
    - i. condensation domain (C) according to SEQ ID NO. 31
    - ii. adenylation domain (A) according to SEQ ID NO. 32
    - iii. thiolation domain (T) according to SEQ ID NO. 33
  - f) elongation modules of non-ribosomal peptide synthetases encoding for the activation and condensation of leucin:
    - i. condensation domain (C 2) according to SEQ ID NO. 34

## 136

- ii. adenylation domain (A 2) according to SEQ ID NO. 35
- iii. thiolation domain (T 2) according to SEQ ID NO. 36
- g) elongation modules of non-ribosomal peptide synthetases encoding for the activation and condensation of tyrosine 1:
  - i. condensation domains (C 3) according to SEQ ID NO. 37
  - ii. adenylation domains (A 3) according to SEQ ID NO. 38
  - iii. thiolation domains (T 3) according to SEQ ID NO. 39
- h) elongation modules of non-ribosomal peptide synthetases encoding for the activation and condensation of tyrosine 2:
  - i. condensation domains (C 4) according to SEQ ID NO. 40
  - ii. adenylation domains (A 4) according to SEQ ID NO. 41
  - iii. thiolation domains (T 4) according to SEQ ID NO. 42
- i) thioesterase (TE) according to SEQ ID NO. 43
- j) monooxygenase (MO) according to SEQ ID NO. 44
- k) two or more peptide repeat spacer sequences (SP1/2) according to SEQ ID NO. 45 and 46.
- l.) an integrated N-methyltransferase domain (MT) within the elongation module (EM) of the NRPS encoding for the activation and condensation of leucin according to SEQ ID 47 and
  - m.) a non-integrated N-methyltransferase (MT 2) according to SEQ ID NO. 48.
- 6. The isolated nucleic acid according to claim 2 wherein, the sequence parts of the nucleic acid encoding the microginin synthetase enzyme complex activities a) through f) are arranged upstream to downstream, respectively.
- 7. A vector comprising the nucleic acid of claim 2.
- 8. A microorganism transformed with the nucleic acid according to claim 2.
- 9. A vector according to claim 7 wherein, the vector is able to replicate autonomously.
- 10. A method of producing a microginin, comprising culturing a cell under conditions under which the cell will produce microginin, wherein said cell is transformed with a nucleic acid encoding the a recombinant microginin synthetase enzyme complex, according to claim 2, and wherein said cell does not produce the microginin in the absence of said nucleic acid, and wherein said cell is cultured in the presence of octanoic acid.
- 11. A microorganism transformed with the vector according to claim 6.

\* \* \* \* \*