

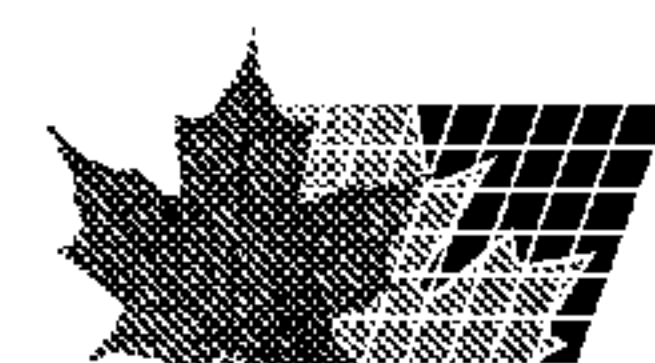


(86) Date de dépôt PCT/PCT Filing Date: 2001/03/08 (87) Date publication PCT/PCT Publication Date: 2001/09/27 (85) Entrée phase nationale/National Entry: 2002/09/05 (86) N° demande PCT/PCT Application No.: US 2001/007478 (87) N° publication PCT/PCT Publication No.: 2001/070776 (30) Priorité/Priority: 2000/03/10 (60/188,362) US	(51) Cl.Int. ⁷ /Int.Cl. ⁷ C07K 14/195 (71) Demandeur/Applicant: TRUSTEES OF TUFTS COLLEGE, US (72) Inventeurs/Inventors: LEVY, STUART B., US; BARBOSA, TERESA M., PT; ALEKSHUN, MICHAEL N., US (74) Agent: BORDEN LADNER GERVAIS LLP

(54) Titre : COMPOSITIONS MAR REGULEES RECEMMENT IDENTIFIEES
(54) Title: NIMR COMPOSITIONS AND THEIR METHODS OF USE

(57) **Abrégé/Abstract:**

Newly identified mar regulated (NIMR) genes and polypeptides are described. In addition, screening assays to identify agents that modulate NIMR activity are provided.



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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
27 September 2001 (27.09.2001)

PCT

(10) International Publication Number
WO 01/70776 A2

(51) International Patent Classification⁷: **C07K 14/195**

(21) International Application Number: PCT/US01/07478

(22) International Filing Date: 8 March 2001 (08.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/188,362 10 March 2000 (10.03.2000) US

(71) Applicant: **TRUSTEES OF TUFTS COLLEGE**
[US/US]; Ballou Hall, 4th Floor, Medford, MA 02155
(US).

(72) Inventors: **LEVY, Stuart, B.**; 144 Warren Avenue,
Boston, MA 02116 (US). **BARBOSA, Teresa, M.**; Rua de
S. Pedro 865, 4510-430 S. Pedro da Cova (PT). **ALEK-**
SHUN, Michael, N.; 44 Hawley Road, Wakefield, MA
01880 (US).

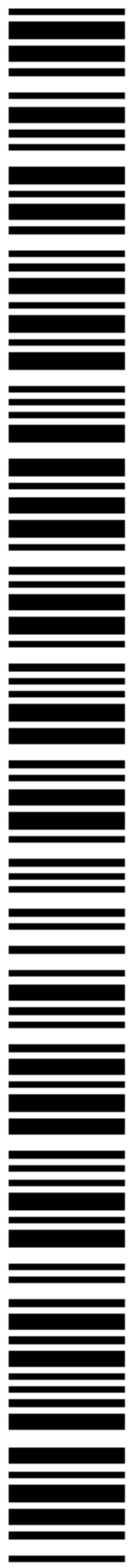
(74) Agents: **HANLEY, Elizabeth, A.** et al.; Lahive & Cock-
field, LLP, 28 State Street, Boston, MA 02109 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished
upon receipt of that report*

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*



WO 01/70776 A2

(54) Title: NIMR COMPOSITIONS AND THEIR METHODS OF USE

(57) Abstract: Newly identified *mar* regulated (NIMR) genes and polypeptides are described. In addition, screening assays to identify agents that modulate NIMR activity are provided.

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D’UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 290

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 290

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

- 1 -

NIMR COMPOSITIONS AND THEIR METHODS OF USE

Related Application Information

This application claims priority to USSN 60/188,362, filed March 10, 2000. The
5 entire contents of this application are hereby incorporated by reference.

Government Funding

This work was funded, in part, by United States Public Health Service Grant
number GM51661. The government may, therefore, have certain rights in this
10 invention.

Background of the Invention

Multidrug resistance in microbes is generally attributed to the acquisition of
multiple transposons and plasmids bearing genetic determinants for different
15 mechanisms of resistance (Gold et al. 1996. *N. Engl. J. Med.* 335:1445). However,
descriptions of intrinsic mechanisms that confer multidrug resistance have begun to
emerge. The first of these was a chromosomally encoded multiple antibiotic resistance
(*mar*) locus in *Escherichia coli* (George and Levy. 1983. *J. Bacteriol.* 155:531;
George and Levy 1983. *J. Bacteriol.* 155:541).

20 The multiple antibiotic resistance (*mar*) locus is a chromosomally encoded locus
that controls an adaptational response to antibiotics and other environmental hazards
(Aleksun, M.N. and Levy, S.B. 1997. *Antimicrob. Agents Chemother.* 10: 2067-2075).
The *mar* locus consists of two divergently positioned transcriptional units that flank a
common promoter/operator region in *E. coli* and *Salmonella typhimurium* (Aleksun and
25 Levy. 1997. *Antimicrobial Agents and Chemother.* 41: 2067) and *Shigella flexneri*
(Barbosa and Levy. 1999. 99th General Meeting of the American Society for
Microbiology (Chicago, IL). Abstract A42, p. 9). One unit encodes MarC, a putative
integral inner membrane polypeptide without any yet apparent function, but which
appears to contribute to the Mar phenotype in some strains. The other unit comprises the
30 *marRAB* operon, encoding the Mar repressor (MarR), which binds *marO* and negatively
regulates expression of *marRAB* (Cohen et al. 1994. *J. Bacteriol.* 175:1484; Martin and

- 2 -

Rosner. 1995. *Proc. Natl. Acad. Sci. USA* 92:5456; Seoane and Levy. 1995. *J. Bacteriol.* 177:530), an activator (MarA), which activates expression of MarRAB and controls expression of other genes on the chromosome, i.e., the *mar* regulon (Cohen et al. 1994. *J. Bacteriol.* 175:1484; Gambino et. al. 1993. *J. Bacteriol.* 175:2888; Seoane
5 and Levy. 1995. *J. Bacteriol.* 177:530), and a putative small polypeptide (MarB) of unknown function. MarA is a member of the XylS/AraC family of transcriptional activators (Gallegos et al. 1993. *Nucleic Acids Res.* 21:807).

The prior art has identified the *mar* regulon as comprising *acrAB*, *fumC*, *inaA*, *marA*, *marB*, *marR*, *ompF*, *ompX*, *sodA*, *tolC*, and *zwf*. Given the role of the *mar* locus
10 in controlling bacterial responses to environmental stress, identification of other genes that are regulated by MarA will be of great benefit in controlling microbes.

Summary

The present invention represents an important advance in controlling microbial
15 adaptation to environmental stress signals by newly identifying genes which respond to high constitutive levels or to overexpression of *marA* and, thus, are important in mediating resistance to and survival in environmental stresses in microbial cells. Further, the instant invention identifies genes under the control of MarA as being important in regulating virulence in microbes. Accordingly, the instant invention
20 provides novel targets (genes and polypeptides) for use in screening assays to identify compounds that modulate microbial adaptation to stress and/or virulence.

In one aspect, the invention provides a method for identifying compounds that modulate an NIMR polypeptide activity comprising:

contacting an NIMR polypeptide with a test compound under conditions which
25 allow interaction of the compound with the polypeptide;

determining the ability of the test compound to modulate the activity of an NIMR polypeptide; and

selecting those compounds that modulate the activity of the NIMR polypeptide to thereby identify compounds that modulate NIMR polypeptide activity.

- 3 -

In one embodiment, the NIMR polypeptide is selected from the group consisting of: b0357, b0447, b0853, b1448, b2530, b2889, b2948, b3469, *mdaB*, *yadG*, *yadH*, *ybjC*, *yfaE*, *yggJ*, and *yhbW*.

In another embodiment, the NIMR polypeptide activity comprises promoting the ability of a microbe to resist an environmental challenge. In another embodiment, the NIMR polypeptide is selected from the group consisting of: *aceG*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*, *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*, *rimK*, *rplE*, *srlA_2*, *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.

In another embodiment, the NIMR polypeptide activity comprises promotion of microbial virulence. In one embodiment, the NIMR polypeptide is selected from the group consisting of: *aceG*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*, *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*, *rimK*, *rplE*, *srlA_2*, *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.

In one embodiment, the step of determining comprises measuring the efflux of the test compound or a marker compound from the cell.

In one embodiment, the step of determining comprises measuring the ability of the microbe to grow or remain viable in the presence of the environmental challenge.

In one embodiment, the NIMR polypeptide is present in a microbial cell.

In another embodiment, the NIMR polypeptide is heterologous to the cell in which it is present.

In another aspect, the invention pertains to a method for identifying compounds that modulate an NIMR polypeptide activity comprising:

contacting an NIMR polypeptide with a test compound under conditions which allow interaction of the compound with the polypeptide;

determining the ability of the test compound to modulate the expression of an NIMR polypeptide; and

selecting those compounds that modulate the expression of the NIMR polypeptide to thereby identify compounds that modulate NIMR polypeptide activity.

- 4 -

In one embodiment, the NIMR polypeptide is selected from the group consisting of: b0357, b0447, b0853, b1448, b2530, b2889, b2948, b3469, *mdaB*, *yadG*, *yadH*, *ybjC*, *yfaE*, *yggJ*, and *yhbW*.

5 In one embodiment, the NIMR polypeptide is selected from the group consisting of: *aceG*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*, *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*, *rimK*, *rplE*, *srlA_2*, *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.

In one embodiment, the step of measuring comprises measuring the amount of RNA produced by the cell.

10 In one embodiment, the step of measuring comprises measuring the amount or activity of a reporter gene product produced by the cell. In another embodiment, the step of measuring comprises detecting the ability of an antibody to bind to the reporter gene product.

In one embodiment, the NIMR polypeptide is present in a cell free system.

15 In one embodiment, the step of determining comprises measuring the ability of the compound to bind to the NIMR polypeptide.

In one aspect, the invention pertains to a method for decreasing the virulence of a microbe comprising exposing the microbe to an environmental challenge and to an agent that modulates the activity of an NIMR polypeptide.

20 In another aspect, the invention pertains to a method for reducing the marA mediated transcription of an NIMR gene comprising exposing the microbe to an environmental challenge and to an agent that modulates the activity of an NIMR polypeptide.

In another aspect, the invention pertains to a method for identifying compounds
25 that modulate activity of an NIMR polypeptide in a microbe comprising: contacting an isolated NIMR nucleic acid molecule with a test compound under conditions which allow interaction of the compound with the nucleic acid molecule; determining the ability of the test compound to bind to the isolated NIMR nucleic acid molecule; and selecting those compounds that bind to the NIMR nucleic acid molecule to thereby
30 identify compounds that modulate activity of an NIMR polypeptide.

- 5 -

In one embodiment, the NIMR polypeptide is selected from the group consisting of: b0357, b0447, b0853, b1448, b2530, b2889, b2948, b3469, *mdaB*, *yadG*, *yadH*, *ybjC*, *yfaE*, *yggJ*, and *yhbW*.

5 In one embodiment, the NIMR polypeptide activity comprises promoting the ability of a microbe to resist an environmental challenge.

In one embodiment, the NIMR polypeptide is selected from the group consisting of: *aceG*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*, *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*, *rimK*, *rplE*, *srlA_2*, *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.

10 In another embodiment, the NIMR polypeptide activity comprises promotion of the virulence of a microbe.

In yet another embodiment, the NIMR polypeptide is selected from the group consisting of: *aceG*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*, *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*, *rimK*, *rplE*, *srlA_2*,
15 *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.

In one embodiment, the environmental challenge is an antibiotic compound.

In another embodiment, the environmental challenge is non-antibiotic compound.

In yet another embodiment, the non-antibiotic compound is a candidate disinfectant or antiseptic compound.

20 In yet another aspect, the invention pertains to a vaccine comprising an NIMR nucleic acid molecule or an NIMR polypeptide and a pharmaceutically acceptable carrier.

In another aspect, the invention pertains to a composition comprising a compound that modulates the activity of an NIMR polypeptide and an antibiotic.

25 In still another aspect, the invention pertains to a composition comprising a compound that modulates the activity of an NIMR polypeptide and a non-antibiotic compound.

In yet another aspect, the invention pertains to a method for reducing the virulence of a microbe in a subject suffering from a microbial infection comprising
30 administering an NIMR modulating agent to the subject such that the virulence of the microbe is reduced.

- 6 -

In another aspect, the invention pertains to a method for treating a microbial infection in a subject comprising administering an NIMR modulating agent to the subject such that the infection is treated.

5 In another aspect, the invention pertains to a method for reducing the infectivity of a microbe on a surface comprising contacting the microbe with an NIMR modulating agent such that the infectivity of the microbe is reduced.

In one embodiment, the microbe is a gram positive bacterium. In another embodiment, the microbe is a gram negative bacterium. In still another embodiment, the microbe is an acid fast bacterium.

10

Brief Description of the Drawings

Figure 1 illustrates a gene expression profile of the *Escherichia coli* MarA regulated genes.

15 Figure 2 illustrates the chromosomal distribution and location of the different members of the *mar* regulon.

Figure 3 illustrates northern blot analysis of NIMR genes.

20 ***Detailed Description***

Although the *mar* regulon was previously identified as being involved in multidrug resistance, the instant invention demonstrates that many more genes of more varied function than previously taught or suggested in the art are under the control, either directly or indirectly, of *marA*. The present invention represents an important advance
25 in controlling microbial adaptation to stress and/or virulence by newly identifying genes that respond to high constitutive expression or to the overexpression of *marA*, and referred to herein as “Newly Identified MarA Responsive (NIMR) genes.” The identification of these genes provides novel targets, both nucleic acid and polypeptide targets, for use in screening assays to identify compounds that modulate microbial
30 responses to environmental stress and, thereby, modulate microbial adaptation to their environment and/or microbial virulence. Compounds identified in such screening assays

- 7 -

can be used, e.g., to improve the activity of antibiotics, to improve the activity of non-antibiotic agents (e.g., disinfectants), and to prevent the MarA induced expression of NIMR genes.

Before further description of the invention, certain terms employed in the
5 specification, examples and appended claims are, for convenience, collected here.

I. Definitions

As used herein the term “newly identified MarA responsive gene (NIMR gene)” includes genes newly identified as responding to high constitutive expression or the
10 overexpression of MarA. Preferably, transcription of these genes is directly modulated by MarA, placing them in the *mar* regulon. As used herein, the term “regulon” includes two or more loci in two or more different operons whose expression is regulated by a common repressor or activator protein. The newly identified *mar* responsive genes are genes whose expression is controlled by MarA, but which had not, prior to the instant
15 invention, been identified as being under the control of this transcriptional activator and had not been previously identified as part of the *mar* regulon. NIMR genes can be either positively or negatively regulated by MarA and can respond directly to MarA or can respond indirectly to MarA, e.g., in response to another protein (e.g., a transcriptional regulator) that directly responds to MarA.

20 NIMR genes do not include genes identified as being part of the “prior art *mar* regulon.” As used herein, the term “prior art *mar* regulon” includes: *acrAB*, *fumC*, *inaA*, *marA*, *marB*, *marR*, *ompF*, *ompX*, *sodA*, *tolC*, and *zwf*. Preferably, NIMR genes include genes that were not previously associated with stress responses in bacteria. For example, preferred, NIMR genes had not previously been identified as being part of the *soxRS*
25 regulon (comprising the *acnA*, *acrAB*, *fumC*, *inaA*, *mdaA*, *ompF*, *ribA*, *sodA*, and *zwf* genes). Particularly preferred NIMR genes had no known function prior to their placement in the *mar* regulon in the instant invention. Exemplary NIMR genes are listed in Table 1 below:

- 8 -

Table 1.

	<i>accB*</i> (AE000404)	b0357*(AE000142)
	<i>aceE*</i> (AE000120)	b0447 (AE000151)
5	<i>aceF*</i> (AE000120)	b0853 (AE000187)
	<i>ackA*</i> (AE000318)	b1448 (AE000241)
	<i>aldA</i> (AE000239)	b2530*(AE000339)
	<i>cobU</i> (AE000291)	b2889 (AE000372)
	<i>fabB*</i> (AE000231)	b2948 (AE000377)
10	<i>fecA*</i> (AE000499)	b3469*(AE000422)
	<i>galK</i> (AE000178)	<i>mdaB</i> (AE000385)
	<i>galT</i> (AE000178)	<i>yadG</i> (AE000122)
	<i>gatA</i> (AE000298)	<i>yadH</i> (AE000122)
	<i>gatC</i> (AE000298)	<i>ybjC</i> (AE000187)
15	<i>glpD*</i> (AE000418)	<i>yfaE</i> (AE000313)
	<i>gltA</i> (AE000175)	<i>yggJ</i> (AE000377)
	<i>gshB</i> (AE000377)	<i>yhbW</i> (AE000397)
	<i>guaB*</i> (AE000337)	<i>hemB</i> (AE000143)
	<i>map</i> (AE000126)	<i>mglB</i> (AE000304)
20	<i>mtr</i> (AE000397)	<i>ndh*</i> (AE000211)
	<i>nfnB</i> (AE000163)	<i>pflB</i> (AE000192)
	<i>pgi</i> (AE000476)	<i>purA*</i> (AE000195)
	<i>ribD</i> (AE000148)	<i>rimK</i> (AE000187)
	<i>rplE*</i> (AE000408)	<i>srlA_2</i> (AE000354)
25	<i>tnaA</i> (AE000448)	<i>tnaL</i> (AE000448)
	<i>tpx</i> (AE000230)	<i>ydeA</i> (AE000250)
	<i>acnA</i> (AE000225)	<i>mdaA</i> (AE000187)
	<i>ribA</i> (AE000226)	

- 9 -

Accession numbers from the *E. coli* K-12 genome project (National Center for Biotechnology Entrez database (<http://www.ncbi.nlm.nih.gov/>)) are given in parentheses after each gene. The sequences for these exemplary NIMR genes are available on GenBank and are presented in the sequence listing part of the description. * Indicates a
5 gene that is down regulated by overexpression of MarA.

As used herein, the language "NIMR genes" also includes NIMR genes having nucleotide sequence similarity to the NIMR genes described above. For example, such genes may be derived from other organisms. For instance, the multiple antibiotic resistance (*mar*) locus, first described in the chromosome of *Escherichia coli*, is also
10 present among other genera of enteric bacteria (Cohen, S. P., Yan, W. & Levy, S. B. (1993) *J Infect. Dis.* 168, 484-488). Molecular characterization of this locus has been performed in *E. coli* (Cohen, S. P., Hachler, H. & Levy, S. B. (1993) *J Bacteriol.* 175, 1484-1492), *Salmonella typhimurium* (Sulavick, M. C., Dazer, M. & Miller, P. F. (1997) *J. Bacteriol.* 179, 1857-1866) and more recently *Shigella flexneri*.

15 NIMR gene sequences are "structurally related" to one or more of the NIMR genes set forth in the Table above. This structural relatedness can be demonstrated by sequence similarity between two NIMR nucleotide sequences or between the amino acid sequences of two NIMR polypeptides. As used herein, the term "NIMR polypeptide" includes polypeptides specified by NIMR genes. NIMR polypeptides have an NIMR
20 activity, e.g., modulate microbial adaptation to environmental stress and/ or microbial virulence.

As used herein, the term "activity" with respect to an NIMR polypeptide includes the modulation of the ability of the microbe to adapt to environmental stress and/or modulation of virulence. In addition, NIMR polypeptides may have additional activities.
25 A used herein, the term "environmental stress" or "environmental challenge" with reference to exposure of a microbe includes agents, which when contacted with a microbe, provoke a stress response in the microbe. Such agents may lead to a decrease in growth, viability, and/or virulence in individual susceptible microbial cells, but also serve as a stimulus for other microbial cells to adapt to the environmental signal e.g., by
30 acting as a selection agent for microbes that have a mutation in a target molecule affected by the stress signal. Thus, in a microbe that is equipped to deal with the

- 10 -

environmental stress (e.g., possesses a phenotype that allows growth in response to the changing environmental conditions brought about by the stress signal), the cell adapts, (e.g. retains its virulence and/or its ability to grow and remain viable when exposed to the environmental stress signal). “Environmental stress” or “environmental challenge”
5 refers to agents that come into contact with a microbe or conditions to which a microbe is exposed that present a challenge to the survival of the microbe. Microbes can contact such environmental stress signals inside (including on the surface of) or outside a mammalian body. For example, microbes (e.g., pathogenic microbes) can be contacted with environmental challenges inside the body or microbes outside the body (e.g.,
10 pathogenic microbes or environmentally important microbes residing on surfaces) can be contacted with environmental challenges outside the body to create an environmental stress.

In one embodiment an environmental stress or challenge is brought about by human intervention, e.g., by exposure of the microbe to a drug as brought about by man
15 (such as a non-antibiotic agent or an antibiotic). For example, such agents include antibiotics or non-antibiotic compounds.

In another embodiment, an environmental stress or challenge is the result of a natural process, e.g., the natural course of an infection, resulting e.g., in exposure of the microbe to natural anti-infective defenses such as antibodies; exposure of a microbe to
20 increased temperature (e.g., during infection); or exposure of the microbe to an environment lacking in cofactors or vitamins.

As used herein, the term “virulence” includes the degree of pathogenicity of an organism. The term virulence encompasses two features of an organism: its infectivity (the ability to colonize a host) and the severity of the disease produced. As used herein,
25 the term “viability” includes the capacity for cell growth. Viable cells may not actively be multiplying, e.g., may be in a quiescent state, but retain the ability to grow when conditions for growth are more favorable.. As used herein, the term “growth” includes the ability to multiply, i.e., by cell division or proliferation.

NIMR polypeptides, before their identification as being regulated by MarA may
30 have been previously found to have one or more other functions, e.g., as set forth in Table 2 below:

Table 2.

Physiological function	NIMR genes
Energy metabolism, carbon	<i>aceE, aceF, ackA, acnA, aldA, fumC, glpD, gltA, mdaA, ndh, pflB, pgi,</i>
Biosynthesis of cofactors, carriers	<i>zwf accB, cobU, hemB, gshB, ribA, ribD</i>
Carbon compound catabolism	<i>Galk, galT</i>
Amino acid biosynthesis and metabolism	<i>TnaA, tnaL</i>
Fatty acid biosynthesis	<i>fabB</i>
Nucleotide biosynthesis	<i>GuaB, purA</i>
Adaptation	<i>inaA</i>
Cell Division	<i>tolC</i>
Transport/binding proteins	<i>gatA, gatC, fecA, mglB, mtr, srlA_2, yadG, yadH, ydeA, b3469</i>
Protection responses	<i>acrA, marA, marB, marR, nfnB, sodA, tpx,</i>
Cell envelope	<i>OmpF, ompX</i>
Ribosome constituents	<i>rimK, rplE</i>
Macromolecule synthesis, modification	<i>map</i>

In isolating or identifying other NIMR molecules, sequence similarity can be shown, e.g., by generating alignments as described in more detail below.

- 5 Preferably, NIMR polypeptides share some amino acid sequence identity with a polypeptide encoded by an NIMR gene set forth in the table above. The nucleic acid sequences of the exemplary NIMR genes set forth in the table above and the polypeptides they encode are available in the art. For example, the nucleic acid and amino acid sequences of the exemplary NIMR genes set forth in Table 1 can be found
- 10 using the accession numbers listed in Table 1 at the NCBI Entrez site (<http://www.ncbi.nlm.nih.gov/>). These sequences are also presented in Appendix A.

- 12 -

As used herein, the term "nucleic acid molecule(s)" includes polyribonucleotides or polydeoxribonucleotides, which may be unmodified RNA or DNA or modified RNA or DNA. As such, "nucleic acid molecule(s)" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded, or triple-stranded regions, or a mixture of single- and double-stranded regions. In addition, "nucleic acid molecule" as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. As used herein, the term "nucleic acid molecule" also includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "nucleic acid molecule(s)" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are nucleic acid molecules as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term "nucleic acid molecule(s)" as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of nucleic acid molecules, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including, for example, simple and complex cells. "Nucleic acid molecule(s)" also embraces short nucleic acid molecules often referred to as oligonucleotide(s).

Preferred NIMR nucleic acid molecules are isolated. An "isolated" nucleic acid molecule is one that is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regard to genomic DNA, (e.g. whether chromosomal or episomal) the term "isolated" includes nucleic acid molecules which are separated from flanking DNA sequences with which the DNA is naturally associated. Preferably, an "isolated" nucleic acid molecule is free of sequences which

- 13 -

naturally flank the nucleic acid molecule (i.e., sequences located at the 5' and 3' ends of the nucleic acid molecule) in the DNA (e.g., chromosomal or episomal) of the organism from which the nucleic acid molecule is derived. As such, isolated DNA is not in its naturally occurring state (although, as described in more detail below, its sequence may be naturally occurring in the sense that has not been altered (e.g., mutated) from its naturally occurring form). For example, in various embodiments, an isolated NIMR nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb, 0.1 kb, or 0.05kb of nucleotide sequences which naturally flank the nucleic acid molecule in DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. An "isolated" NIMR nucleic acid molecule may, however, be linked to other nucleotide sequences that do not normally flank the NIMR sequences in genomic DNA (e.g., the NIMR nucleotide sequences may be linked to vector sequences). In certain preferred embodiments, an "isolated" nucleic acid molecule, such as a cDNA molecule, also may be free of other cellular material. However, it is not necessary for the NIMR nucleic acid molecule to be free of other cellular material to be considered "isolated" (e.g., an NIMR DNA molecule separated from other chromosomal DNA and inserted into another bacterial cell would still be considered to be "isolated").

As used herein, "polypeptide(s)" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds. "Polypeptide(s)" refers to both short chains, commonly referred to as peptides, oligopeptides and oligomers and to longer chains generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene encoded amino acids. "Polypeptide(s)" include those modified either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide

- 14 -

may contain many types of modifications. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains, and the amino or carboxyl termini. Modifications include, for example, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins, such as arginylation, and ubiquitination. See, for instance, *Proteins--Structure And Molecular Properties*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993) and Wold, F., *Posttranslational Protein Modifications: Perspectives and Prospects*, pgs. 1-12 in *Posttranslational Covalent Modification Of Proteins*, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., *Meth. Enzymol.* 182:626-646 (1990) and Rattan et al., *Protein Synthesis: Posttranslational Modifications and Aging*, *Ann. N.Y. Acad. Sci.* 663: 48-62 (1992). Polypeptides may be branched or cyclic, with or without branching. Cyclic, branched and branched circular polypeptides may result from post-translational natural processes and may be made by entirely synthetic methods, as well.

As used herein, an "isolated polypeptide" or "isolated protein" refers to a polypeptide or protein that is substantially free of other polypeptides, proteins, cellular material and culture medium when isolated from cells or produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. An "isolated" or "purified" polypeptide or biologically active portion thereof is substantially free of cellular material or other contaminating polypeptides from the cell or tissue source from which the NIMR polypeptide is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially

- 15 -

free of cellular material" includes preparations of NIMR polypeptide in which the polypeptide is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of NIMR polypeptide having less than about
5 30% (by dry weight) of non- NIMR polypeptide (also referred to herein as a "contaminating polypeptide"), more preferably less than about 20% of non- NIMR polypeptide, still more preferably less than about 10% of non- NIMR polypeptide, and most preferably less than about 5% non- NIMR polypeptide. When the NIMR polypeptide or biologically active portion thereof is recombinantly produced, it is also
10 preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the polypeptide preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of NIMR polypeptide in which the polypeptide is separated from
15 chemical precursors or other chemicals that are involved in the synthesis of the polypeptide. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of NIMR polypeptide having less than about 30% (by dry weight) of chemical precursors or non- NIMR chemicals, more preferably less than about 20% chemical precursors or non- NIMR chemicals, still more
20 preferably less than about 10% chemical precursors or non- NIMR chemicals, and most preferably less than about 5% chemical precursors or non- NIMR chemicals.

Preferred NIMR nucleic acid molecules and polypeptides are "naturally occurring." As used herein, a "naturally-occurring" molecule refers to an NIMR molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural
25 NIMR polypeptide). In addition, naturally or non-naturally occurring variants of these polypeptides and nucleic acid molecules which retain the same functional activity, *e.g.*, the ability to modulate adaptation to stress and/or virulence in a microbe. Such variants can be made, *e.g.*, by mutation using techniques that are known in the art. Alternatively, variants can be chemically synthesized.

- 16 -

As used herein the term "variant(s)" includes nucleic acid molecules or polypeptides that differ in sequence from a reference nucleic acid molecule or polypeptide, but retain its essential properties. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference nucleic acid molecule. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, and/or deletions in any combination. A variant of a nucleic acid molecule or polypeptide may be naturally occurring, such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of nucleic acid molecules and polypeptides may be made by mutagenesis techniques, by direct synthesis, and by other recombinant methods known to skilled artisans.

For example, it will be understood that the NIMR polypeptides described herein are also meant to include equivalents thereof. Such variants can be made, e.g., by mutation using techniques that are known in the art. Alternatively, variants can be chemically synthesized. For instance, mutant forms of NIMR polypeptides which are functionally equivalent, (e.g., have the ability to bind to DNA and to regulate transcription from an operon) can be made using techniques which are well known in the art. Mutations can include, e.g., at least one of a discrete point mutation which can give rise to a substitution, or by at least one deletion or insertion. For example, random mutagenesis can be used. Mutations can also be made by random mutagenesis or using cassette mutagenesis. For the former, the entire coding region of a molecule is mutagenized by one of several methods (chemical, PCR, doped oligonucleotide synthesis) and that collection of randomly mutated molecules is subjected to selection or screening procedures. In the latter, discrete regions of a polypeptide, corresponding either to defined structural or functional determinants are subjected to saturating or semi-random mutagenesis and these mutagenized cassettes are re-introduced into the context

- 17 -

of the otherwise wild type allele. In one embodiment, PCR mutagenesis can be used. For example, Megaprimer PCR can be used (O.H. Landt, 1990. Gene 96:125-128).

In certain embodiments, such variants have at least about 25, 30, 35, 40, 45, 50, or 60% or more amino acid identity with a naturally occurring NIMR polypeptide. In preferred embodiments, such variants have at least about 70% amino acid identity with a naturally occurring NIMR polypeptide. In more preferred embodiments, such variants have at least about 80% amino acid identity with a naturally occurring NIMR polypeptide. In particularly preferred embodiments, such variants have at least about 90% amino acid identity and preferably at least about 95% amino acid identity with a naturally occurring NIMR polypeptide.

In yet other embodiments, a nucleic acid molecule encoding a variant of an NIMR polypeptide is capable of hybridizing under stringent conditions to a nucleic acid molecule encoding a naturally occurring NIMR polypeptide.

Preferred NIMR nucleic acid molecules and NIMR polypeptides are "naturally occurring." As used herein, a "naturally-occurring" molecule refers to an NIMR polypeptide encoded by a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural NIMR polypeptide). Such molecules can be obtained from other microbes, *e.g.*, based on their sequence similarity to the NIMR molecules described herein.

In addition, naturally or non-naturally occurring variants of these polypeptides and nucleic acid molecules which retain the same functional activity, *e.g.*, the ability to modulate microbial responses to environmental stress and, thereby, modulate microbial adaptation to stress and/or microbial virulence are also within the scope of the invention. Such variants can be made, *e.g.*, by mutation using techniques which are known in the art. Alternatively, variants can be chemically synthesized.

As used herein, "heterologous DNA" or "heterologous nucleic acid" includes DNA that does not occur naturally in the cell (*e.g.*, as part of the genome) in which it is present or which is found in a location or locations in the genome that differs from that in which it occurs in nature or which is operatively linked to DNA to which it is not normally linked in nature (*i.e.*, a gene that has been operatively linked to a heterologous promoter). Heterologous DNA is 1) not naturally occurring in a particular position (*e.g.*, at a particular position in the genome) or 2) is not endogenous to the cell into which it is

- 18 -

introduced, but has been obtained from another cell. Heterologous DNA can be from the same species or from a different species. Any DNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which is expressed is herein encompassed by the term heterologous DNA.

5 The terms "heterologous protein", "recombinant protein", and "exogenous protein" are used interchangeably throughout the specification and refer to a polypeptide which is produced by recombinant DNA techniques, wherein generally, DNA encoding the polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. That is, the polypeptide is
10 expressed from a heterologous nucleic acid molecule.

 The term "interact" includes close contact between molecules that results in a measurable effect, e.g., on the conformation and/or activity of at least one of the molecules involved in the interaction. For example, a first molecule can be said to interact with a second when it inhibits the binding of the second molecule to a target
15 (e.g., a DNA or polypeptide target) to which that second molecule normally binds, or when it alters the activity of the second molecule, e.g., by steric interaction with a domain of the second molecule that mediates its activity. For example, compounds can interact with (e.g., by binding) to an NIMR polypeptide and alter the activity of the NIMR polypeptide or can interact with (e.g., by binding) to an NIMR nucleic acid
20 molecule and alter transcription of an NIMR polypeptide from that nucleic acid molecule.

 As used herein, the term "NIMR binding polypeptide" includes polypeptides that normally interact with NIMR nucleic acid molecules or NIMR polypeptides under physiological conditions in a cell, e.g., and alter transcription of an NIMR nucleic acid
25 molecule or activity of an NIMR polypeptide.

 As used herein, the term "drug" includes antibiotic agents and non-antibiotic agents. The term "drug" includes antiinfective compounds which are static or cidal for microbes, e.g., an antimicrobial compound which inhibits the growth and/or viability of a microbe. Preferred antiinfective compounds increase the susceptibility of microbes to
30 antibiotics or decrease the infectivity or virulence of a microbe. The term "drug" includes the antimicrobial agents such as disinfectants, antiseptics, and surface delivered

- 19 -

compounds. For example, antibiotics or other types of antibacterial compounds, including agents which induce oxidative stress, and organic solvents are included in this term. The term "drug" also includes biocides. The term "biocide" is art recognized and includes an agent that is thought to kill a cell "non-specifically," or a broad spectrum
5 agent whose mechanism of action is unknown as well as drugs that are known to be target-specific (e.g., triclosan). Examples of biocides include paraben, chlorbutanol, phenol, alkylating agents such as ethylene oxide and formaldehyde, halides, mercurials and other heavy metals, detergents, acids, alkalis, and chlorhexidine. Other biocidal agents include: pine oil, quaternary amine compounds such as alkyl dimethyl benzyl
10 ammonium chloride, chloroxylol, chlorhexidine, cyclohexidine, triclocarbon, and disinfectants. The term "bactericidal" refers to an agent that can kill a bacterium; "bacteriostatic" refers to an agent that inhibits the growth of a bacterium.

The term "antibiotic" is art recognized and includes antimicrobial agents synthesized by an organism in nature and isolated from this natural source, and
15 chemically synthesized drugs. The term includes but is not limited to: polyether ionophores such as monensin and nigericin; macrolide antibiotics such as erythromycin and tylosin; aminoglycoside antibiotics such as streptomycin and kanamycin; β -lactam antibiotics (having a β lactam ring) such as penicillin and cephalosporin; and polypeptide antibiotics such as subtilisin and neosporin. Semi-synthetic derivatives of
20 antibiotics, and antibiotics produced by chemical methods are also encompassed by this term. Chemically-derived antimicrobial agents such as isoniazid, trimethoprim, quinolones, fluoroquinolones and sulfa drugs are considered antibacterial drugs, and the term antibiotic includes these. It is within the scope of the screens of the present invention to include compounds derived from natural products and compounds that are
25 chemically synthesized.

The phrase "non-antibiotic agent" includes agents that are not art recognized as being antibiotics. Exemplary non-antibiotic agents include, e.g., biocides, disinfectants or antiinfectives. Non antibiotic agents also include compounds incorporated into consumer goods, e.g., for topical use on a subject or as cleaning products. In contrast to
30 the term "biocide," an antibiotic or an "anti-microbial drug approved for human use" is considered to have a specific molecular target in a microbial cell. Preferably a microbial

- 20 -

target of a therapeutic agent is sufficiently different from its physiological counterpart in a subject in need of treatment that the antibiotic or drug has minimal adverse effects on the subject.

The term "microbe" includes microorganisms expressing or made to express an NMIR polypeptide. "Microbes" are of some economic importance, e.g., are environmentally important or are important as human pathogens. For example, in one embodiment microbes cause environmental problems, e.g., fouling or spoilage, or perform useful functions such as breakdown of plant matter. In another embodiment, microbes are organisms that live in or on mammals and are medically important. Preferably microbes are unicellular and include bacteria, fungi, or protozoa. In another embodiment, microbes suitable for use in the invention are multicellular, e.g., parasites or fungi. In preferred embodiments, microbes are pathogenic for humans, animals, or plants. Microbes may be used as intact cells or as sources of materials for cell-free assays as described herein.

As used herein the term "reporter gene" includes any gene that encodes an easily detectable product that is operably linked to a promoter. By operably linked it is meant that under appropriate conditions an RNA polymerase may bind to the promoter of the regulatory region and proceed to transcribe the nucleotide sequence of the reporter gene. In certain embodiments, however, it may be desirable to include other sequences, e.g., transcriptional regulatory sequences, in the reporter gene construct. For example, modulation of the activity of the promoter may be affected by altering the RNA polymerase binding to the promoter region, or, alternatively, by interfering with initiation of transcription or elongation of the mRNA. Thus, sequences which are herein collectively referred to as transcriptional regulatory elements or sequences may also be included in the reporter gene construct. In addition, the construct may include sequences of nucleotides that alter translation of the resulting mRNA, thereby altering the amount of reporter gene product.

As used herein the term "test compound" includes agent(s) that are tested using the assays of the invention to determine whether they modulate the activity or expression of an NMIR polypeptide. More than one compound, e.g., a plurality of compounds, can

- 21 -

be tested at the same time for their ability to modulate the activity or expression of an NIMR polypeptide sequence in a screening assay.

Test compounds that can be assayed in the subject assays include antibiotic and non-antibiotic compounds. In one embodiment, test compounds include candidate
5 detergent or disinfectant compounds. Exemplary compounds which can be screened for activity include, but are not limited to, peptides, non-peptidic compounds, nucleic acids, carbohydrates, small organic molecules (e.g., polyketides), and natural product extract libraries. The term "non-peptidic compound" is intended to encompass compounds that are comprised, at least in part, of molecular structures different from naturally-occurring
10 L-amino acid residues linked by natural peptide bonds. However, "non-peptidic compounds" are intended to include compounds composed, in whole or in part, of peptidomimetic structures, such as D-amino acids, non-naturally-occurring L-amino acids, modified peptide backbones and the like, as well as compounds that are composed, in whole or in part, of molecular structures unrelated to naturally-occurring
15 L-amino acid residues linked by natural peptide bonds. "Non-peptidic compounds" also are intended to include natural products.

As used herein, the term "antibody" is intended to include immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which binds (immunoreacts with) an
20 antigen, such as Fab and F(ab')₂ fragments, single chain antibodies, intracellular antibodies, scFv, Fd, or other fragments. Preferably, antibodies of the invention bind specifically or substantially specifically to NIMR molecules. The terms "monoclonal antibodies" and "monoclonal antibody composition", as used herein, refer to a population of antibody molecules that contain only one species of an antigen binding site
25 capable of immunoreacting with a particular epitope of an antigen, whereas the term "polyclonal antibodies" and "polyclonal antibody composition" refer to a population of antibody molecules that contain multiple species of antigen binding sites capable of interacting with a particular antigen. A monoclonal antibody composition thus typically display a single binding affinity for a particular antigen with which it immunoreacts.

- 22 -

The phrase “specifically” with reference to binding, recognition, or reactivity of antibodies includes antibodies which bind to a naturally occurring NIMR molecule, but are substantially unreactive with other unrelated molecules. Preferably, such antibodies bind to an NIMR molecule (or its homolog from another species) and bind to non-NIMR molecules (or unrelated NIMR molecules) with only background binding. Antibodies specific for NIMR family molecules from one source may or may not be reactive with NIMR molecules from different species. Antibodies specific for naturally occurring NIMR molecules may or may not bind to mutant forms of such molecules. Assays to determine affinity and specificity of binding are known in the art, including competitive and non-competitive assays. Assays of interest include ELISA, RIA, flow cytometry, etc.

II. Compositions Which Modulate Antibiotic Resistance

A. Nucleic Acid Molecules

In one aspect, the invention provides isolated nucleic acid molecules comprising or consisting essentially NIMR nucleotide sequences. In another aspect, the invention provides nucleic acid molecules consisting of NIMR nucleotide sequences. Exemplary NIMR molecules are shown in Table 1.

NIMR genes have structural similarity (e.g., to the sequence shown in Table 1) and, preferably, encode NIMR polypeptides having an NIMR polypeptide activity. For example, in one embodiment, an NIMR polypeptide is capable of modulating microbial responses to environmental stress and, thereby, modulating microbial adaptation to stress and/or microbial virulence. Preferably, NIMR polypeptides modulate resistance to drugs. In one embodiment, NIMR polypeptides modulate resistance to non-antibiotic compounds. In another embodiment, NIMR polypeptides modulate resistance to antibiotics.

There is a known and definite correspondence between the amino acid sequence of a particular protein and the nucleotide sequences that can code for the protein, as defined by the genetic code (shown below). Likewise, there is a known and definite correspondence between the nucleotide sequence of a particular nucleic acid molecule

- 23 -

and the amino acid sequence encoded by that nucleic acid molecule, as defined by the genetic code.

GENETIC CODE

5	Alanine (Ala, A)	GCA, GCC, GCG, GCT
	Arginine (Arg, R)	AGA, ACG, CGA, CGC, CGG, CGT
	Asparagine (Asn, N)	AAC, AAT
	Aspartic acid (Asp, D)	GAC, GAT
	Cysteine (Cys, C)	TGC, TGT
10	Glutamic acid (Glu, E)	GAA, GAG
	Glutamine (Gln, Q)	CAA, CAG
	Glycine (Gly, G)	GGA, GGC, GGG, GGT
	Histidine (His, H)	CAC, CAT
	Isoleucine (Ile, I)	ATA, ATC, ATT
15	Leucine (Leu, L)	CTA, CTC, CTG, CTT, TTA, TTG
	Lysine (Lys, K)	AAA, AAG
	Methionine (Met, M)	ATG
	Phenylalanine (Phe, F)	TTC, TTT
	Proline (Pro, P)	CCA, CCC, CCG, CCT
20	Serine (Ser, S)	AGC, AGT, TCA, TCC, TCG, TCT
	Threonine (Thr, T)	ACA, ACC, ACG, ACT
	Tryptophan (Trp, W)	TGG
	Tyrosine (Tyr, Y)	TAC, TAT
	Valine (Val, V)	GTA, GTC, GTG, GTT
25	Termination signal (end)	TAA, TAG, TGA

An important and well known feature of the genetic code is its redundancy, whereby, for most of the amino acids used to make proteins, more than one coding nucleotide triplet may be employed (illustrated above). Therefore, a number of different nucleotide sequences may code for a given amino acid sequence. Such nucleotide sequences are considered functionally equivalent since they result in the production of the same amino acid sequence in all organisms (although certain organisms may translate some sequences more efficiently than they do others). Moreover, occasionally, a methylated variant of a purine or pyrimidine may be found in a given nucleotide sequence. Such methylations do not affect the coding relationship between the trinucleotide codon and the corresponding amino acid.

- 24 -

In view of the foregoing, the nucleotide sequence of a DNA or RNA molecule coding for an NIMR polypeptide of the invention (or a portion thereof) can be used to derive the NIMR amino acid sequence, using the genetic code to translate the DNA or RNA molecule into an amino acid sequence. Likewise, for any NIMR -amino acid
5 sequence, corresponding nucleotide sequences that can encode an NIMR protein can be deduced from the genetic code (which, because of its redundancy, will produce multiple nucleic acid sequences for any given amino acid sequence). Thus, description and/or disclosure herein of an NIMR related nucleotide sequence should be considered to also include description and/or disclosure of the amino acid sequence encoded by the
10 nucleotide sequence. Similarly, description and/or disclosure of an NIMR amino acid sequence herein should be considered to also include description and/or disclosure of all possible nucleotide sequences that can encode the amino acid sequence.

One aspect of the invention pertains to isolated nucleic acid molecules that encode NIMR proteins or biologically active portions thereof, as well as nucleic acid
15 fragments sufficient for use as hybridization probes to identify NIMR -encoding nucleic acids (e.g., NIMR mRNA) and fragments for use as PCR primers for the amplification or mutation of NIMR nucleic acid molecules. It will be understood that in discussing the uses of NIMR nucleic acid molecules, e.g., as shown in Table 1, that fragments of such nucleic acid molecules as well as full length NIMR nucleic acid molecules can be used.

20 A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having the nucleotide sequence of an NIMR molecule shown in Table 1, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, using all or portion of an NIMR nucleic acid sequence as a hybridization probe, NIMR nucleic acid molecules can be isolated using
25 standard hybridization and cloning techniques (e.g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Moreover, a nucleic acid molecule encompassing all or a portion of an NIMR nucleotide sequence can be isolated by the polymerase chain reaction (PCR) using
30 synthetic oligonucleotide primers designed based upon an NIMR nucleotide sequence (e.g., from a different species of microbe).

- 25 -

A nucleic acid molecule of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR and/or RT PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA
5 sequence analysis. Furthermore, oligonucleotides corresponding to NIMR nucleotide sequences can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of a nucleotide
10 sequence of an NIMR gene presented in Table 1 or a portion of the nucleotide sequence. A nucleic acid molecule which is complementary to the nucleotide sequence of an NIMR gene shown in Table 1 is one which is sufficiently complementary to the nucleotide sequence of an NIMR gene presented in Table 1, such that it can hybridize to the nucleotide sequence of an NIMR gene shown in Table 1, thereby forming a stable
15 duplex.

In addition to the nucleic acid molecule shown in Table 1, other NIMR nucleotide sequences of the invention are "structurally related" (i.e., share sequence identity with) the NIMR nucleotide sequence of the NIMR molecules listed in Table 1. Such sequence similarity can be shown, e.g., by optimally aligning the NIMR nucleotide
20 sequence with a putative NIMR nucleotide sequence using an alignment program for purposes of comparison and comparing corresponding positions. In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence of one of the molecules listed in Table 1.

In still another preferred embodiment, an isolated nucleic acid molecule of the
25 present invention comprises a nucleotide sequence which is at least about 25, 30, 35, 40, 45, 50, or 60% or more homologous to a naturally occurring NIMR polypeptide. In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which is at least about 25, 30, 35, 40, 45, 50, or 60% or more amino acid identity with a naturally occurring NIMR polypeptide. In another embodiment, an
30 isolated nucleic acid molecule of the invention comprises a nucleotide sequence which is at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to a

- 26 -

nucleotide sequence (e.g., to the entire length of a nucleotide sequence) of an NIMR molecule shown in Table 1 or a portion thereof.

In other embodiments, a nucleic acid molecule of the invention has at least 25, 30, 35, 40, 45, 50, 60, or 70% identity, more preferably 80% identity, and even more preferably 90% identity with a nucleic acid molecule comprising: at least about 100, 200, 300, 400, 500, 600, or at about 700 nucleotides of an NIMR molecule listed in Table 1.

Sequence similarity can be shown, e.g., by optimally aligning NIMR nucleotide or amino acid sequences for purposes of comparison using an alignment program and comparing corresponding positions of the sequences. To determine the degree of similarity between sequences, they can be aligned for optimal comparison purposes (e.g., gaps may be introduced in the sequence of one polypeptide or nucleic acid molecule for optimal alignment with the other polypeptide or nucleic acid molecule with which they are to be compared). The amino acid residues or bases at a given position are then compared with the corresponding amino acid residue or base in the sequence with which they are being compared. When a position in one sequence is occupied by the same amino acid residue or by the same base as the corresponding position in the other sequence, then the sequences are identical at that position. If amino acid residues are not identical, they may be similar. As used herein, an amino acid residue is "similar" to another amino acid residue if the two amino acid residues are members of the same family of residues having similar side chains. Families of amino acid residues having similar side chains have been defined in the art (see, for example, Altschul et al. 1990. *J. Mol. Biol.* 215:403) including basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan.) The degree (percentage) of similarity between sequences, therefore, is a function of the number of identical or similar positions shared by two sequences (*i.e.*, % homology = # of identical or similar

- 27 -

positions/total # of positions x 100). Alignment strategies are well known in the art; see, for example, Altschul et al. *supra* for optimal sequence alignment.

Nucleic acid molecules that exist as an active functional unit, e.g., mRNA molecules, will be expected to have a higher degree of structural identity among
5 homologs. It will be understood that among divergent organisms, there will be a lower degree of structural relatedness among the nucleic acid molecules that encode functional homologs.

Preferably, NIMR polypeptides share some amino acid sequence similarity with a polypeptide encoded by an NIMR gene of a molecule listed in Table 1. The nucleic
10 acid and/or amino acid sequences of an NIMR gene or polypeptide (e.g., as provided above) can be used as "query sequence" to perform a search against databases (e.g., either public or private such as <http://www.tigr.org>) to, for example, identify other NIMR genes (or polypeptides) having related sequences. For example, such searches can be performed, e.g., using the NBLAST and XBLAST programs (version 2.0) of
15 Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to the above NIMR nucleic acid molecules. BLAST polypeptide searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to NIMR polypeptide
20 molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

25 However, it will be understood that the level of sequence identity among microbial genes, even though members of the same family, is not necessarily high. This is particularly true in the case of divergent genomes where the level of sequence identity may be low, e.g., less than 20% (e.g., *B. burgdorferi* as compared e.g., to *B. subtilis*). Accordingly, structural similarity among NIMR- molecules can also be determined
30 based on "three-dimensional correspondence" of amino acid residues. As used herein, the language "three-dimensional correspondence" is meant to includes residues which

- 28 -

spatially correspond, e.g., are in the same functional position of an NIMR polypeptide member as determined, e.g., by x-ray crystallography, but which may not correspond when aligned using a linear alignment program. The language "three-dimensional correspondence" also includes residues which perform the same function, e.g., bind to
5 DNA or bind the same cofactor, as determined, e.g., by mutational analysis.

Nucleic acid molecules that differ in nucleotide sequence from those NIMR molecules listed in Table 1 due to degeneracy of the genetic code, and thus encode the same NIMR protein as are encompassed by the invention. Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide
10 sequence encoding a protein having an amino acid sequence of an NIMR molecule listed in Table 1.

In addition to the nucleotide sequences of the NIMR molecules shown in Table 1, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of a given NIMR polypeptide may exist
15 within a population of organisms. Such nucleotide variations and resulting amino acid polymorphisms in NIMR genes that are the result of natural allelic variation and that do not alter the functional activity of an NIMR polypeptide are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding functional NIMR polypeptides but
20 which have a nucleotide sequence which differs from an NIMR nucleotide sequence of a molecule listed in Table 1 are intended to be within the scope of the invention. Nucleic acid molecules encoding functional homologs of the NIMR proteins listed in Table 1, e.g., from different species, and thus which have a nucleotide sequence which differs from the NIMR sequence of the NIMR molecules listed in Table 1 are intended to be
25 within the scope of the invention. Given the list of NIMR genes set forth in Table 1, NIMR homologs can be readily identified by one of ordinary skill in the art, e.g., by structural similarity to the NIMR nucleotide sequences provided using standard techniques.

For example, NIMR nucleic acid molecules can be identified as being
30 structurally similar to the exemplary NIMR gene set forth herein based on their ability to hybridize to the nucleic acid molecule listed in Table 1 under stringent conditions. For

- 29 -

example, an NIMR DNA can be isolated from a DNA library using all or portion of a nucleotide sequence of an NIMR molecule from Table 1 as a hybridization probe and standard hybridization techniques (e.g., as described in Sambrook, J., *et al. Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989*; Cohen et al. 1993. *J. of Infectious Diseases. 168:484*)).

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 30%, 40%, 50%, or 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% homologous to each other typically remain hybridized to each other. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of a molecule from Table 1 or its complement corresponds to a naturally-occurring nucleic acid molecule. Such stringent conditions are known to those skilled in the art and can be found e.g., in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Conditions for hybridizations are largely dependent on the melting temperature T_m that is observed for half of the molecules of a substantially pure population of a double-stranded nucleic acid. T_m is the temperature in °C at which half the molecules of a given sequence are melted or single-stranded. For nucleic acids of sequence 11 to 23 bases, the T_m can be estimated in degrees C as $2(\text{number of A+T residues}) + 4(\text{number of C+G residues})$. Hybridization or annealing of nucleic acid molecules should be conducted at a temperature lower than the T_m , e.g., 15°C, 20°C, 25°C or 30°C lower than the T_m . The effect of salt concentration (in M of NaCl) can also be calculated, see for example, Brown, A., "Hybridization" pp. 503-506, in *The Encyclopedia of Molec. Biol.*, J. Kendrew, Ed., Blackwell, Oxford (1994).

- 30 -

In addition, NIMR genes can be identified by overexpressing transcriptional activators related to MarA in other microbes and identifying the genes whose expression is controlled by overexpression of the MarA homolog, using techniques similar to those set forth in the instant examples.

5 Moreover, the nucleic acid molecules of the invention can comprise only a portion of a full length NIMR nucleic acid sequence. For example a fragment can be used as a probe or primer or a fragment encoding a biologically active portion of an NIMR protein. The nucleotide sequence of the NIMR genes allows for the generation of probes and primers designed for use in identifying and/or cloning other NIMR
10 polypeptides, as well as NIMR homologues from other species. The probe/primer typically comprises a substantially purified oligonucleotide. In one embodiment, the oligonucleotide comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, 75, or 100 consecutive nucleotides of a sense
15 sequence of an NIMR molecule from Table 1 or of a naturally occurring allelic variant or mutant thereof. In another embodiment, a nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 200, 300, 400, 500, 600 or 700 nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of Table 1 or the complement thereof.

20 Moreover, a nucleic acid molecule encompassing all or a portion of an NIMR gene can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon the sequence of an NIMR molecule listed in Table 1. For example, RNA can be isolated from cells (*e.g.*, by the guanidinium-thiocyanate extraction procedure of Chirgwin *et al.* (1979) *Biochemistry* 18: 5294-5299) and cDNA can be
25 prepared using reverse transcriptase (*e.g.*, Moloney MLV reverse transcriptase, available from Gibco/BRL, Bethesda, MD; or AMV reverse transcriptase, available from Seikagaku America, Inc., St. Petersburg, FL). Synthetic oligonucleotide primers for PCR amplification can be designed based upon an NIMR nucleotide sequence. A nucleic acid molecule of the invention can be amplified using cDNA or, alternatively,
30 genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be sequenced

- 31 -

directly or cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to an NIMR nucleotide sequence can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

5 In addition to naturally-occurring allelic variants of NIMR sequences that may exist in a population, the skilled artisan will further appreciate that minor changes may be introduced by mutation into an NIMR nucleotide sequences, *e.g.*, of a molecule listed in Table 1, thereby leading to changes in the amino acid sequence of the encoded polypeptide, without altering the functional activity of an NIMR polypeptide. For
10 example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues may be made in the sequence of an NIMR molecule of Table 1. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of an NIMR nucleic acid molecule (*e.g.*, the sequence of an NIMR molecule listed in Table 1) without altering the functional activity of an NIMR molecule.
15 Exemplary residues which are non-essential and, therefore, amenable to substitution, can be identified by one of ordinary skill in the art, *e.g.*, by performing an amino acid alignment of NIMR molecules (*e.g.*, NIMR homologs from different species) and determining residues that are not conserved or by alanine scanning mutagenesis. Such residues, because they have not been conserved, are more likely amenable to
20 substitution.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding NIMR proteins that contain changes in amino acid residues that are not essential for an NIMR activity. Such NIMR proteins differ in amino acid sequence from an NIMR molecule listed in Table 1, yet retain an inherent NIMR activity. An isolated
25 nucleic acid molecule encoding a non-natural variant of an NIMR polypeptide can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of an NIMR molecule of Table 1 such that one or more amino acid substitutions, additions or deletions are introduced into the encoded polypeptide. Mutations can be introduced into an NIMR molecule by standard techniques, such as
30 site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more non-essential amino acid residues. A

- 32 -

"conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art, including basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid),
5 uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a nonessential amino acid residue in an NIMR polypeptide
10 is preferably replaced with another amino acid residue from the same side chain family.

Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an NIMR coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for activity, to identify mutants that retain functional activity. Following mutagenesis, the encoded NIMR mutant polypeptide can
15 be expressed recombinantly in a host cell and the functional activity of the mutant polypeptide can be determined using assays available in the art for assessing an NIMR activity.

Yet another aspect of the invention pertains to isolated nucleic acid molecules encoding an NIMR fusion polypeptide. Such nucleic acid molecules, comprising at least
20 a first nucleotide sequence encoding a full-length (an entire) NIMR protein, polypeptide or peptide having an NIMR activity operatively linked to a second nucleotide sequence encoding a non- NIMR protein, polypeptide or peptide, can be prepared by standard recombinant DNA techniques.

In addition to the nucleic acid molecules encoding NIMR proteins described
25 above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a polypeptide, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can
30 hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire NIMR coding strand, or only to a portion thereof. In one

- 33 -

embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding NIMR. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is
5 antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding NIMR. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids.

With the coding strand sequences encoding NIMR molecules disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson
10 and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of NIMR mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of NIMR mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of NIMR mRNA. An antisense oligonucleotide can
15 be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to
20 increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-
30 methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine,

- 34 -

uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense
5 nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest).

The antisense nucleic acid molecules of the invention are typically administered
10 to a subject or generated *in situ* such that they hybridize with or bind to cellular nucleic acid molecules to thereby inhibit expression of the polypeptide, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the
15 major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site.

Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens
20 expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a
25 strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier et al. (1987) *Nucleic Acids*.
30 *Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-

- 35 -

methyribonucleotide (Inoue et al. (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are
5 capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave NIMR mRNA transcripts to thereby inhibit translation of NIMR mRNA. A ribozyme having specificity for an NIMR -encoding nucleic acid can be designed based upon the
10 nucleotide sequence of SEQ ID NO:1. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an NIMR-encoding mRNA. See, e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742. Alternatively, NIMR mRNA can be used to select a catalytic RNA having a
15 specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of NIMR (e.g., the NIMR promoter and/or enhancers) to form triple helical structures that prevent transcription of the NIMR gene
20 in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(Alekshun, M. A. & Levy, S. B. (1999) *J. Bacteriol.* 181, 4669-4672):569-84; Helene, C. et al. (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

In yet another embodiment, the NIMR nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to
25 improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. et al. (1996) *Bioorganic & Medicinal Chemistry* 4 (George, A. M. & Levy, S. B. (1983) *J. Bacteriol.* 155, 541-548): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics,
30 e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral

- 36 -

backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. et al. (1996) *supra*; Perry-O'Keefe et al. Proc. Natl. Acad. Sci. 93: 14670-675.

5 PNAs of NIMR nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of NIMR nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-
10 directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (e.g., S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. et al. (1996) *supra*; Perry-O'Keefe *supra*).

In another embodiment, PNAs of NIMR molecules can be modified, (e.g., to
15 enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of NIMR nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (e.g.,
20 RNase H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and
25 Finn P.J. et al. (1996) *Nucleic Acids Res.* 24 (Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P. & Kushner, S. R. (1989) *J. Bacteriol.* 171, 4617-4622): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the
30 PNA and the 5' end of DNA (Mag, M. et al. (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule

- 37 -

with a 5' PNA segment and a 3' DNA segment (Finn P.J. et al. (1996) *supra*).

Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. et al. (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide may include other appended groups
5 such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al. (1989) *Proc. Natl. Acad. Sci. US.* 86:6553-6556; Lemaitre et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-
10 triggered cleavage agents (See, e.g., Krol et al. (1988) *Bio-Techniques* 6:958-976) or intercalating agents. (See, e.g., Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

15 *B. NIMR Polypeptides, Fragments Thereof, and Anti-NIMR Antibodies*

One aspect of the invention pertains to isolated NIMR polypeptides, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-NIMR antibodies.

In one embodiment, native NIMR polypeptides can be isolated from cells or
20 tissue sources by an appropriate purification scheme using standard polypeptide purification techniques. In another embodiment, NIMR polypeptides are produced by recombinant DNA techniques. Alternative to recombinant expression, an NIMR polypeptide or polypeptide can be synthesized chemically using standard peptide synthesis techniques. It will be understood that in discussing the uses of NIMR
25 polypeptides, e.g., as shown in Table 1, that fragments of such polypeptides that are not full length NIMR polypeptides as well as full length NIMR polypeptides can be used.

Preferably, the NIMR polypeptides comprise the amino acid sequence encoded by the nucleotide sequence of an NIMR molecule or a portion thereof. In another preferred embodiment, the polypeptide comprises the amino acid sequence of an NIMR
30 molecule listed in Table 1 or a portion thereof.

- 38 -

Preferred NIMR polypeptides are naturally occurring. In other embodiments, the polypeptide has at least about 25, 30, 35, 40, 45, 50, or 60% or more amino acid identity with a naturally occurring NIMR polypeptide. Preferably, the polypeptide has at least about 70% amino acid identity, more preferably 80%, and even more preferably, 90% or 95% amino acid identity with the amino acid sequence of an NIMR molecule shown in Table 1 or a portion thereof. Preferred portions of NIMR polypeptide molecules are biologically active, i.e., encode a portion of the NIMR polypeptide having the ability to modulate microbial responses to environmental stress and, thereby, modulate microbial adaptation to stress and/or microbial virulence.

10 In addition, naturally or non-naturally occurring variants of these polypeptides and nucleic acid molecules which retain the same functional activity, e.g., the ability to modulate drug resistance in a cell are also within the scope of the invention. Such variants can be made, e.g., by mutation using techniques which are known in the art. Alternatively, variants can be chemically synthesized.

15 For example, it will be understood that the NIMR polypeptides described herein also encompass equivalents thereof. For instance, mutant forms of NIMR polypeptides which are functionally equivalent, (e.g., modulate resistance to environmental challenge) can be made using techniques which are well known in the art. Mutations can include, e.g., at least one of a discrete point mutation which can give rise to a substitution, or by 20 at least one deletion or insertion. For example, random mutagenesis can be used. Mutations can be made by random mutagenesis or using cassette mutagenesis. For the former, the entire coding region of a molecule is mutagenized by one of several methods (chemical, PCR, doped oligonucleotide synthesis) and that collection of randomly mutated molecules is subjected to selection or screening procedures. In the latter, 25 discrete regions of a polypeptide, corresponding either to defined structural or functional determinants (e.g., the extracellular, transmembrane, or cytoplasmic domain) are subjected to saturating or semi-random mutagenesis and these mutagenized cassettes are re-introduced into the context of the otherwise wild type allele. In one embodiment, PCR mutagenesis can be used. For example, Megaprimer PCR can be used (O.H. Landt, 30 1990. Gene 96:125-128).

- 39 -

In addition to full length NIMR polypeptides, fragments of NIMR polypeptides and their use are also within the scope of the invention. As used herein, a fragment of an NIMR polypeptide refers to a portion of a full-length NIMR polypeptide which is useful in a screening assay to identify compounds which modulate a biological activity of an NIMR polypeptide (e.g., alter the ability of an NIMR polypeptide to influence drug resistance in a microbe). Accordingly, isolated NIMR polypeptides for use in the instant screening assays can be full length NIMR polypeptides or fragments thereof. Thus, an isolated NIMR polypeptide can comprise, consist essentially of, or consist of an amino acid sequence derived from the full length amino acid sequence of an NIMR polypeptide, provided that it retains an NIMR polypeptide activity.

Portions of the above described polypeptides suitable for use in the claimed assays, such as those which retain their function (e.g., the ability to modulate drug resistance, the ability to modulate drug efflux from a cell, or those which are critical for binding to other molecules (such as DNA, proteins, or compounds) can be easily determined by one of ordinary skill in the art, e.g, using standard truncation or mutagenesis techniques and used in the instant assays. Exemplary techniques are described by Gallegos et al. (1996. J. Bacteriol. 178:6427). In addition, biologically active portions of an NIMR polypeptide include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the NIMR polypeptide, which include fewer amino acids than the full length NIMR polypeptides, and exhibit at least one activity of an NIMR polypeptide are also the subject of the invention.

Other fragments include, for example, truncation polypeptides having a portion of an amino acid sequence of an NIMR molecule shown in Table 1, or of variants thereof, such as a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus. Degradation forms of the polypeptides of the invention in a host cell are also preferred. Further preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions,

- 40 -

flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps
5 can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence. The residues at
10 corresponding positions are then compared and when a position in one sequence is occupied by the same residue as the corresponding position in the other sequence, then the molecules are identical at that position. The percent identity between two sequences, therefore, is a function of the number of identical positions shared by two sequences (*i.e.*, % identity = # of identical positions/total # of positions x 100). The percent identity
15 between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which are introduced for optimal alignment of the two sequences. As used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology".

The comparison of sequences and determination of percent identity between two
20 sequences can be accomplished using a mathematical algorithm. A non-limiting example of a mathematical algorithm utilized for comparison of sequences is the algorithm of Karlin and Altschul (1990) Proc. Natl. Acad. Sci. USA 87:2264, modified as in Karlin and Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) of
25 Altschul, et al. (1990) J. Mol. Biol. 215:403. BLAST nucleotide searches can be performed with the NBLAST program score=100, wordlength=12 to obtain nucleotide sequences homologous to the nucleic acid molecules of the invention. BLAST polypeptide searches can be performed with the XBLAST program, score=50, wordlength =3 to obtain amino acid sequences homologous to the polypeptide molecules
30 of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) Nucleic Acids Research

- 41 -

25(Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P. & Kushner, S. R. (1989)
J. Bacteriol. 171, 4617-4622):3389. When utilizing BLAST and Gapped BLAST
programs, the default parameters of the respective programs (e.g., XBLAST and
NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>. Another preferred, non-
5 limiting algorithm utilized for the comparison of sequences is the algorithm of Myers
and Miller, CABIOS (1988). Such an algorithm is incorporated into the ALIGN
program (version 2.0) which is part of the GCG sequence alignment software package.
When utilizing the ALIGN program for comparing amino acid sequences, a PAM120
weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

10 Another non-limiting example of a mathematical algorithm utilized for the
alignment of polypeptide sequences is the Lipman-Pearson algorithm (Lipman and
Pearson (1985) *Science* 227:1435). When using the Lipman-Pearson algorithm, a
PAM250 weight residue table, a gap length penalty of 12, a gap penalty of 4, and a
Kutple of 2 can be used. A preferred, non-limiting example of a mathematical algorithm
15 utilized for the alignment of nucleic acid sequences is the Wilbur-Lipman algorithm
(Wilbur and Lipman (1983) *Proc. Natl. Acad. Sci. USA* 80:726). When using the
Wilbur-Lipman algorithm, a window of 20, gap penalty of 3, Ktuple of 3 can be used.
Both the Lipman-Pearson algorithm and the Wilbur-Lipman algorithm are incorporated,
for example, into the MEGALIGN program (e.g., version 3.1.7) which is part of the
20 DNASTAR sequence analysis software package.

Additional algorithms for sequence analysis are known in the art, and include
ADVANCE and ADAM., described in Torelli and Robotti (1994) *Comput. Appl. Biosci.*
10:3; and FASTA, described in Pearson and Lipman (1988) *PNAS* 85:2444.

In a preferred embodiment, the percent identity between two amino acid
25 sequences is determined using the GAP program in the GCG software package, using
either a Blosom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6,
or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the
percent identity between two nucleotide sequences is determined using the GAP
program in the GCG software package, using a NWSgapdna. CMP matrix and a gap
30 weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6.

- 42 -

Protein alignments can also be made using the Geneworks global polypeptide alignment program (e.g., version 2.5.1) with the cost to open gap set at 5, the cost to lengthen gap set at 5, the minimum diagonal length set at 4, the maximum diagonal offset set at 130, the consensus cutoff set at 50% and utilizing the Pam 250 matrix.

5 The nucleic acid and polypeptide sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the
10 NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to NIMR nucleic acid molecules of the invention. BLAST polypeptide searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to NIMR polypeptide molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can
15 be utilized as described in Altschul et al., (1997) *Nucleic Acids Res.* 25(Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P. & Kushner, S. R. (1989) *J. Bacteriol.* 171, 4617-4622):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. For example, the nucleotide sequences of the invention can be analyzed using the
20 default Blastn matrix 1-3 with gap penalties set at: existence 11 and extension 1. The amino acid sequences of the invention can be analyzed using the default settings: the Blosum62 matrix with gap penalties set at existence 11 and extension 1. See <http://www.ncbi.nlm.nih.gov>.

The invention also provides NIMR chimeric or fusion polypeptides. As used
25 herein, an NIMR "chimeric polypeptide" or "fusion polypeptide" comprises an NIMR polypeptide operatively linked to a non- NIMR polypeptide. An " NIMR polypeptide" refers to a polypeptide having an amino acid sequence corresponding to NIMR polypeptide, whereas a "non-NIMR polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a polypeptide which is not substantially
30 homologous to the NIMR polypeptide, e.g., a polypeptide which is different from the NIMR polypeptide and which is derived from the same or a different organism. Within

- 43 -

an NIMR fusion polypeptide the NIMR polypeptide can correspond to all or a portion of an NIMR polypeptide. In a preferred embodiment, an NIMR fusion polypeptide comprises at least one biologically active portion of an NIMR polypeptide. Within the fusion polypeptide, the term "operatively linked" is intended to indicate that the NIMR
5 polypeptide and the non-NIMR polypeptide are fused in-frame to each other. The non-NIMR polypeptide can be fused to the N-terminus or C-terminus of the NIMR polypeptide.

For example, in one embodiment, the fusion polypeptide is a GST-NIMR member fusion polypeptide in which the NIMR member sequences are fused to the C-
10 terminus of the GST sequences. In another embodiment, the fusion polypeptide is an NIMR -HA fusion polypeptide in which the NIMR member nucleotide sequence is inserted in a vector such as pCEP4-HA vector (Herrscher, R.F. *et al.* (1995) *Genes Dev.* 9:3067-3082) such that the NIMR member sequences are fused in frame to an influenza hemagglutinin epitope tag. Such fusion polypeptides can facilitate the purification of a
15 recombinant NIMR polypeptide.

Fusion polypeptides and peptides produced by recombinant techniques may be secreted and isolated from a mixture of cells and medium containing the polypeptide or peptide. Alternatively, the polypeptide or peptide may be retained cytoplasmically and the cells harvested, lysed and the polypeptide isolated. A cell culture typically includes
20 host cells, media and other byproducts. Suitable media for cell culture are well known in the art. Polypeptides can be isolated from cell culture media, host cells, or both using techniques known in the art for purifying polypeptides and peptides. Techniques for transfecting host cells and purifying polypeptides and peptides are known in the art.

Preferably, an NIMR fusion polypeptide of the invention is produced by standard
25 recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and
30 enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR

- 44 -

amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide or an HA epitope tag). A NIMR encoding nucleic acid molecule can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the NIMR polypeptide.

In another embodiment, the fusion polypeptide is an NIMR polypeptide containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of NIMR can be increased through use of a heterologous signal sequence. The NIMR fusion polypeptides of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. Use of NIMR fusion polypeptides may be useful therapeutically for the treatment of infection. Moreover, the NIMR-fusion polypeptides of the invention can be used as immunogens to produce anti- NIMR antibodies in a subject.

Preferably, an NIMR chimeric or fusion polypeptide of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A NIMR-

- 45 -

encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the NIMR polypeptide.

The present invention also pertains to variants of the NIMR polypeptides which function as either NIMR agonists (mimetics) or as NIMR antagonists. Variants of the NIMR polypeptides can be generated by mutagenesis, e.g., discrete point mutation or truncation of an NIMR polypeptide. An agonist of the NIMR polypeptides can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of an NIMR polypeptide. An antagonist of an NIMR polypeptide can inhibit one or more of the activities of the naturally occurring form of the NIMR polypeptide by, for example, competitively modulating a cellular activity of an NIMR polypeptide. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the polypeptide has fewer side effects in a subject relative to treatment with the naturally occurring form of the NIMR polypeptide.

In one embodiment, the invention pertains to derivatives of NIMR which may be formed by modifying at least one amino acid residue of NIMR by oxidation, reduction, or other derivatization processes known in the art.

In one embodiment, variants of an NIMR polypeptide which function as either NIMR agonists (mimetics) or as NIMR antagonists can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of an NIMR polypeptide for NIMR polypeptide agonist or antagonist activity. In one embodiment, a variegated library of NIMR variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of NIMR variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential NIMR sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion polypeptides (e.g., for phage display) containing the set of NIMR sequences therein. There are a variety of methods which can be used to produce libraries of potential NIMR variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA

- 46 -

synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential NIMR sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang, S.A.

5 (1983) *Tetrahedron* 39:3; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477).

In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an NIMR coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the
10 double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of
15 various sizes of the NIMR polypeptide.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of
20 NIMR polypeptides. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene
25 whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify NIMR variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave et al. (1993) *Protein Engineering* 6(Cohen, S. P., Hachler, H. & Levy, S. B. (1993) *J Bacteriol.* 175,
30 1484-1492):327-331).

- 47 -

In one embodiment, cell based assays can be exploited to analyze a variegated NIMR library. For example, a library of expression vectors can be transfected into a cell line which ordinarily synthesizes and secretes NIMR. The transfected cells are then cultured such that NIMR and a particular mutant NIMR are secreted and the effect of expression of the mutant on NIMR activity in cell supernatants can be detected, e.g., by any of a number of enzymatic assays. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of NIMR activity, and the individual clones further characterized.

In addition to NIMR polypeptides comprising only naturally-occurring amino acids, NIMR peptidomimetics are also provided. Peptide analogs are commonly used in the pharmaceutical industry as non-peptide drugs with properties analogous to those of the template peptide. These types of non-peptide compound are termed "peptide mimetics" or "peptidomimetics" (Fauchere, J. (1986) *Adv. Drug Res.* 15: 29; Veber and Freidinger (1985) *TINS* p.392; and Evans et al. (1987) *J. Med. Chem* 30: 1229, which are incorporated herein by reference) and are usually developed with the aid of computerized molecular modeling.

Peptide mimetics that are structurally similar to therapeutically useful peptides may be used to produce an equivalent therapeutic or prophylactic effect. Generally, peptidomimetics are structurally similar to a paradigm polypeptide (i.e., a polypeptide that has a biological or pharmacological activity), such as NIMR, but have one or more peptide linkages optionally replaced by a linkage selected from the group consisting of: -CH₂NH-, -CH₂S-, -CH₂-CH₂-, -CH=CH- (cis and trans), -COCH₂-, -CH(OH)CH₂-, and -CH₂SO-, by methods known in the art and further described in the following references: Spatola, A.F. in "Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins," B. Weinstein, eds., Marcel Dekker, New York, p. 267 (1983); Spatola, A. F., Vega Data (March 1983), Vol. 1, Issue 3, "Peptide Backbone Modifications" (general review); Morley, J. S., *Trends Pharm Sci* (1980) pp. 463-468 (general review); Hudson, D. et al., *Int J Pept Prot Res* (1979) 14:177-185 (-CH₂NH-, CH₂CH₂-); Spatola, A. F. et al., *Life Sci* (1986) 38:1243-1249 (-CH₂-S); Hann, M. M., *J Chem Soc Perkin Trans I* (1982) 307-314 (-CH-CH-, cis and trans); Almquist, R. G. et al., *J Med Chem* (1980) 23:1392-1398 (-COCH₂-); Jennings-White, C. et al., *Tetrahedron Lett* (1982) 23:2533 (-

- 48 -

COCH₂-); Szelke, M. et al., European Appln. EP 45665 (1982) CA: 97:39405 (1982)(-CH(OH)CH₂-); Holladay, M. W. et al., Tetrahedron Lett (1983) 24:4401-4404 (-C(OH)CH₂-); and Hruby, V. J., Life Sci (1982) 31:189-199 (-CH₂-S-); each of which is incorporated herein by reference. A particularly preferred non-peptide linkage is -

5 CH₂NH-

Such peptide mimetics may have significant advantages over polypeptide embodiments, including, for example: more economical production, greater chemical stability, enhanced pharmacological properties (half-life, absorption, potency, efficacy, etc.), altered specificity (e.g., a broad-spectrum of biological activities), reduced

10 antigenicity, and others. Labeling of peptidomimetics usually involves covalent attachment of one or more labels, directly or through a spacer (e.g., an amide group), to non-interfering position(s) on the peptidomimetic that are predicted by quantitative structure-activity data and/or molecular modeling. Such non-interfering positions generally are positions that do not form direct contacts with the macromolecules(s) to

15 which the peptidomimetic binds to produce the therapeutic effect. Derivatization (e.g., labelling) of peptidomimetics should not substantially interfere with the desired biological or pharmacological activity of the peptidomimetic.

Systematic substitution of one or more amino acids of an NIMR amino acid sequence with a D-amino acid of the same type (e.g., D-lysine in place of L-lysine)

20 may be used to generate more stable peptides. In addition, constrained peptides comprising an NIMR amino acid sequence or a substantially identical sequence variation may be generated by methods known in the art (Rizo and Gierasch (1992) Ann. Rev. Biochem. 61: 387, incorporated herein by reference); for example, by adding internal cysteine residues capable of forming intramolecular disulfide bridges

25 which cyclize the peptide.

The amino acid sequences of NIMR polypeptides identified herein will enable those of skill in the art to produce polypeptides corresponding to NIMR peptide sequences and sequence variants thereof. Such polypeptides may be produced in prokaryotic or eukaryotic host cells by expression of polynucleotides encoding an

30 NIMR peptide sequence, frequently as part of a larger polypeptide. Alternatively, such peptides may be synthesized by chemical methods. Methods for expression of

- 49 -

heterologous polypeptides in recombinant hosts, chemical synthesis of polypeptides, and in vitro translation are well known in the art and are described further in Maniatis et al., *Molecular Cloning: A Laboratory Manual* (1989), 2nd Ed., Cold Spring Harbor, N.Y.; Berger and Kimmel, *Methods in Enzymology*, Volume 152, Guide to Molecular Cloning Techniques (1987), Academic Press, Inc., San Diego, Calif.; Merrifield, J. (1969) *J. Am. Chem. Soc.* 91: 501; Chaiken I. M. (1981) *CRC Crit. Rev. Biochem.* 11: 255; Kaiser et al. (1989) *Science* 243: 187; Merrifield, B. (1986) *Science* 232: 342; Kent, S. B. H. (1988) *Ann. Rev. Biochem.* 57: 957; and Offord, R. E. (1980) *Semisynthetic Proteins*, Wiley Publishing, which are incorporated herein by reference).

Peptides can be produced, typically by direct chemical synthesis, and used e.g., as agonists or antagonists of an NIMR molecule, e.g., to modulate binding of an NIMR polypeptide and a molecule with which it normally interacts. Peptides can be produced as modified peptides, with nonpeptide moieties attached by covalent linkage to the N-terminus and/or C-terminus. In certain preferred embodiments, either the carboxy-terminus or the amino-terminus, or both, are chemically modified. The most common modifications of the terminal amino and carboxyl groups are acetylation and amidation, respectively. Amino-terminal modifications such as acylation (e.g., acetylation) or alkylation (e.g., methylation) and carboxy-terminal-modifications such as amidation, as well as other terminal modifications, including cyclization, may be incorporated into various embodiments of the invention. Certain amino-terminal and/or carboxy-terminal modifications and/or peptide extensions to the core sequence can provide advantageous physical, chemical, biochemical, and pharmacological properties, such as: enhanced stability, increased potency and/or efficacy, resistance to serum proteases, desirable pharmacokinetic properties, and others. Peptides may be used therapeutically, e.g., to treat infection.

An isolated NIMR polypeptide, or a portion or fragment thereof, can also be used as an immunogen to generate antibodies that bind NIMR using standard techniques for polyclonal and monoclonal antibody preparation. A full-length NIMR polypeptide can be used or, alternatively, the invention provides antigenic peptide fragments of NIMR for use as immunogens. The antigenic peptide of NIMR preferably comprises at least 8 amino acid residues and encompasses an epitope of NIMR such that an antibody raised

- 50 -

against the peptide forms a specific immune complex with NIMR. More preferably, the antigenic peptide comprises at least 10 amino acid residues, even more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

5 Alternatively, an antigenic peptide fragment of an NIMR polypeptide can be used as the immunogen. An antigenic peptide fragment of an NIMR polypeptide typically comprises at least 8 amino acid residues of an amino acid sequence of an NIMR molecule of Table 1 and encompasses an epitope of an NIMR polypeptide such that an antibody raised against the peptide forms an immune complex with an NIMR
10 molecule. Preferred epitopes encompassed by the antigenic peptide are regions of NIMR that are located on the surface of the polypeptide, e.g., hydrophilic regions. In one embodiment, an antibody binds substantially specifically to an NIMR polypeptide. In another embodiment, an antibody binds specifically to an NIMR polypeptide.

 In one embodiment such epitopes can be specific for an NIMR polypeptide from
15 one species (i.e., an antigenic peptide that spans a region of an NIMR polypeptide that is not conserved across species is used as immunogen; such non conserved residues can be determined using an alignment such as that provided herein). A standard hydrophobicity analysis of the polypeptide can be performed to identify hydrophilic regions.

 Accordingly, another aspect of the invention pertains to the use of anti- NIMR
20 antibodies. Polyclonal anti-NIMR antibodies can be prepared as described above by immunizing a suitable subject with an NIMR immunogen. The anti- NIMR antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized an NIMR polypeptide. If desired, the antibody molecules directed against an NIMR polypeptide
25 can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as polypeptide A chromatography to obtain the IgG fraction. At an appropriate time after immunization, e.g., when the anti- NIMR antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique
30 originally described by Kohler and Milstein (1975, *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol* 127:539-46; Brown *et al.* (1980) *J Biol Chem*

- 51 -

255:4980-83; Yeh *et al.* (1976) *PNAS* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques.

5 As an alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-NIMR antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with an NIMR to thereby isolate immunoglobulin library members that bind an NIMR polypeptide. Kits for generating and screening phage display libraries are
10 commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP*TM *Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* International Publication No. WO
15 92/18619; Dower *et al.* International Publication No. WO 91/17271; Winter *et al.* International Publication WO 92/20791; Markland *et al.* International Publication No. WO 92/15679; Breitling *et al.* International Publication WO 93/01288; McCafferty *et al.* International Publication No. WO 92/01047; Garrard *et al.* International Publication No. WO 92/09690; Ladner *et al.* International Publication No. WO 90/02809; Fuchs *et al.*
20 (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum Antibod Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J Mol Biol* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *PNAS* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc Acid Res* 19:4133-4137;
25 Barbas *et al.* (1991) *PNAS* 88:7978-7982; and McCafferty *et al.* *Nature* (1990) 348:552-554.

Additionally, recombinant anti-NIMR antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of
30 the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in

- 52 -

Robinson *et al.* International Patent Publication PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT Application WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *PNAS* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *PNAS* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl Cancer Inst.* 80:1553-1559; Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti- NIMR antibody (*e.g.*, monoclonal antibody) can be used to isolate an NIMR polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. Anti- NIMR antibodies can facilitate the purification of natural NIMR polypeptides from cells and of recombinantly produced NIMR polypeptides expressed in host cells. Moreover, an anti- NIMR antibody can be used to detect an NIMR polypeptide (*e.g.*, in a cellular lysate or cell supernatant). Detection may be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Accordingly, in one embodiment, an anti- NIMR antibody of the invention is labeled with a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials and radioactive materials.

25 **III. Microbes**

Numerous different microbes are suitable for use as sources of NIMR nucleic acid molecules or polypeptides, as host cells, and in testing for compounds in the screening assays described herein, *e.g.*, for testing for compounds that modulate the activity and/or expression of an NIMR polypeptides. The term "microbe" includes microorganisms having an NIMR polypeptide or those that can be engineered to express such a molecule for the purposes of developing a screening assay. Preferably "microbe"

- 53 -

refers to unicellular prokaryotic or eukaryotic microbes including bacteria, fungi, or protozoa. In another embodiment, microbes suitable for use in the invention are multicellular, e.g., parasites or fungi. In preferred embodiments, microbes are pathogenic for humans, animals, or plants. In other embodiments, microbes causing
5 environmental problems, e.g., fouling or spoilage or that perform useful functions such as breakdown of plant matter are also preferred. As such, any of these disclosed microbes may be used as intact cells or as sources of materials for cell-free assays as described herein.

In preferred embodiments, microbes for use in the claimed methods are bacteria,
10 either Gram-negative or Gram-positive bacteria. In a preferred embodiment, any bacteria that are shown to become resistant to drugs, preferably antibiotics, are appropriate for use in the claimed methods.

In preferred embodiments, microbes are bacteria from the family *Enterobacteriaceae*. In more preferred embodiments bacteria of a genus selected from
15 the group consisting of: *Escherichia*, *Proteus*, *Salmonella*, *Klebsiella*, *Shigella*, *Providencia*, *Enterobacter*, *Burkholderia*, *Pseudomonas*, *Acinetobacter*, *Aeromonas*, *Haemophilus*, *Yersinia*, *Neisseria*, and *Erwinia*, *Rhodopseudomonas*, or *Burkholderia*.

In yet other embodiments, the microbes are Gram-positive bacteria and are from a genus selected from the group consisting of: *Lactobacillus*, *Azorhizobium*,
20 *Streptomyces*, *Pediococcus*, *Photobacterium*, *Bacillus*, *Enterococcus*, *Staphylococcus*, *Clostridium*, *Streptococcus*, *Butyrivibrio*, *Sphingomonas*, *Rhodococcus*, or *Streptomyces*

In yet other embodiments, the microbes are acid fast bacilli, e.g., from the genus *Mycobacterium*.

In still other embodiments, the microbes are, e.g., selected from a genus selected
25 from the group consisting of: *Methanobacterium*, *Sulfolobus*, *Archaeoglobus*, *Rhodobacter*, or *Sinorhizobium*.

In other embodiments, the microbes are fungi. In a preferred embodiment the fungus is from the genus *Mucor* or *Candida*, e.g., *Mucor racemosus* or *Candida albicans*.

30 In yet other embodiments, the microbes are protozoa. In a preferred embodiment the microbe is a malaria or cryptosporidium parasite.

- 54 -

IV. Vectors and Host Cells

Preferred NIMR polypeptides for use in screening assays are "isolated" or recombinant polypeptides. In one embodiment, NIMR polypeptides can be made from isolated nucleic acid molecules. Nucleic acid molecules encoding NIMR polypeptides
5 can be used for screening or can be used to produce NIMR polypeptides for use in the instant assays. For example, nucleic acid molecules encoding an NIMR polypeptide can be isolated (e.g., isolated from the sequences which naturally flank it in the chromosome and from cellular components) and can be used to produce an NIMR polypeptide. In one embodiment, a nucleic acid molecule which has been (George, A. M. & Levy, S. B.
10 (1983) *J. Bacteriol.* 155, 541-548) amplified *in vitro* by, for example, polymerase chain reaction (PCR); (Cohen, S. P., Yan, W. & Levy, S. B. (1993) *J Infect. Dis.* 168, 484-488) recombinantly produced by cloning, or (Cohen, S. P., Hachler, H. & Levy, S. B. (1993) *J Bacteriol.* 175, 1484-1492) purified, as by cleavage and gel separation; or
15 synthesized by, for example, chemical synthesis can be used to produce NIMR polypeptides.

NIMR polypeptides can be expressed in a modified form. For example, for secretion of the translated polypeptide into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals
20 may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals. Polypeptides of the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction
25 chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

- 55 -

For recombinant production, host cells can be genetically engineered to incorporate nucleic acid molecules of the invention. In one embodiment nucleic acid molecules specifying NIMR polypeptides can be placed in a vector. The term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid molecule to which it has been linked. The term "expression vector" or "expression system" includes any vector, (e.g., a plasmid, cosmid or phage chromosome) containing a gene construct in a form suitable for expression by a cell (e.g., linked to a promoter). In the present specification, "plasmid" and "vector" are used interchangeably, as a plasmid is a commonly used form of vector. Moreover, the invention is intended to include other vectors which serve equivalent functions. A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids.

Appropriate vectors are widely available commercially and it is within the knowledge and discretion of one of ordinary skill in the art to choose a vector which is appropriate for use with a given host cell. The sequences encoding NIMR polypeptides can be introduced into a cell on a self-replicating vector or may be introduced into the chromosome of a microbe using homologous recombination or by an insertion element such as a transposon.

The expression system constructs may contain control regions that regulate expression. "Transcriptional regulatory sequence" is a generic term to refer to DNA sequences, such as initiation signals, enhancers, operators, and promoters, which induce or control transcription of polypeptide coding sequences with which they are operably linked. It will also be understood that a recombinant gene encoding an NIMR polypeptide can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the

- 56 -

naturally-occurring NIMR gene. Exemplary regulatory sequences are described in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). For instance, any of a wide variety of expression control sequences, that control the expression of a DNA sequence when operatively linked to it, may be used in these vectors to express DNA sequences encoding the NIMR polypeptides of this invention.

Generally, any system or vector suitable to maintain, propagate or express nucleic acid molecules and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al., *Molecular Cloning, A Laboratory Manual*, (supra).

Exemplary expression vectors for expression of a gene encoding an NIMR polypeptide and capable of replication in a bacterium, e.g., a gram positive, gram negative, or in a cell of a simple eukaryotic fungus such as a *Saccharomyces* or, *Pichia*, or in a cell of a eukaryotic organism such as an insect, a bird, a mammal, or a plant, are known in the art. Such vectors may carry functional replication-specifying sequences (replicons) both for a host for expression, for example a *Streptomyces*, and for a host, for example, *E. coli*, for genetic manipulations and vector construction. See e.g. U.S.P.N 4,745,056. Suitable vectors for a variety of organisms are described in Ausubel, F. et al., *Short Protocols in Molecular Biology*, Wiley, New York (1995), and for example, for *Pichia*, can be obtained from Invitrogen (Carlsbad, CA).

Useful expression control sequences, include, for example, the early and late promoters of SV40, adenovirus or cytomegalovirus immediate early promoter, the *lac* system, the *trp* system, the TAC or TRC system, T7 promoter whose expression is directed by T7 RNA polymerase, the major operator and promoter regions of phage lambda, the control regions for fd coat polypeptide, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α -mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations

- 57 -

thereof. A useful translational enhancer sequence is described in U.S. patent number 4,820,639.

It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the type of polypeptide desired to be expressed. Representative examples of appropriate hosts include bacterial cells, such as gram positive, gram negative cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

In preferred embodiments, cells used to express NIMR polypeptides for purification or for use in screening assays, e.g., host cells, comprise a mutation which renders any endogenous NIMR polypeptide nonfunctional or causes the endogenous polypeptide to not be expressed. In other embodiments, mutations may also be made in other related genes of the host cell, such that there will be no interference from the endogenous host loci.

Introduction of a nucleic acid molecule into the host cell ("transformation") can be effected by methods described in many standard laboratory manuals, such as Davis et al., *Basic Methods In Molecular Biology*, (1986) and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). Examples include electroporation, phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Purification of an NIMR polypeptides, e.g., recombinantly expressed polypeptides, can be accomplished using techniques known in the art. For example, if the NIMR polypeptide is expressed in a form that is secreted from cells, the medium can be collected. Alternatively, if the NIMR polypeptide is expressed in a form that is retained by cells, the host cells can be lysed to release the NIMR polypeptide. Such spent medium or cell lysate can be used to concentrate and purify the NIMR polypeptide. For example, the medium or lysate can be passed over a column, e.g., a column to which antibodies specific for the NIMR polypeptide have been bound.

- 58 -

Alternatively, such antibodies can be specific for a non-NIMR polypeptide which has been fused to the NIMR polypeptide (e.g., as a tag) to facilitate purification of the NIMR polypeptide. Other means of purifying NIMR polypeptides are known in the art.

5 *V. Uses Of NIMR Compositions*

The NIMR modulating agents (e.g., nucleic acid molecules, polypeptides, variants, polypeptide homologues, NIMR agonists or antagonists, and antibodies described herein) can be used in one or more of the following methods: a) methods of treatment, e.g., a) treatment of infection and disinfection of surfaces; b) screening assays;
10 c) use in vaccines, d) diagnostic assays, and the like. The isolated nucleic acid molecules of the invention can be used, for example, to express NIMR polypeptide (e.g., in a host cell in gene therapy applications), to detect NIMR mRNA (e.g., in a biological sample) or a genetic alteration in an NIMR gene, and to modulate NIMR activity, as described further below. In addition, the NIMR polypeptides can be used, e.g., to screen
15 for naturally occurring NIMR binding polypeptides, to screen for drugs or compounds which modulate NIMR activity (e.g., are agonists or antagonists of NIMR activity), as well as to treat disorders that would benefit from modulation of NIMR, e.g., infection with a microbe. The NIMR modulating agents can be used to treat infection (e.g., alone or in combination with a second drug, e.g., an antibiotic) or to reduce contamination
20 (e.g., alone or in combination with a non-antibiotic agent). NIMR modulating agents can also be used to alter MarA regulation of NIMR genes. For example, such agents can be used to downregulate genes that are normally upregulated by MarA or to upregulate genes that are normally downregulated by MarA. Moreover, the anti-NIMR antibodies of the invention can be used to modulate NIMR activity and to detect and isolate NIMR
25 polypeptides, regulate the bioavailability of NIMR polypeptides, and modulate NIMR activity.

A. Methods of Treatment

The subject compositions can be used in treating disorders that would benefit
30 from modulation of an NIMR polypeptide activity, e.g., in treating a subject having an infection with a microbe.

- 59 -

As used herein the term "infection" includes the presence of a microbe in or on a subject which, if its growth were inhibited, would result in a benefit to the subject. As such, the term "infection" in addition to referring to the presence of pathogens also includes normal flora which is not desirable, e.g., on the skin of a burn patient or in the
5 gastrointestinal tract of an immunocompromised patient. As used herein, the term "treating" refers to the administration of a compound to a subject, for prophylactic and/or therapeutic purposes. The term "administration" includes delivery to a subject, e.g., by any appropriate method which serves to deliver the drug to the site of the infection. Administration of the drug can be, e.g., oral, intravenous, or topical (as
10 described in further detail below). Drugs can also be contacted with microbes that are not present in the body, but are present in the environment, e.g., on surfaces.

Methods of modulating expression and/or activity of an NIMR polypeptide in a microbial cell are useful in modulation, e.g., of microbial adaptation to environmental stress and/or modulation of microbial virulence. Generally, it is desirable to increase
15 expression and/or activity of those genes that are downmodulated by overexpression of MarA and to decrease the expression and/or activity of those genes that are upmodulated by overexpression of MarA.

Exemplary NIMR downmodulatory agents include: antisense NIMR nucleic acid molecules, anti-NIMR antibodies, dominant negative NIMR mutants, NIMR
20 antagonists, or compounds which downmodulate NIMR activity identified using the subject screening assays. Additionally or alternatively, compounds which downmodulate NIMR activity can be designed using approaches known in the art.

Exemplary NIMR stimulatory agents include active NIMR polypeptide molecules and nucleic acid molecules encoding NIMR that are introduced into a cell to
25 increase NIMR activity in the cell.

The modulatory methods of the invention can be performed in vitro or in vivo.

NIMR modulating agents can be used alone, in combination with other NIMR modulating agents (e.g., that modulate the same or a different NIMR molecule), or with other drugs (e.g., antibiotic or non-antibiotic drugs).

- 60 -

In one embodiment, an NIMR modulating agent can be administered to a subject alone, e.g., prior to administration of an antibiotic agent in order to increase the efficacy of the antibiotic. In one embodiment, an NIMR modulating agent can be administered to a subject in combination with an antibiotic agent in order to increase the efficacy of the
5 antibiotic.

In another embodiment, an NIMR modulating agent or agents can be used to disinfect surfaces, e.g., in combination with a non-antibiotic agent such as a biocide, in order to increase the effectiveness of the non-antibiotic agent.

In one embodiment, a "combination product" can be formulated comprising an
10 NIMR modulating agent and a non-antibiotic agent, e.g., a disinfectant for decontamination of surfaces or a consumer product (e.g., a detergent, soap, deodorant, mouthwash, toothpaste, or lotion).

B. Uses in Identifying NIMR Agonists and Antagonists

15 The invention provides a method (also referred to herein as a "screening assay") to identify those which modulate (enhance (agonists) or block (antagonists)) the action of NIMR polypeptides or nucleic acid molecules, particularly those compounds that are bacteriostatic and/or bactericidal or prevent the infectious process. The subject screening assays can be used to identify modulators, i.e., candidate or test compounds or
20 agents (e.g., peptides, peptidomimetics, small molecules or other drugs) which modulate NIMR polypeptides, i.e., have a stimulatory or inhibitory effect on, for example, NIMR polypeptide expression or NIMR polypeptide activity. Test compounds may be natural substrates and ligands or may be structural or functional mimetics. See, e.g., Coligan et al., Current Protocols in Immunology 1(2): Chapter 5 (1991).

25 NIMR polypeptide agonists and antagonists can be assayed in a variety of ways. For example, in one embodiment, the invention provides for methods for identifying a compound which modulates an NIMR molecule, e.g., by measuring the ability of the compound to interact with an NIMR nucleic acid molecule or an NIMR polypeptide or the ability of a compound to modulate the activity or expression of an NIMR
30 polypeptide. Furthermore, the ability of a compound to modulate the binding of an

- 61 -

NIMR polypeptide or NIMR nucleic acid molecule to a molecule to which they normally bind, e.g., an NIMR binding polypeptide can be tested.

Compounds for testing in the instant methods can be derived from a variety of different sources and can be known or can be novel. Preferably, a screening assay is performed to test the activity of a compound not previously known to have the activity tested for. Each of the NIMR sequences provided herein may be used in the discovery and development of antibacterial compounds. The NIMR polypeptide or portions thereof, upon expression, can be used as a target for the screening of antibacterial drugs. In another embodiment, antisense nucleic acid molecules or nucleic acid molecules that encode for dominant negative NIMR mutants can also be tested in the subject assays.

In one embodiment, libraries of compounds are tested in the instant methods. In another embodiment, known compounds are tested in the instant methods. In another embodiment, compounds among the list of compounds generally regarded as safe (GRAS) by the Environmental Protection Agency are tested in the instant methods.

In one embodiment, a library of compounds can be screened in the subject assays. A recent trend in medicinal chemistry includes the production of mixtures of compounds, referred to as libraries. While the use of libraries of peptides is well established in the art, new techniques have been developed which have allowed the production of mixtures of other compounds, such as benzodiazepines (Bunin et al. 1992. *J. Am. Chem. Soc.* 114:10987; DeWitt et al. 1993. *Proc. Natl. Acad. Sci. USA* 90:6909) peptoids (Zuckermann. 1994. *J. Med. Chem.* 37:2678) oligocarbamates (Cho et al. 1993. *Science*. 261:1303), and hydantoins (DeWitt et al. supra). Rebek et al. have described an approach for the synthesis of molecular libraries of small organic molecules with a diversity of 10^4 - 10^5 (Carell et al. 1994. *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell et al. *Angew. Chem. Int. Ed. Engl.* 1994. 33:2061).

The compounds for screening in the assays of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide

- 62 -

libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. *Anticancer Drug Des.* 1997. 12:145).

Exemplary compounds which can be screened for activity include, but are not
5 limited to, peptides, nucleic acids, carbohydrates, small organic molecules (e.g., polyketides) (Cane et al. 1998. *Science* 282:63), and natural product extract libraries. In one embodiment, the test compound is a peptide or peptidomimetic. In another, preferred embodiment, the compounds are small, organic non-peptidic compounds.

Other exemplary methods for the synthesis of molecular libraries can be found in
10 the art, for example in: Erb et al. 1994. *Proc. Natl. Acad. Sci. USA* 91:11422; Horwell et al. 1996 *Immunopharmacology* 33:68; and in Gallop et al. 1994. *J. Med. Chem.* 37:1233. Libraries of compounds may be presented in solution (e.g., Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP
15 '409), plasmids (Cull et al. (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla et al. (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra*). Other types of peptide libraries may also be expressed, see, for example, U.S. Patents 5,270,181 and 5,292,646). In still another embodiment,
20 combinatorial polypeptides can be produced from a cDNA library.

The efficacy of the agonist or antagonist can be assessed by generating dose response curves from data obtained using various concentrations of the test modulating agent. Moreover, a control assay can also be performed to provide a baseline for comparison. As described in more detail below, either whole cell or cell free assay
25 systems can be employed.

1. Whole Cell Assays

In one embodiment of the invention, the subject screening assays can be performed using whole cells. In one embodiment of the invention, the step of
30 determining whether a compound reduces the activity or expression of an NIMR polypeptide comprises contacting a cell expressing an NIMR polypeptide with a

- 63 -

compound and measuring the ability of the compound to modulate the activity or expression of an NIMR polypeptide.

In another embodiment, modulators of NIMR polypeptide expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of NIMR polypeptide mRNA or protein in the cell is determined. The level of expression of NIMR polypeptide mRNA or protein in the presence of the candidate compound is compared to the level of expression of NIMR polypeptide mRNA or polypeptide in the absence of the candidate compound. The candidate compound can then be identified as a modulator of NIMR polypeptide expression based on this comparison. For example, when expression of NIMR polypeptide mRNA or protein is greater (e.g., statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of NIMR polypeptide mRNA or protein expression. Alternatively, when expression of NIMR polypeptide mRNA or protein is less (e.g., statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of NIMR mRNA or protein expression. The level of NIMR mRNA or protein expression in the cells can be determined by methods described herein for detecting NIMR mRNA or protein.

To measure expression of an NIMR polypeptide, transcription of an NIMR nucleic acid molecule gene can be measured in control cells which have not been treated with the compound and compared with that of test cells which have been treated with the compound. For example, cells which express endogenous NIMR polypeptides or which are engineered to express or overexpress recombinant NIMR polypeptides can be caused to express or overexpress a recombinant NIMR polypeptide in the presence and absence of a test modulating agent of interest, with the assay scoring for modulation in NIMR polypeptide responses by the target cell mediated by the test agent. For example, as with the cell-free assays, modulating agents which produce a change, e.g., a statistically significant change in NIMR polypeptide -dependent responses (either an increase or decrease) can be identified.

- 64 -

Recombinant expression vectors that can be used for expression of NIMR polypeptide are known in the art (see discussions above). In one embodiment, within the expression vector the NIMR polypeptide -coding sequences are operatively linked to regulatory sequences that allow for constitutive or inducible expression of NIMR polypeptide in the indicator cell(s). Use of a recombinant expression vector that allows for constitutive or inducible expression of NIMR polypeptide in a cell is preferred for identification of compounds that enhance or inhibit the activity of NIMR polypeptide. In an alternative embodiment, within the expression vector the NIMR polypeptide coding sequences are operatively linked to regulatory sequences of the endogenous NIMR polypeptide gene (*i.e.*, the promoter regulatory region derived from the endogenous gene). Use of a recombinant expression vector in which NIMR polypeptide expression is controlled by the endogenous regulatory sequences is preferred for identification of compounds that enhance or inhibit the transcriptional expression of NIMR polypeptide.

15 In one embodiment, the level of transcription can be determined by measuring the amount of RNA produced by the cell. For example, the RNA can be isolated from cells which express an NIMR polypeptide and that have been incubated in the presence or absence of compound. Northern blots using probes specific for the sequences to be detected can then be performed using techniques known in the art. Numerous other, art-recognized techniques can be used. For example, western blot analysis can be used to test for NIMR. For example, in another embodiment, transcription of specific RNA molecules can be detected using the polymerase chain reaction, for example by making cDNA copies of the RNA transcript to be measured and amplifying and measuring them. In another embodiment, RNase protection assays, such as S1 nuclease mapping or

25 RNase mapping can be used to detect the level of transcription of a gene. In another embodiment, primer extension can be used.

In yet other embodiments, the ability of a compound to induce a change in transcription or translation of an NIMR polypeptide can be accomplished by measuring the amount of NIMR polypeptide produced by the cell. Polypeptides which can be

30 detected include any polypeptides which are produced upon the activation of an NIMR responsive promoter, including, for example, both endogenous sequences and reporter

- 65 -

gene sequences. In one embodiment, the amount of polypeptide made by a cell can be detected using an antibody against that polypeptide. In other embodiments, the activity of such a polypeptide can be measured.

In one embodiment, other sequences which are regulated by an NIMR promoter (e.g., a promoter having sequence identity with a promoter that regulates expression of an NIMR gene set forth in Table 1) can be detected. In one embodiment, sequences not normally regulated by an NIMR promoter can be assayed by linking them to a promoter that regulates transcription of an NIMR polypeptide.

In preferred embodiments, to provide a convenient readout of the transcription from an NIMR promoter, such a promoter is linked to a reporter gene, the transcription of which is readily detectable. For example, a bacterial cell, e.g., an *E. coli* cell, can be transformed as taught in Cohen et al. 1993. J. Bacteriol. 175:7856.

Examples of reporter genes include, but are not limited to, CAT (chloramphenicol acetyl transferase) (Alton and Vapnek (1979), Nature 282: 864-869) luciferase, and other enzyme detection systems, such as beta-galactosidase; firefly luciferase (deWet et al. (1987), Mol. Cell. Biol. 7:725-737); bacterial luciferase (Engebrecht and Silverman (1984), PNAS 1: 4154-4158; Baldwin et al. (1984), Biochemistry 23: 3663-3667); PhoA, alkaline phosphatase (Toh et al. (1989) Eur. J. Biochem. 182: 231-238, Hall et al. (1983) J. Mol. Appl. Gen. 2: 101), human placental secreted alkaline phosphatase (Cullen and Malim (1992) Methods in Enzymol. 216:362-368) and green fluorescent polypeptide (U.S. patent 5,491,084; WO96/23898).

In yet another embodiment, the ability of a compound to modulate an NIMR polypeptide activity, (e.g., to modulate microbial responses to environmental stress and, thereby, modulate microbial adaptation to stress and/or microbial virulence) can be tested by measuring the ability of the compound to affect the ability of a microbe to adapt to a drug, e.g. by testing the ability of the microbe to grow in the presence of the drug. For example, the ability of a test compound to modulate the minimal inhibitory concentration (MIC) of the indicator compound can be tested. Such an assay can be performed using a standard methods, e.g., an antibiotic disc assay or an automated growth assay, e.g., using a system such as one commercially available from Viteck . In one embodiment, the method comprises detecting the ability of the compound to

- 66 -

modulate growth of a microbe in the presence of one or more non-antibiotic agents. In another embodiment, the method comprises detecting the ability of the compound to modulate growth of a microbe in the presence of one or more antibiotics.

In another embodiment, the ability of a test compound to modulate the efflux of a
5 drug from the cell can be tested. In this method, microbes are contacted with a test compound with or without an indicator compound (an indicator compound is one which is normally exported by the cell). The ability of a test compound to inhibit the activity of an efflux pump is demonstrated by determining whether the intracellular concentration of the test compound or the indicator compound (e.g., a drug or a dye) is
10 elevated in the presence of the test compound. If the intracellular concentration of the indicator compound is increased in the presence of the test compound as compared to the intracellular concentration in the absence of the test compound, then the test compound can be identified as an inhibitor of an efflux pump. Thus, one can determine whether or not the test compound is an inhibitor of an efflux pump by showing that the test
15 compound affects the ability of an efflux pump present in the microbe to export the indicator compound.

The "intracellular concentration" of an indicator compound includes the concentration of the indicator compound inside the outermost membrane of the microbe. The outermost membrane of the microbe can be, e.g., a cytoplasmic membrane. In the
20 case of Gram-negative bacteria, the relevant "intracellular concentration" is the concentration in the cellular space in which the indicator compound localizes, e.g., the cellular space which contains a target of the indicator compound.

In one embodiment, the method comprises detecting the ability of the compound to reduce antibiotic resistance in a microbe. For example, in one embodiment, the
25 indicator compound comprises an antibiotic and the effect of the test compound on the intracellular concentration of antibiotic in the microbe is measured. In one embodiment, an increase in the intracellular concentration of antibiotic can be measured directly, e.g., in an extract of microbial cells. For example, accumulation of a radiolabelled antibiotic can be determined using standard techniques. For instance, microbes can be contacted
30 with a radiolabelled antibiotic as an indicator composition in the presence and absence of a test compound. The concentration of the antibiotic inside the cells can be measured at

- 67 -

equilibrium by harvesting cells from the two groups (with and without test compound) and cell associated radioactivity measured with a liquid scintillation counter. In another embodiment, an increase in the intracellular concentration of antibiotic can be measured indirectly, e.g., by a showing that a given concentration of antibiotic when contacted
5 with the microbe is sufficient to inhibit the growth of the microbe in the presence of the test compound, but not in the absence of the test compound.

In another embodiment, measurement of the intracellular concentration of an indicator compound can be facilitated by using an indicator compound which is readily detectable by spectroscopic means. Such a compound may be, for example, a dye, e.g., a
10 basic dye, or a fluorophore. Exemplary indicator compounds include: acridine, ethidium bromide, gentian violet, malachite green, methylene blue, benzenyl viologen, bromothymol blue, toluidine blue, methylene blue, rose bengal, alcyan blue, ruthenium red, fast green, aniline blue, xylene cyanol, bromophenol blue, coomassie blue, bromocresol purple, bromocresol green, trypan blue, and phenol red.

15 In such an assay, the effect of the test compound on the ability of the cell to export the indicator compound can be measured spectroscopically. For example, the intracellular concentration of the dye or fluorophore can be determined indirectly, by determining the concentration of the indicator compound in the suspension medium or by determining the concentration of the indicator compound in the cells. This can be
20 done, e.g., by extracting the indicator compound from the cells or by visual inspection of the cells themselves.

In another embodiment, the presence of an indicator compound in a microbe can be detected using a reporter gene which is sensitive to the presence of the indicator compound. Exemplary reporter genes are known in the art. For example, a reporter gene
25 can provide a colorimetric read out or an enzymatic read out of the presence of an indicator compound. In yet another embodiment, a reporter gene whose expression is inducible by the presence of a drug in a microbe can be used. For example, a microbe can be grown in the presence of a drug with and without a putative test compound. In cells in which the efflux pump is inhibited, the concentration of the drug will be
30 increased and the reporter gene construct will be expressed. By this method, efflux

- 68 -

pump inhibitors are identified by their ability to inhibit the export rate of the drug and, thus, to induce reporter gene expression.

In another embodiment, a primary screening assay is used in which an indicator compound which does not comprise an antibiotic is employed. In one embodiment, upon the identification of a test compound that increases the intracellular concentration of the test compound, a secondary screening assay is performed in which the effect of the same test compound on susceptibility to the drug of interest, e.g., antibiotic resistance, is measured.

In yet another embodiment, the ability of a compound to modulate the binding of an NIMR polypeptide to an NIMR binding polypeptide can be determined. NIMR binding polypeptides can be identified using techniques which are known in the art. For example, in the case of binding polypeptides that interact with NIMR polypeptides, interaction trap assays or two hybrid screening assays can be used.

NIMR binding polypeptides can be identified e.g., e.g., by using an NIMR polypeptides or portions thereof of the invention as a "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Biotechniques* 14:920-924; Iwabuchi et al. (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with NIMR polypeptides ("NIMR -binding polypeptides") and are involved in NIMR activity. Such NIMR family-binding polypeptides are also likely to be involved in the propagation of signals by the NIMR polypeptides or to associate with NIMR polypeptides and enhance or inhibit their activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for an NIMR polypeptide is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming an NIMR

- 69. -

polypeptide-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be
5 detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the polypeptide which interacts with the NIMR polypeptide.

NIMR binding polypeptides may also be identified in other ways. For example, a library of molecules can be tested for the presence of NIMR binding polypeptides. In
10 one embodiment, the library of molecules can be tested by expressing them in an expression vector, e.g., a bacteriophage. Bacteriophage can be made to display on their surface a plurality of polypeptide sequences, each polypeptide sequence being encoded by a nucleic acid contained within the bacteriophage. The phage expressing these candidate NIMR binding polypeptides can be tested for the ability to bind an
15 immobilized NIMR polypeptide, to obtain those polypeptides having affinity for the NIMR polypeptide. For example, the method can comprise: contacting the immobilized NIMR polypeptide with a sample of the library of bacteriophage so that the NIMR polypeptide can interact with the different polypeptide sequences and bind those having affinity for the NIMR polypeptide to form a set of complexes consisting of immobilized
20 NIMR polypeptide and bound bacteriophage. The complexes which have not formed a complex can be separated. The complexes of NIMR polypeptide and bound bacteriophage can be contacted with an agent that dissociates the bound bacteriophage from the complexes; and the dissociated bacteriophage can be isolated and the sequence of the nucleic acid molecule encoding the displayed polypeptide obtained, so that amino
25 acid sequences of displayed polypeptides with affinity for NIMR polypeptides are obtained.

In the case of NIMR nucleic acid molecules, NIMR binding polypeptides can be identified, e.g., by contacting an NIMR nucleotide sequence with candidate NIMR binding polypeptides (e.g., in the form of microbial extract) under conditions which
30 allow interaction of components of the extract with the NIMR nucleotide sequence. The

- 70 -

ability of the NIMR nucleotide sequence to interact with the components can then be measured to thereby identify a polypeptide that binds to an NIMR nucleotide sequence.

2. Cell-Free Assays

5 The subject screening methods can involve cell-free assays, e.g., using high-throughput techniques. For example, to screen for agonists or antagonists, a synthetic reaction mix comprising an NIMR molecule and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule that may be an agonist or antagonist. In one embodiment, the reaction mix can further comprise a
10 cellular compartment, such as a membrane, cell envelope or cell wall, or a combination thereof. The ability of the test compound to agonize or antagonize the NIMR polypeptide is reflected in decreased binding of the NIMR polypeptide to an NIMR binding polypeptide or in a decrease in NIMR polypeptide activity.

 In many drug screening programs which test libraries of modulating agents and
15 natural extracts, high throughput assays are desirable in order to maximize the number of modulating agents surveyed in a given period of time. Assays which are performed in cell-free systems, such as may be derived with purified or semi-purified proteins, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is
20 mediated by a test modulating agent. Moreover, the effects of cellular toxicity and/or bioavailability of the test modulating agent can be generally ignored in the *in vitro* system.

 In one embodiment, the ability of a compound to modulate the activity of an NIMR polypeptide is accomplished using isolated NIMR polypeptides or NIMR nucleic
25 acid molecule in a cell-free system. In such an assay, the step of measuring the ability of a compound to modulate the activity of the NIMR polypeptide is accomplished, for example, by measuring direct binding of the compound to an NIMR polypeptide or NIMR nucleic acid molecule or the ability of the compound to alter the ability of the NIMR polypeptide to bind to a molecule to which the NIMR polypeptide normally binds
30 (e.g., protein or DNA).

- 71 -

In yet another embodiment, an assay of the present invention is a cell-free assay in which an NIMR polypeptide or portion thereof is contacted with a test compound and the ability of the test compound to bind to the NIMR polypeptide or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of an NIMR polypeptide can be accomplished, for example, by determining the ability of the NIMR polypeptide to bind to an NIMR target molecule by one of the methods described above for determining direct binding. Determining the ability of the NIMR polypeptide to bind to an NIMR target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo et al. (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another embodiment, the cell-free assay involves contacting an NIMR polypeptide or biologically active portion thereof with a known compound which binds the NIMR polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the NIMR polypeptide, wherein determining the ability of the test compound to interact with the NIMR polypeptide comprises determining the ability of the NIMR polypeptide to preferentially bind to or modulate the activity of an NIMR target molecule.

The cell-free assays of the present invention are amenable to use of both soluble and/or membrane-bound forms of proteins (e.g., NIMR polypeptides or NIMR binding polypeptides). In the case of cell-free assays in which a membrane-bound form of a polypeptide is used it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the polypeptide is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)_n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate

- 72 -

(CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

For example, compounds can be tested for their ability to directly bind to an NIMR nucleic acid molecule or an NIMR polypeptide or portion thereof, e.g., by using
5 labeled compounds, e.g., radioactively labeled compounds. For example, an NIMR polypeptide sequence can be expressed by a bacteriophage. In this embodiment, phage which display the NIMR polypeptide would then be contacted with a compound so that the polypeptide can interact with and potentially form a complex with the compound. Phage which have formed complexes with compounds can then be separated from those
10 which have not. The complex of the polypeptide and compound can then be contacted with an agent that dissociates the bacteriophage from the compound. Any compounds that bound to the polypeptide can then be isolated and identified.

In another embodiment, the ability of a compound to bind to an NIMR nucleic acid molecule can be measured. For example, gel shift assays or restriction enzyme
15 protection assays can be used. Gel shift assays separate polypeptide-DNA complexes from free DNA by non-denaturing polyacrylamide gel electrophoresis. In such an experiment, the level of binding of a compound to DNA can be determined and compared to that in the absence of compound. Compounds which change the level of this binding are selected in the screen as modulating the activity of an NIMR
20 polypeptide.

Other methods of assaying the ability of proteins to bind to DNA, e.g., DNA footprinting, and nuclease protection are also well known in the art and can be used to test the ability of a compound to bind to an NIMR nucleotide sequence.

In another embodiment, the invention provides a method for identifying
25 compounds that modulate antibiotic resistance by assaying for test compounds that bind to NIMR nucleic acid molecules and interfere, e.g., with gene transcription.

In another embodiment, an NIMR nucleic acid molecule and an NIMR binding polypeptide that normally binds to that nucleotide sequence are contacted with a test compound to identify compounds that block the interaction of an NIMR nucleic acid
30 molecule and an NIMR binding polypeptide. For example, in one embodiment, the NIMR nucleotide sequence and/or the NIMR binding polypeptide are contacted under

- 73 -

conditions which allow interaction of the compound with at least one of the NIMR nucleic acid molecule and the NIMR binding polypeptide. The ability of the compound to modulate the interaction of the NIMR nucleotide sequence with the NIMR binding polypeptide is indicative of its ability to modulate an NIMR polypeptide activity.

- 5 Determining the ability of the NIMR polypeptide to bind to or interact with an NIMR binding polypeptide can be accomplished, e.g., by direct binding. In a direct binding assay, the NIMR polypeptide could be coupled with a radioisotope or enzymatic label such that binding of the NIMR polypeptide to an NIMR polypeptide target molecule can be determined by detecting the labeled NIMR polypeptide in a complex.
- 10 For example NIMR polypeptides can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, NIMR polypeptide molecules can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an
- 15 appropriate substrate to product.

- Typically, it will be desirable to immobilize either NIMR polypeptide, an NIMR binding polypeptide or a compound to facilitate separation of complexes from uncomplexed forms, as well as to accommodate automation of the assay. Binding of NIMR polypeptide to an upstream or downstream binding polypeptide, in the presence
- 20 and absence of a candidate agent, can be accomplished in any vessel suitable for containing the reactants. Examples include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows the polypeptide to be bound to a matrix. For example, glutathione-S-transferase/ NIMR polypeptide (GST/ NIMR polypeptide) fusion proteins can be
- 25 adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates, e.g. an ^{35}S -labeled, and the test modulating agent, and the mixture incubated under conditions conducive to complex formation, e.g., at physiological conditions for salt and pH, though slightly more stringent conditions may be desired. Following incubation, the
- 30 beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly (e.g. beads placed in scintillant), or in the supernatant after

- 74 -

the complexes are subsequently dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-PAGE, and the level of NIMR polypeptide-binding polypeptide found in the bead fraction quantitated from the gel using standard electrophoretic techniques.

5 Other techniques for immobilizing proteins on matrices are also available for use in the subject assay. For instance, either an NIMR polypeptide or polypeptide to which it binds can be immobilized utilizing conjugation of biotin and streptavidin. For instance, biotinylated NIMR polypeptide molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, 10 Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with NIMR polypeptide but which do not interfere with binding of upstream or downstream elements can be derivatized to the wells of the plate, and NIMR polypeptide trapped in the wells by antibody conjugation. As above, preparations of an NIMR polypeptide - 15 binding polypeptide and a test modulating agent are incubated in the NIMR polypeptide-presenting wells of the plate, and the amount of complex trapped in the well can be quantitated. Exemplary methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the NIMR binding polypeptide, or which are 20 reactive with NIMR polypeptide and compete with the binding polypeptide; as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the binding polypeptide, either intrinsic or extrinsic activity. In the instance of the latter, the enzyme can be chemically conjugated or provided as a fusion protein with the NIMR binding polypeptide. To illustrate, the NIMR polypeptide can be chemically cross- 25 linked or genetically fused with horseradish peroxidase, and the amount of protein trapped in the complex can be assessed with a chromogenic substrate of the enzyme, e.g. 3,3'-diamino-benzadine tetrahydrochloride or 4-chloro-1-naphthol. Likewise, a fusion protein comprising the protein and glutathione-S-transferase can be provided, and complex formation quantitated by detecting the GST activity using 1-chloro-2,4- 30 dinitrobenzene (Habig et al (1974) J Biol Chem 249:7130).

- 75 -

For processes which rely on immunodetection for quantitating one of the proteins trapped in the complex, antibodies against the polypeptide, such as anti- NIMR polypeptide antibodies, can be used. Alternatively, the polypeptide to be detected in the complex can be "epitope tagged" in the form of a fusion protein which includes, in addition to the NIMR polypeptide sequence, a second polypeptide for which antibodies are readily available (e.g. from commercial sources). For instance, the GST fusion proteins described above can also be used for quantification of binding using antibodies against the GST moiety. Other useful epitope tags include myc-epitopes (e.g., see Ellison et al. (1991) J Biol Chem 266:21150-21157) which includes a 10-residue sequence from c-myc, as well as the pFLAG system (International Biotechnologies, Inc.) or the pEZZ-protein A system (Pharmacia, NJ).

It is also within the scope of this invention to determine the ability of a compound to modulate the interaction between NIMR polypeptide and its target molecule, without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of NIMR polypeptide with its target molecule without the labeling of either NIMR polypeptide or the target molecule. McConnell, H. M. et al. (1992) *Science* 257:1906-1912. As used herein, a "microphysiometer" (e.g., Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between compound and receptor.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in methods of reducing drug resistance in microbes, e.g., in vivo or ex vivo. For example, an agent identified as described herein (e.g., an NIMR modulating agent such as an antisense NIMR nucleic acid molecule, an NIMR agonist or antagonist, or an NIMR -specific antibody) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Additionally, such agents can be used in methods of treatment (in vivo or ex vivo) or in methods of reducing resistance to

- 76 -

drugs in the environment. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

C. Vaccines

5 Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal, comprising inoculating the individual with an NIMR modulating agent, or a fragment or variant thereof, adequate to produce an immune response and/or to augment an immune response (e.g., an antibody and/or T cell immune response) to ameliorate or prevent infection with a
10 microbe comprising an NIMR polypeptide. The invention also relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector to direct expression of an NIMR molecule, or a fragment or a variant thereof, for expressing an NIMR molecule, or a fragment or a variant thereof in vivo in order to induce an immunological response, such as, to produce antibody
15 and/or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to ameliorate an ongoing infection or to prevent infection. One way of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise, e.g., DNA, RNA, a modified nucleic acid, or a DNA/RNA hybrid.

20 A further aspect of the invention relates to an immunological composition which, when introduced into an individual, induces an immunological response. Such a composition can comprise, e.g., an isolated NIMR polypeptide or an NIMR nucleic acid molecule. The immunologic composition may be used therapeutically or prophylactically and may be dominated by either a humoral response or a cellular immune response.

25 In one embodiment, an NIMR polypeptide or a fragment thereof may be fused with a second polypeptide, which may not by itself produce antibodies, but is capable of stabilizing the first polypeptide and enhancing immunogenic and protective properties. Thus fused recombinant polypeptide, preferably further comprises an antigenic co-protein, such as lipoprotein D from *Hemophilus influenzae*, Glutathione-S-transferase
30 (GST) or beta-galactosidase, relatively large second proteins which solubilize the polypeptide and facilitate production and purification of an NIMR molecule to which

- 77 -

they are fused. Moreover, the second polypeptide may act as an adjuvant in the sense of providing a generalized stimulation of the immune system. The second polypeptide may be attached to either the amino or carboxy terminus of the NIMR polypeptide.

The use of a nucleic acid molecule of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff et al., Hum Mol Genet 1992, 1:363, Manthorpe et al., Hum. Gene Ther. 1993:4, 419), delivery of DNA complexed with specific polypeptide carriers (Wu et al., J Biol Chem. 1989: 264,16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, PNAS USA, 1986:83,9551), encapsulation of DNA in various forms of liposomes (Kaneda et al., Science 1989:243,375), particle bombardment (Tang et al., Nature 1992, 356:152, Eisenbraun et al., DNA Cell Biol 1993, 12:791) and in vivo infection using cloned retroviral vectors (Seeger et al., PNAS USA 1984:81,5849).

In one embodiment, immunostimulatory DNA sequences, such as those described in Sato, Y. et al. Science 273: 352 (1996) can be used in connection with the instant invention.

In one embodiment, a vaccine formulation comprises an immunogenic recombinant polypeptide of the invention together with a suitable carrier. Preferably, such vaccines are administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems, alum, or other systems known in the art. The dosage will depend on the specific activity of the vaccine and on the status of the patient and can be readily determined by routine experimentation.

- 78 -

VI. Compositions Comprising NIMR Modulating Agents

The compositions of the invention can comprise at least one NIMR modulating agent and one or more pharmaceutically acceptable carriers (additives) and/or diluents. A composition can also include a second antimicrobial agent, e.g., an antimicrobial
5 compound, preferably an antibiotic or a non-antibiotic agent.

As described in detail below, the compositions can be formulated for administration in solid or liquid form, including those adapted for the following:
(George, A. M. & Levy, S. B. (1983) *J. Bacteriol.* 155, 541-548) oral administration, for example, drenches (aqueous or non-aqueous solutions or suspensions), tablets, boluses,
10 powders, granules, pastes; (Cohen, S. P., Yan, W. & Levy, S. B. (1993) *J Infect. Dis.* 168, 484-488) parenteral administration, for example, by subcutaneous, intramuscular or intravenous injection as, for example, a sterile solution or suspension; (Cohen, S. P., Hachler, H. & Levy, S. B. (1993) *J Bacteriol.* 175, 1484-1492) topical application, for example, as a cream, ointment or spray applied to the skin; (Sulavick, M. C., Dazer, M.
15 & Miller, P. F. (1997) *J. Bacteriol.* 179, 1857-1866) intravaginally or intrarectally, for example, as a pessary, cream, foam, or suppository; or (Cohen, S. P., Levy, S. B., Foulds, J. & Rosner, J. L. (1993) *J. Bacteriol* 175, 7856-7862) aerosol, for example, as an aqueous aerosol, liposomal preparation or solid particles containing the compound.

The phrase "pharmaceutically-acceptable carrier" as used herein means a
20 pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, solvent or encapsulating material, involved in carrying or transporting the antimicrobial agents or compounds of the invention from one organ, or portion of the body, to another organ, or portion of the body without affecting its biological effect. Each carrier should be "acceptable" in the sense of being compatible
25 with the other ingredients of the composition and not injurious to the subject. Some examples of materials which can serve as pharmaceutically-acceptable carriers include: (George, A. M. & Levy, S. B. (1983) *J. Bacteriol.* 155, 541-548) sugars, such as lactose, glucose and sucrose; (Cohen, S. P., Yan, W. & Levy, S. B. (1993) *J Infect. Dis.* 168, 484-488) starches, such as corn starch and potato starch; (Cohen, S. P., Hachler, H. &
30 Levy, S. B. (1993) *J Bacteriol.* 175, 1484-1492) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; (Sulavick, M. C.,

- 79 -

- Dazer, M. & Miller, P. F. (1997) *J. Bacteriol.* 179, 1857-1866) powdered tragacanth; (Cohen, S. P., Levy, S. B., Foulds, J. & Rosner, J. L. (1993) *J. Bacteriol* 175, 7856-7862) malt; (Aleksun, M. A. & Levy, S. B. (1999) *J. Bacteriol.* 181, 4669-4672) gelatin; (George, A. M. & Levy, S. B. (1983) *J. Bacteriol.* 155, 531-540) talc;
- 5 (Oethinger, M., Podglajen, I., Kern, W. V. & Levy, S. B. (1998) *Antimicrob. Agents Chemother.* 42, 2089-2094) excipients, such as cocoa butter and suppository waxes; (Asako, H., Nakajima, K., Kobayashi, K., Kobayashi, M. & Aono, R. (1997) *Appl. Environ. Microbiol* 63, 1428-1433) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (White, D. G., Goldman, J. D., Demple,
- 10 B. & Levy, S. B. (1997) *J. Bacteriol.* 179, 6122-6126) glycols, such as propylene glycol; (Ariza, R. R., Cohen, S. P., Bachawat, N., Levy, S. B. & Demple, B. (1994) *J. Bacteriol.* 176, 143-148) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; (McMurry, L. M., Oethinger, M. & Levy, S. B. (1998) *FEMS Microbiol. Lett.* 166, 305-309) esters, such as ethyl oleate and ethyl laurate; (Moken, M. C., McMurry, L.
- 15 M. & Levy, S. B. (1997) *Antimicrob. Agents Chemother.* 41, 2770-2772) agar; (Martin, R. G., Gillette, W. K., Rhee, S. & Rosner, J. L. (1999) *Mol. Microbiol.* 34, 431-441) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (Maneewannakul, K. & Levy, S. B. (1996) *Antimicrob. Agents Chemother.* 40, 1695-1698) alginic acid; (Seoane, A. S. & Levy, S. B. (1995) *J. Bacteriol.* 177, 530-535)
- 20 pyrogen-free water; (Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P. & Kushner, S. R. (1989) *J. Bacteriol.* 171, 4617-4622) isotonic saline; (Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) in *Molecular Cloning. A Laboratory Manual*, eds. Cold Spring Harbor Laboratory Press (Cold Spring Harbor, NY)) Ringer's solution; (Blattner, F. R., Plunkett, G. I. I., Bloch, C. A., Perna, N., Burland, V., Riley, M.,
- 25 Collado-Vides, J., Glasner, J. D., Rode, C. K. M., G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997) *Science* 277, 1453-1462) ethyl alcohol; (Tao, H., Bausch, C., Richmond, C., Blattner, F. R. & Conway, T. (1999) *J. Bacteriol.* 181, 6425-6440) phosphate buffer solutions; and (Aleksun, M. N. & Levy, S. B. (1997) *Antimicrob. Agents Chemother.* 41, 2067-2075)
- 30 other non-toxic compatible substances employed in pharmaceutical compositions.
- Proper fluidity can be maintained, for example, by the use of coating materials, such as

- 80 -

lecithin, by the maintenance of the required particle size in the case of dispersions, and by the use of surfactants.

These compositions may also contain additional agents, such as preservatives, wetting agents, emulsifying agents and dispersing agents. Prevention of the action of microorganisms may be ensured by the inclusion of various antibacterial and antifungal agents, for example, paraben, chlorobutanol, phenol sorbic acid, and the like. It may also be desirable to include isotonic agents, such as sugars, sodium chloride, and the like into the compositions. In addition, prolonged absorption of the injectable pharmaceutical form may be brought about by the inclusion of agents which delay absorption such as aluminum monostearate and gelatin.

In some cases, in order to prolong the effect of a drug, it is desirable to slow the absorption of the drug from subcutaneous or intramuscular injection. This may be accomplished by the use of a liquid suspension of crystalline or amorphous material having poor water solubility. The rate of absorption of the drug then depends upon its rate of dissolution which, in turn, may depend upon crystal size and crystalline form. Alternatively, delayed absorption of a parenterally-administered drug form is accomplished by dissolving or suspending the drug in an oil vehicle.

Compositions of the present invention may be administered to epithelial surfaces of the body orally, parenterally, topically, rectally, nasally, intravaginally, intracisternally. They are of course given by forms suitable for each administration route. For example, they are administered in tablets or capsule form, by injection, inhalation, eye lotion, ointment, etc., administration by injection, infusion or inhalation; topical by lotion or ointment; and rectal or vaginal suppositories.

The phrases "parenteral administration" and "administered parenterally" as used herein means modes of administration other than enteral and topical administration, usually by injection, and includes, without limitation, intravenous, intramuscular, intraarterial, intrathecal, intracapsular, intraorbital, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intraarticular, subcapsular, subarachnoid, intraspinal and intrasternal injection and infusion.

- 81 -

The phrases "systemic administration," "administered systemically," "peripheral administration" and "administered peripherally" as used herein mean the administration of a sucrose octasulfate and/or an antibacterial or a contraceptive agent, drug or other material other than directly into the central nervous system, such that it enters the
5 subject's system and, thus, is subject to metabolism and other like processes, for example, subcutaneous administration.

In some methods, the compositions of the invention can be topically administered to any epithelial surface. An "epithelial surface" according to this invention is defined as an area of tissue that covers external surfaces of a body, or which
10 and lines hollow structures including, but not limited to, cutaneous and mucosal surfaces. Such epithelial surfaces include oral, pharyngeal, esophageal, pulmonary, ocular, aural, nasal, buccal, lingual, vaginal, cervical, genitourinary, alimentary, and anorectal surfaces.

Compositions can be formulated in a variety of conventional forms employed for
15 topical administration. These include, for example, semi-solid and liquid dosage forms, such as liquid solutions or suspensions, suppositories, douches, enemas, gels, creams, emulsions, lotions, slurries, powders, sprays, lipsticks, foams, pastes, toothpastes, ointments, salves, balms, douches, drops, troches, chewing gums, lozenges, mouthwashes, rinses.

20 Conventionally used carriers for topical applications include pectin, gelatin and derivatives thereof, polylactic acid or polyglycolic acid polymers or copolymers thereof, cellulose derivatives such as methyl cellulose, carboxymethyl cellulose, or oxidized cellulose, guar gum, acacia gum, karaya gum, tragacanth gum, bentonite, agar, carbomer, bladderwrack, ceratonia, dextran and derivatives thereof, ghatti gum,
25 hectorite, ispaghula husk, polyvinylpyrrolidone, silica and derivatives thereof, xanthan gum, kaolin, talc, starch and derivatives thereof, parafin, water, vegetable and animal oils, polyethylene, polyethylene oxide, polyethylene glycol, polypropylene glycol, glycerol, ethanol, propanol, propylene glycol (glycols, alcohols), fixed oils, sodium, potassium, aluminum, magnesium or calcium salts (such as chloride, carbonate,
30 bicarbonate, citrate, gluconate, lactate, acetate, gluceptate or tartrate).

- 82 -

Such compositions can be particularly useful, for example, for treatment or prevention of an unwanted infections e.g., of the oral cavity, including cold sores, infections of eye, the skin, or the lower intestinal tract. Standard composition strategies for topical agents can be applied to the antimicrobial compounds, or pharmaceutically acceptable salts thereof in order to enhance the persistence and residence time of the drug, and to improve the prophylactic efficacy achieved.

For topical application to be used in the lower intestinal tract or vaginally, a rectal suppository, a suitable enema, a gel, an ointment, a solution, a suspension or an insert can be used. Topical transdermal patches may also be used. Transdermal patches have the added advantage of providing controlled delivery of the compositions of the invention to the body. Such dosage forms can be made by dissolving or dispersing the agent in the proper medium.

Compositions of the invention can be administered in the form of suppositories for rectal or vaginal administration. These can be prepared by mixing the agent with a suitable non-irritating carrier which is solid at room temperature but liquid at rectal temperature and therefore will melt in the rectum or vagina to release the drug. Such materials include cocoa butter, beeswax, polyethylene glycols, a suppository wax or a salicylate, and which is solid at room temperature, but liquid at body temperature and, therefore, will melt in the rectum or vaginal cavity and release the active agent.

Compositions which are suitable for vaginal administration also include pessaries, tampons, creams, gels, pastes, foams, films, or spray compositions containing such carriers as are known in the art to be appropriate. The carrier employed in the sucrose octasulfate /contraceptive agent should be compatible with vaginal administration and/or coating of contraceptive devices. Combinations can be in solid, semi-solid and liquid dosage forms, such as diaphragm, jelly, douches, foams, films, ointments, creams, balms, gels, salves, pastes, slurries, vaginal suppositories, sexual lubricants, and coatings for devices, such as condoms, contraceptive sponges, cervical caps and diaphragms.

For ophthalmic applications, the pharmaceutical compositions can be formulated as micronized suspensions in isotonic, pH adjusted sterile saline, or, preferably, as solutions in isotonic, pH adjusted sterile saline, either with or without a preservative

- 83 -

such as benzylalkonium chloride. Alternatively, for ophthalmic uses, the compositions can be formulated in an ointment such as petrolium. Exemplary ophthalmic compositions include eye ointments, powders, solutions and the like.

Powders and sprays can contain, in addition to sucrose octasulfate and/or
5 antibiotic or contraceptive agent(s), carriers such as lactose, talc, silicic acid, aluminum hydroxide, calcium silicates and polyamide powder, or mixtures of these substances. Sprays can additionally contain customary propellants, such as chlorofluorohydrocarbons and volatile unsubstituted hydrocarbons, such as butane and propane.

10 Ordinarily, an aqueous aerosol is made by formulating an aqueous solution or suspension of the agent together with conventional pharmaceutically acceptable carriers and stabilizers. The carriers and stabilizers vary with the requirements of the particular compound, but typically include nonionic surfactants (Tweens, Pluronic, or polyethylene glycol), innocuous proteins like serum albumin, sorbitan esters, oleic acid,
15 lecithin, amino acids such as glycine, buffers, salts, sugars or sugar alcohols. Aerosols generally are prepared from isotonic solutions.

Compositions of the invention can also be orally administered in any orally-acceptable dosage form including, but not limited to, capsules, cachets, pills, tablets, lozenges (using a flavored basis, usually sucrose and acacia or tragacanth), powders,
20 granules, or as a solution or a suspension in an aqueous or non-aqueous liquid, or as an oil-in-water or water-in-oil liquid emulsion, or as an elixir or syrup, or as pastilles (using an inert base, such as gelatin and glycerin, or sucrose and acacia) and/or as mouth washes and the like, each containing a predetermined amount of sucrose octasulfate and/or antibiotic or contraceptive agent(s) as an active ingredient. A compound may also
25 be administered as a bolus, electuary or paste. In the case of tablets for oral use, carriers which are commonly used include lactose and corn starch. Lubricating agents, such as magnesium stearate, are also typically added. For oral administration in a capsule form, useful diluents include lactose and dried corn starch. When aqueous suspensions are required for oral use, the active ingredient is combined with emulsifying and suspending
30 agents. If desired, certain sweetening, flavoring or coloring agents may also be added.

- 84 -

Tablets, and other solid dosage forms, such as dragees, capsules, pills and granules, may be scored or prepared with coatings and shells, such as enteric coatings and other coatings well known in the pharmaceutical-formulating art. They may also be formulated so as to provide slow or controlled release of the active ingredient therein

5 using, for example, hydroxypropylmethyl cellulose in varying proportions to provide the desired release profile, other polymer matrices, liposomes and/or microspheres. They may be sterilized by, for example, filtration through a bacteria-retaining filter, or by incorporating sterilizing agents in the form of sterile solid compositions which can be dissolved in sterile water, or some other sterile injectable medium immediately before

10 use. These compositions may also optionally contain opacifying agents and may be of a composition that they release the active ingredient(s) only, or preferentially, in a certain portion of the gastrointestinal tract, optionally, in a delayed manner. Examples of embedding compositions which can be used include polymeric substances and waxes. The active ingredient can also be in micro-encapsulated form, if appropriate, with one or

15 more of the above-described excipients.

Liquid dosage forms for oral administration include pharmaceutically acceptable emulsions, microemulsions, solutions, suspensions, syrups and elixirs. In addition to the active ingredient, the liquid dosage forms may contain inert diluents commonly used in the art, such as, for example, water or other solvents, solubilizing agents and emulsifiers,

20 such as ethyl alcohol, isopropyl alcohol, ethyl carbonate, ethyl acetate, benzyl alcohol, benzyl benzoate, propylene glycol, 1,3-butylene glycol, oils (in particular, cottonseed, groundnut, corn, germ, olive, castor and sesame oils), glycerol, tetrahydrofuryl alcohol, polyethylene glycols and fatty acid esters of sorbitan, and mixtures thereof.

Besides inert diluents, the oral compositions can also include adjuvants such as

25 wetting agents, emulsifying and suspending agents, sweetening, flavoring, coloring, perfuming and preservative agents.

Suspensions, in addition to the antimicrobial agent(s) may contain suspending agents as, for example, ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters, microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-

30 agar and tragacanth, and mixtures thereof.

- 85 -

Sterile injectable forms of the compositions of this invention can be aqueous or oleaginous suspensions. These suspensions may be formulated according to techniques known in the art using suitable dispersing or wetting agents and suspending agents.

Wetting agents, emulsifiers and lubricants, such as sodium lauryl sulfate and magnesium stearate, as well as coloring agents, release agents, coating agents, sweetening, flavoring and perfuming agents, preservatives and antioxidants can also be present in the compositions.

The sterile injectable preparation may also be a sterile injectable solution or suspension in a nontoxic parenterally-acceptable diluent or solvent, for example as a solution in 1,3-butanediol. Among the acceptable vehicles and solvents that may be employed are water, Ringer's solution and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose, any bland fixed oil may be employed including synthetic mono-or di-glycerides. Fatty acids, such as oleic acid and its glyceride derivatives are useful in the preparation of injectables, as are natural pharmaceutically-acceptable oils, such as olive oil or castor oil, especially in their polyoxyethylated versions. These oil solutions or suspensions may also contain a long-chain alcohol diluent or dispersant.

In the case of modulators of the activity and/or expression of NIMR molecules which are nucleic acid molecules, the optimal course of administration of the oligomers may vary depending upon the desired result or on the subject to be treated. As used in this context "administration" refers to contacting cells with oligomers, e.g., in vivo or ex vivo. The dosage of nucleic molecule may be adjusted to optimally regulate expression of a protein translated from a target mRNA, e.g., as measured by a readout of RNA stability or by a therapeutic response, without undue experimentation. For example, expression of the protein encoded by the nucleic acid can be measured to determine whether or dosage regimen needs to be adjusted accordingly. In addition, an increase or decrease in RNA and/or protein levels in a cell or produced by a cell can be measured using any art recognized technique. By determining whether transcription has been decreased, the effectiveness of the molecule can be determined.

- 86 -

As used herein, "pharmaceutically acceptable carrier" includes appropriate solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any
5 conventional media or agent is incompatible with the active ingredient, it can be used in the therapeutic compositions. Supplementary active ingredients can also be incorporated into the compositions.

Compositions may be incorporated into liposomes or liposomes modified with polyethylene glycol or admixed with cationic lipids for parenteral administration.
10 Incorporation of additional substances into the liposome, for example, antibodies reactive against membrane proteins found on specific target microbes, can help target the molecule to specific cell types.

Moreover, the present invention provides for administering the subject compositions with an osmotic pump providing continuous infusion of the compositions,
15 for example, as described in Rataiczak et al. (1992 *Proc. Natl. Acad. Sci. USA* 89:11823-11827). Such osmotic pumps are commercially available, e.g., from Alzet Inc. (Palo Alto, Calif.). Topical administration and parenteral administration in a cationic lipid carrier are preferred.

With respect to *in vivo* applications, the formulations of the present invention can
20 be administered to a patient in a variety of forms adapted to the chosen route of administration, namely, parenterally, orally, or intraperitoneally. Parenteral administration, which is preferred, includes administration by the following routes: intravenous; intramuscular; interstitially; intraarterially; subcutaneous; intra ocular; intrasynovial; trans epithelial, including transdermal; pulmonary via inhalation;
25 ophthalmic; sublingual and buccal; topically, including ophthalmic; dermal; ocular; rectal; and nasal inhalation via insufflation. Intravenous administration is preferred among the routes of parenteral administration.

Pharmaceutical preparations for parenteral administration include aqueous solutions of the active compounds in water-soluble or water-dispersible form. In
30 addition, suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for

- 87 -

example, sesame oil, or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol, and/or dextran, optionally, the suspension may also contain stabilizers.

5 Drug delivery vehicles can be chosen e.g., for *in vitro*, for systemic, or for topical administration. These vehicles can be designed to serve as a slow release reservoir or to deliver their contents directly to the target cell. An advantage of using some direct delivery drug vehicles is that multiple molecules are delivered per uptake. Such vehicles have been shown to increase the circulation half-life of drugs that would otherwise be
10 rapidly cleared from the blood stream. Some examples of such specialized drug delivery vehicles which fall into this category are liposomes, hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres.

The subject compositions may be incorporated into liposomes or liposomes modified with polyethylene glycol or admixed with cationic lipids for parenteral
15 administration. Incorporation of additional substances into the liposome, for example, antibodies reactive against membrane proteins found on specific target microbes, can help target the compositions to specific cell types.

Moreover, the present invention provides for administering the subject compositions with an osmotic pump providing continuous infusion of nucleic acid
20 molecules, for example, as described in Rataiczak et al. (1992 *Proc. Natl. Acad. Sci. USA* 89:11823-11827). Such osmotic pumps are commercially available, e.g., from Alzet Inc. (Palo Alto, Calif.). Topical administration and parenteral administration in a cationic lipid carrier are preferred.

With respect to *in vivo* applications, the formulations of the present invention can
25 be administered to a patient in a variety of forms adapted to the chosen route of administration, namely, parenterally, orally, or intraperitoneally. Parenteral administration, which is preferred, includes administration by the following routes: intravenous; intramuscular; interstitially; intraarterially; subcutaneous; intra ocular; intrasynovial; trans epithelial, including transdermal; pulmonary via inhalation;
30 ophthalmic; sublingual and buccal; topically, including ophthalmic; dermal; ocular;

- 88 -

rectal; and nasal inhalation via insufflation. Intravenous administration is preferred among the routes of parenteral administration.

Pharmaceutical preparations for parenteral administration include aqueous solutions of the active compounds in water-soluble or water-dispersible form. In addition, suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil, or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol, and/or dextran, optionally, the suspension may also contain stabilizers.

Drug delivery vehicles can be chosen e.g., for *in vitro*, for systemic, or for topical administration. These vehicles can be designed to serve as a slow release reservoir or to deliver their contents directly to the target cell. An advantage of using some direct delivery drug vehicles is that multiple molecules are delivered per uptake. Such vehicles have been shown to increase the circulation half-life of drugs that would otherwise be rapidly cleared from the blood stream. Some examples of such specialized drug delivery vehicles which fall into this category are liposomes, hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres.

The described compositions may be administered systemically to a subject. Systemic absorption refers to the entry of drugs into the blood stream followed by distribution throughout the entire body. Administration routes which lead to systemic absorption include: intravenous, subcutaneous, intraperitoneal, and intranasal. Each of these administration routes delivers the compositions to accessible diseased cells. Following subcutaneous administration, the therapeutic agent drains into local lymph nodes and proceeds through the lymphatic network into the circulation. The rate of entry into the circulation has been shown to be a function of molecular weight or size. The use of a liposome or other drug carrier localizes the compositions at the lymph node. The nucleic acid molecule can be modified to diffuse into the cell, or the liposome can directly participate in the delivery of the composition into the cell.

- 89 -

For prophylactic applications, the pharmaceutical composition of the invention can be applied prior to physical contact with a microbe. The timing of application prior to physical contact can be optimized to maximize the prophylactic effectiveness of the compound. The timing of application will vary depending on the mode of

5 administration, the epithelial surface to which it is applied, the surface area, doses, the stability and effectiveness of composition under the pH of the epithelial surface, the frequency of application, e.g., single application or multiple applications. Preferably, the timing of application can be determined such that a single application of composition is sufficient. One skilled in the art will be able to determine the most appropriate time

10 interval required to maximize prophylactic effectiveness of the compound.

One of ordinary skill in the art can determine and prescribe the effective amount of the pharmaceutical composition required. For example, one could start doses at levels lower than that required in order to achieve the desired therapeutic effect and gradually increase the dosage until the desired effect is achieved. In general, a suitable daily dose

15 of a composition of the invention will be that amount of the composition which is the lowest dose effective to produce a therapeutic effect. Such an effective dose will generally depend upon the factors described above. It is preferred that administration be intravenous, intracoronary, intramuscular, intraperitoneal, or subcutaneous.

Another aspect of the invention pertains to kits for carrying out the screening

20 assays or modulatory methods of the invention. For example, a kit for carrying out a screening assay of the invention can include a cell comprising an NIMR polypeptide, means for determining NIMR polypeptide activity and instructions for using the kit to identify modulators of NIMR activity.

In another embodiment, the invention provides a kit for carrying out a

25 modulatory method of the invention. The kit can include, for example, a modulatory agent of the invention (e.g., an NIMR inhibitory or stimulatory agent) in a suitable carrier and packaged in a suitable container with instructions for use of the modulatory agent to modulate NIMR expression or activity.

The practice of the present invention will employ, unless otherwise indicated,

30 conventional techniques of cell biology, cell culture, molecular biology, genetics, microbiology, recombinant DNA, and immunology, which are within the skill of the art.

- 90 -

Such techniques are explained fully in the literature. See, for example, *Genetics*; *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, J. *et al.* (Cold Spring Harbor Laboratory Press (1989)); *Short Protocols in Molecular Biology*, 3rd Ed., ed. by Ausubel, F. *et al.* (Wiley, NY (1995)); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed. (1984)); Mullis *et al.* U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. (1984)); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y); *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London (1987)); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds. (1986)); and Miller, J. *Experiments in Molecular Genetics* (Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1972)).

The contents of all references, pending patent applications and published patents, cited throughout this application are hereby expressly incorporated by reference. Each reference disclosed herein is incorporated by reference herein in its entirety. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety.

The invention is further illustrated by the following examples, which should not be construed as further limiting.

20

EXAMPLES

Example 1.

The following materials and methods were used in the examples:

25

Bacterial strains, plasmids and growth conditions. *E. coli* K-12 strain AG100 (George, A. M. & Levy, S. B. (1983) *J. Bacteriol.* 155, 541-548) was used for the PCR amplification of specific DNA probes. *E. coli* AG100Kan, an isogenic strain of AG100 containing a 1.2 kb kanamycin resistance cassette in the place of the *mar* locus (Maneewannakul, K. & Levy, S. B. (1996) *Antimicrob. Agents Chemother.* 40, 1695-1698) was used in all the experiments described. pAS10 (Seoane, A. S. & Levy, S. B. (1995) *J Bacteriol.* 177, 530-535), derived from the temperature-sensitive pMAK705

30

- 91 -

(Chl^R) (Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P. & Kushner, S. R. (1989) *J. Bacteriol.* 171, 4617-4622), carries a 2.5 kb PCR amplified fragment containing the *marCORAB* sequence bearing the *marR5* mutation, which produces no MarR and thus mediates constitutive MarA expression.

5 Bacterial strains were grown in Luria Bertani (LB) media (composition per litre: 10 g tryptone, 10 g NaCl, 5 g yeast extract) at 30°C with vigorous aeration. *E. coli* AG100Kan cells were made competent by the standard CaCl₂ method (Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) in *Molecular Cloning. A Laboratory Manual*, eds. Cold Spring Harbor Laboratory Press (Cold Spring Harbor, NY)) and transformants
10 containing the plasmids pMAK705 or pAS10 were maintained in the presence of 25 µg ml⁻¹ chloramphenicol (Sigma, St. Louis, MO).

RNA extraction. Total RNA was isolated by a modification of the hot-acidic phenol extraction method (Sigma-Genosys Biotechnologies, Inc., The Woodlands, TX).

15 Overnight cultures were diluted 250-fold in fresh LB medium, and grown to mid-logarithmic phase ($A_{530} = 0.35-0.40$). Bacterial pellets from 5 ml cell cultures were harvested at 4°C, and resuspended in 250 µl ice-cold resuspension buffer (0.3 M sucrose-10 mM sodium acetate, pH 4.2) and 37.5 µl of ice-cold 0.5 M EDTA. After incubation on ice for 5 min, cells were lysed by adding 375 µl lysis buffer (2%, sodium
20 dodecyl sulphate, 10 mM sodium acetate, pH 4.2) and heating at 65°C for 3 min. The suspension was extracted three times with 700 µl of pre-warmed acidic phenol (65°C) (Sigma) and the aqueous phase was extracted, first with 700 µl of a mixture of acidic phenol:chlorophorm:isoamyl alcohol (25:24:1), and then with an equal volume of chlorophorm:isoamylalcohol (24:1). The RNA in the aqueous phase was ethanol
25 precipitated at -80°C, and the RNA pellet rinsed with 70% ethanol and resuspended in 100 µl of RNase-free water (Ambion Inc., Austin, TX). Samples were treated with DNaseI (amplification grade, Life Technologies Inc., Gaithersburg, MD), following the manufacturer's instructions, to eliminate DNA contamination. The absence of genomic DNA in the RNA was confirmed by examining samples of the RNA in non-denaturing
30 1.2%, agarose gels, and by performing PCR on DNase treated RNA samples using

- 92 -

primers known to target the genomic DNA. The RNA concentration was determined spectrophotometrically (Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) in *Molecular Cloning. A Laboratory Manual*, eds. Cold Spring Harbor Laboratory Press (Cold Spring Harbor, NY)).

5

Preparation of labeled cDNA and hybridization to the arrays. Labeled cDNA was prepared using the *E. coli* cDNA labeling primers (Sigma-Genosys) following the manufacturer's instructions. The primers were annealed to 1 µg of total RNA in the presence of 333 µM dATP, dCTP and dTTP and 1x reverse transcriptase buffer at 90°C for 2 min. The mixture was cooled to 42°C and 50 U AMV reverse transcriptase (Boehringer-Mannheim, Indianapolis, IN) and 20 µCi ³²P-α-dTP (2,000 Ci/mmol) (New England Nuclear, Boston, MA) were added. Incubation was at 42°C for 2h 30 min. The unincorporated nucleotides were removed from the labeled cDNA using a NucTrap probe purification column (Stratagene, La Jolla, CA) prior to hybridization.

15

Hybridization of the purified labeled cDNA to the Panorama *E. coli* Gene arrays (Sigma-Genosys) was performed in roller bottles following the manufacturer's instructions. Essentially, arrays were pre-wet in 2x SSPE and then pre-hybridized for ~ 2 h at 65°C in 5 ml pre-warmed hybridization solution (5x SSPE, 2% SDS, 1x Denhardt's reagent and 100 µg ml⁻¹ denatured salmon sperm DNA). Denatured labeled cDNA in 5 ml hybridization solution replaced the prehybridization solution and hybridization proceeded for ~ 18 h at 65°C. The arrays were washed 3x with 50 ml wash buffer (0.5x SSPE-0.2% SDS) at room temperature for 3 min intervals and 3x with 100 ml pre-warmed (65°C) wash buffer for 20 min intervals. Hybridizing signals on the membrane were visualized by exposure to Kodak BioMax MR X-ray film and to a Kodak storage phosphorimager screen SO230 (Molecular Dynamics, Sunnivale, CA). Phosphor screens were scanned, after 1 to 3 days exposure, at 50 micron pixel resolution in a Storm 860 phosphorimaging instrument (Molecular Dynamics). Arrays were stripped by immersing the membranes in a boiling solution of 0.5% SDS (w/v) and removal of the probe was confirmed before reuse as described above.

30

Description and quantification of the arrays. The Panorama *E. coli* Gene Arrays (Sigma-Genosys) contain 4,290 PCR-amplified Orfs of the *E. coli* K-12 (MG1655) genome (Blattner, F. R., Plunkett, G. I. I., Bloch, C. A., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. M., G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997) *Science* 277, 1453-1462), spotted in duplicate (see Tao et al. (Tao, H., Bausch, C., Richmond, C., Blattner, F. R. & Conway, T. (1999) *J. Bacteriol.* 181, 6425-6440) for a more detailed description of the arrays).

10 Quantification of the hybridizing signals in the phosphoimager file was carried out by Sigma-Genosys using the Array Vision® software (Imaging Research, Inc.). The relative pixel values for the duplicate spots of each gene were averaged and normalized by expressing the averaged spot signal as a percentage to the signal from the averaged pixel values of the genomic DNA spots in the respective field where each gene
15 was printed (Fig. 1). In figure 1, The ratio between these values in samples from cells expressing or lacking MarA represented the fold change in gene expression. Background values were determined for each field in each array by averaging the pixel values of the empty spaces located in the same secondary grid as the genomic DNA (Fig. 1). Genes whose averaged pixel values were close to background (less than a 2-fold difference
20 from background values) in both experimental and control samples were not considered here. Identical arrays were probed with labeled ³²P-cDNA populations prepared from total RNA from *mar*-deleted, AG100Kan[pMAK705] (panel A) and *mar*-expressing, AG100Kan[pAS10] (panel B) strains. Columns (1-24) and rows (A-P) forming the primary grid in Field 1 of the autoradiogram are labeled. Fields 2 and 3 are similar in
25 format to Field I and are not shown. The four spots in the four corners of each field are genomic DNA. Boxes underneath correspond to expanded views of representative areas shown in (A) and (B) where changes in expression levels are visible for several genes (7 of the differentially expressed genes are labeled as examples).

- 94 -

All the genes identified by computing analysis as members of the *mar* regulon were confirmed by visual analysis of autoradiograms of the arrays in three independent experiments. Only those genes which satisfied both criteria were classified as members of the *mar* regulon.

5

Northern blot analysis. Duplicate samples of DNaseI treated total RNA (5-10 µg) were fractionated electrophoretically on 1-1.2%, denaturing formaldehyde-agarose gels, and RNA was transferred to nylon membranes (Hybond-N, Amersham Life Science Inc., Arlington Heights, IL) using established capillary blotting methods in 10x SSC (Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) in *Molecular Cloning. A Laboratory Manual*, eds. Cold Spring Harbor Laboratory Press (Cold Spring Harbor, NY).). DNA probes for specific *E. coli* genes were amplified by PCR from *E. coli* AG100 chromosomal DNA using the appropriate *E. coli* ORFmer PCR primer pairs (Sigma-Genosys), according to the supplier specifications. After amplification, the PCR products were purified from agarose gels using the Qiaex II gel extraction kit (Qiagen Inc., Valencia, CA) and quantified by comparison to DNA size standards (Life Technologies) of known concentration. Labeling of DNA probes with [³²P]-dCTP (New England Nuclear) using the RTS RadPrime DNA labeling system (Life Technologies) was carried out according to the manufacturer's instructions. Hybridizations were performed using standard procedures at 65°C (Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) in *Molecular Cloning. A Laboratory Manual*, eds. Cold Spring Harbor Laboratory Press (Cold Spring Harbor, NY).), and RNA membranes were washed at high stringency for 15 min intervals, four times in 2x SSC buffer/0.1% SDS and 2 to 4 times in 0.1x SSC buffer/0.1% SDS. Hybridizing bands were visualized as described for the *E. coli* gene macroarrays.

DNA manipulations. Genomic and plasmid DNA were purified from *E. coli* strains using the QIAamp Tissue kit and the QIAprep spin Miniprep kit (Qiagen) respectively, following manufacturer's instructions.

- 95 -

Example 1. **Identification of genes regulated by MarA.** DNA macroarrays, constructed for *E. coli*, which contain most of the genomic Orfs (Blattner, F. R., Plunkett, G. I. I., Bloch, C. A., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. M., G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A.,
 5 Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997) *Science* 277, 1453-1462), allowed studies of expression of the complete genome in the presence or absence of MarA. *E. coli* AG100K a strain (Maneewannakul, K. & Levy, S. B. (1996) *Antimicrob. Agents Chemother.* 40, 1695-1698) bearing only plasmid pMAK705 represented the control, i.e. deficient in *mar* expression. The experimental strain AG100Kan[pAS10]
 10 containing the pMAK705-derived plasmid pAS10, which expresses MarA constitutively (Seoane, A. S. & Levy, S. B. (1995) *J Bacteriol.* 177, 530-535). Antibiotic susceptibility assayed using the E-test method showed the expected increase (~ 4-20 fold) in resistance in the *mar* expressing strain as compared to the control to the antibiotics tested, including norfloxacin, nalidixic acid, tetracycline and ampicillin (data not shown).
 15 ³³P-labeled cDNAs prepared from RNA extracted from *mar*-deleted and *mar*-expressing strains were hybridized to paired macroarrays and phosphorimager files and autoradiograms were obtained (Fig. 1). Previously ~ 15 genes were known to be regulated by MarA (Aleksun, M. N. & Levy, S. B. (1997) *Antimicrob. Agents Chemother.* 41, 2067-2075). The gene macroarrays identified a total of 62 genes
 20 responsive to *mar*-regulation in logarithmic phase: 47 induced and 15 repressed (Table 3). Only those findings detected in all three experiments were included in the list.

The signals for the three genes encoded by the *marRAB* operon were easily detected in the cDNA from the *mar*-expressing but not from the *mar*-deleted strain (Fig. 1). This finding was reassuring given that cDNAs from genes belonging to the same
 25 family of homologues (e.g. *soxS* and *rob* for *marA*) could have caused some level of non-specific binding (Richmond, C. S., Glasner, J. D., Mau, R., Jin, H. & Blattner, F. R. (1999) *Nucleic Acids Res.* 27, 3821-3835). For *marR*, *marA* and *marB*, the expression was 31-fold, 35-fold and 12-fold higher (averaged values) than in control samples (Table 3). Although the signal for *marB* expression was consistently less than the signals for
 30 *marR* and *marA* expression the meaning is unclear. Since the spotted PCR products differ in length (which has an effect in hybridizing intensities, (Richmond, C. S.,

- 96 -

Glasner, J. D., Mau, R., Jin, H. & Blattner, F. R. (1999) *Nucleic Acids Res.* 27, 3821-3835)), and because the efficiency of reverse transcription will vary between different RNAs, the results do not allow comparative analysis between different genes. The expression of the divergent *marC*, (referred to as *ydeB* in GenBank), was close to background in the experimental sample. Thus it does not appear to be affected by MarA under these conditions.

The *mar*-regulated genes identified are dispersed throughout the chromosome and are involved in a wide range of cell functions (Fig. 2, Table 2). In Figure 2, the internal circle represents the chromosome of *E. coli* K-12 MG1655 divided in intervals of 1 minute, while the external is divided in intervals of 100,000 nucleotide residues (adapted from Blattner et al. (Blattner, F. R., Plunkett, G. I. I., Bloch, C. A., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. M., G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997) *Science* 277, 1453-1462)). Genes induced by *mar* are plotted to face the exterior of the chromosome and genes repressed by *mar* are plotted to face the interior of the chromosome. Bold faced genes read in the clockwise direction, while regular font represents those genes on the opposite strand (Blattner, F. R., Plunkett, G. I. I., Bloch, C. A., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. M., G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997) *Science* 277, 1453-1462). Genes that are in the immediate vicinity of each other were placed together over the same designation line.

In addition to changing the expression of genes with known functions, MarA also changed the expression of genes yet uncharacterized. For instance the gene bO447 encodes a putative LRP-like transcriptional regulator, *yadG* encodes a putative ATP-binding component of a transport system, while bI448 and *yggJ* have no known homologues. It is not clear how all these genes relate to each other in the development of the Mar phenotype. *gshB* is involved in the synthesis of glutathione, which is part of the cell's antioxidants defenses (Hidalgo, E. & Demple, B. (1995) in *Regulation of gene expression in Escherichia coli.*, eds. Lin, E. C. & Lynch, A. S. (R. G. Landes Company, Austin), pp. 433-450), and among other functions, is involved in the reduction of OxyR to its normal redox state (Chater, K. F. & Nikaido, H. (1999) *Curr. Opin. Microbiol.* 2,

- 97 -

121-125) and in the detoxification of toxic electrophiles (Ferguson, G. P. (1999) *Trends Microbiol.* 7, 242-247). The induction of *gshB* by MarA could help to explain why resistance to oxidative stress is a Mar phenotype.

5 **Example 2. Confirmation of previously identified *mar* regulated genes.** The differential expression of most of the genes previously identified as part of the, *mar* regulon, e.g. *inaA*, *sodA*, *ompF*, *zwf* and *fumC* (Ariza, R. R., Cohen, S. P., Bachhawat, N., Levy, S. B. & Demple, B. (1994) *J Bacteriol.* 176, 143-148, Greenberg, J. T., Chou, J. H., Monach, P. A. & Demple, B. (1991) *J. Bacteriol.* 173, 4433-4439, Jair, K.-W.,
10 Martin, R. G., Rosner, J. L., Fujita, N., Ishihama, A. & Wolf, J. R. E. (1995) *J. Bacteriol.* 177, 7100-7104, Rosner, J. L. & Slonczewski, J. L. (1994) *J. Bacteriol.* 176, 6262-6269), was confirmed in the current study (Table 3). A major role in the Mar phenotype is played by the efflux system *acrAB*, which acts by pumping toxic compounds out of the cell (White, D. G., Goldman, J. D., Demple, B. & Levy, S. B. (1997) *J. Bacteriol.* 179, 6122-6126, Moken, M. C., McMurry, L. M. & Levy, S. B. (1997) *Antimicrob. Agents Chemother.* 41, 2770-2772, Okusu, H., Ma, D. & Nikaido, H. (1996) *J. Bacteriol.* 178, 306-308). An increase in the expression of the *acrA* gene of the *acrAB* operon was also observed (Table 3), however the expression values for *acrB* were not above background. As described earlier for *marB*, this kind of finding is not fully
15 understood, but could arise from differential processing of the polycistronic transcript and/or by slight differences in transcript stability.

Previous studies suggest co-ordinate activation of TolC and the AcrAB efflux pump in the development of the Mar phenotype, particularly in the context of organic solvent tolerance (Fralick, J. A. (1996) *J Bacteriol.* 178, 5803-5805, Aono, R.,
25 Tsukagoshi, N. & Yamamoto, M. (1998) *J. Bacteriol.* 180, 938-944). Changes in the expression of outer membrane proteins (e.g. increased OmpX, and decreased OmpF and LamB) have also been reported in *E. coli marR* mutants and wild type strains over-expressing MarA (Aono, R., Tsukagoshi, N. & Yamamoto, M. (1998) *J. Bacteriol.* 180, 938-944). MarA expression is shown herein to increase the transcription of both
30 *tolC* and *ompX* (Table 3). Although a decrease in the levels of *ompF*, was observed, there was no evidence for a similar decrease in *lamB* expression, suggesting that LamB

- 98 -

may not be the underproduced protein identified in the earlier study (Aono, R., Tsukagoshi, N. & Yamamoto, M. (1998) *J. Bacteriol.* 180, 938-944).

Transcription of the previously identified *mlr1* (b1451) and *mlr2* (b0603) genes (Seoane, A. S. & Levy, S. B. (1995) *J Bacteriol.* 177, 530-535) was increased in the *mar* expression strain in two experiments, but appeared to be unaffected in a third experiment, so they were not included in Table 3. Expression of the *slp* gene, previously described as repressed by MarA (Seoane, A. S. & Levy, S. B. (1995) *J Bacteriol.* 177, 530-535) was so low that any *mar*-mediated changes would have been difficult to detect. This latter observation may reflect these experiments being performed on cells in mid-logarithmic phase while *slp* is a stationary phase inducible gene. Since the identity of the two *mar*-responsive genes *soi-17* and *soi-19* (Greenberg, J. T., Chou, J. H., Monach, P. A. & Demple, B. (1991) *J. Bacteriol.* 173, 4433-4439) remains to be determined, their differential expression could not be confirmed by the macroarrays analysis.

Example 3. Relationship between *soxRS* and *mar* regulons. SoxS is the ractivator of the *soxRS* regulon (Demple, B. (1996) *Gene* 179, 53-57), which mediates a cellular response to oxidative stress, and, like MarA, is a member of the XylS/AraC of transcriptional activators (Gallegos, M.-T., Schleif, R., Bairoch). Many oxidative stress genes, that are known to respond to SoxS, are also responsive to MarA (Jair, K.-W., Martin, R. G., Rosner, J. L., Fujita, N., Ishihama, A. & Wolf, J. R. E. (1995) *J. Bacteriol.* 177, 7100-7104, Miller, P. F., Gambino, L. F., Sulavik, M. C. & Gracheck, S. J. (1994) *Antimicrob. Agents Chemother.* 38, 1773-1779). Conversely, SoxS is able to confer a Mar phenotype via activation of genes that are under the control of MarA (Ariza, R. R., Cohen, S. P., Bachbawat, N., Levy, S. B. & Demple, B. (1994) *J Bacteriol.* 176, 143-148, Greenberg, J. T., Chou, J. H., Monach, P. A. & Demple, B. (1991) *J. Bacteriol.* 173, 4433-4439). Genes known to be regulated directly or indirectly by both the MarA and SoxS regulators include *zwf*, *fpr*, *fumC*, *micF*, *nfo*, *inaA*, *sodA* and *acrA* (Ariza, R. R., Cohen, S. P., Bachbawat, N., Levy, S. B. & Demple, B. (1994) *J Bacteriol.* 176, 143-148, Greenberg, J. T., Chou, J. H., Monach, P. A. & Demple, B. (1991) *J. Bacteriol.*

- 99 -

173, 4433-4439, Jair, K.-W., Martin, R. G., Rosner, J. L., Fujita, N., Ishihama, A. & Wolf, J. R. E. (1995) *J. Bacteriol* 177, 7100-7104, Rosner, J. L. & Slonczewski, J. L. (1994) *J. Bacteriol.* 176, 6262-6269, Liochev, S.I. & Fridovich, I. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5892-5896, Ma, D., Alberti, M., Lynch, C., Nikaido, H. & Hearst, J. E. (1996) *Mol. Microbiol.* 19, 101-112). The positive regulation of *zwf*, *fumC*, *acrA*, *inaA* and *sodA* by *mar*, and also the down-regulation of *ompF* is confirmed by these results. However, although binding of MarA to *nfo* and *fpr* was shown in cell-free studies (Jair, K.-W., Martin, R. G., Rosner, J. L., Fujita, N., Ishihama, A. & Wolf, J. R. E. (1995) *J. Bacteriol* 177, 7100-7104), no significant change in expression of these two genes was detected using the experimental conditions employed here.

Other findings revealed further overlap between the *mar* and *soxRS* regulons. The levels of aconitase (*acnA*), GTP cyclohydrolase II (*ribA*) genes, and the major oxygen insensitive nitroreductase (*nfsA/mdaA*), previously known to be under the control of *soxRS* (Gruer, M. J. & Guest, J. R. (1994) *Microbiology* 140, 2531-2541, Koh, Y. S., Chung, W-H., Lee, J.-H. & Roe, J.-H. (1999) *Mol. Gen. Cent.*, 374-380, Liochev, S.I., Hausladen, A. & Fridovich, I. (1999) *Proc. Natl. Acad. Sci. USA* 96, 3537-3539), were observed to be increased in *mar*-expressing strains (Table 3). While NfsA was shown to be the major isoenzyme affected by paraquat, the oxygen sensitive NAD(P)H nitroreductase B, *nfnB* (also designated *nfsB*), was shown to be slightly induced (Liochev, S.I., Hausladen, A. & Fridovich, I. (1999) *Proc. Natl. Acad. Sci. USA* 96, 3537-3539). *nfnB*, like *nfsA*, is under the positive control of *mar* (Table 3).

nfsA was initially designated *mdaA* (modulator of drug activity), as one of two genes associated with bacterial resistance to tumoricidal compounds (Chatterjee, P. K. & Sternberg, N. L. (1995) *Proc. Natl. Acad. Sci. USA* 92, 8950-8954). The other gene, designated *mdaB*, was also found to be affected by *mar* (Table 3). Information about *mdaB* is very limited, and its function remains unknown. These findings provide suggestive evidence for a putative physiological role in protection against environmental stresses.

The exact mechanisms for the overlapping regulation by MarA and SoxS are still poorly understood. Multiple antibiotic resistance encoded by the *soxRS* locus appeared partly dependent on an intact *mar* locus; strains overexpressing SoxS showed increased

- 100 -

levels of *mar RAB* transcription (Miller, P. F., Gambino, L. F., Sulavik, M. C. & Gracheck, S. J. (1994) *Antimicrob. Agents Chemother.* 38, 1773-1779). On the other hand, other work showed that regulation of some genes by *mar* and by *soxRS* can occur through independent pathways, e.g. *inaA* (Rosner, J. L. & Slonczewski, J. L. (1994) *J. Bacteriol.* 176, 6262-6269). An effect of *mar* on *soxRS* has not been detected and no up-regulation of *soxS* expression by *mar* was observed. Therefore, MarA appears to operate independently of SoxS.

Rob, a MarA/SoxS homologue, is also able to bind to promoters of genes belonging to the *mar*-regulon and overexpression of this protein leads to multiple antibiotic resistance and organic solvent tolerance in *E. coli* (Ariza, R. R., Li, Z., Ringstad, N. & Demple, B. (1995) *J. Bacteriol.* 177, 1655-1661, Jair, K. W., Yu, X., Skarstad, K., Thony, B., Fujita, N., Ishihama, A. & Wolf, R. E. J. (1996) *J. Bacteriol.* 178, 2507-2513). No substantial change in expression of *rob* by MarA was found.

Example 4. *mar* regulation of operons and co-transcribed units. Some of the *mar*-regulated genes were clustered in discrete regions, as part of documented or predicted operons (Blattner, F. R., Plunkett, G. I. I., Bloch, C. A., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. M., G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997) *Science* 277, 1453-1462) (Fig. 2). Interestingly, considerable variability in the levels of expression of different genes from the same operon was observed, and therefore only some of these genes were eligible for listing in Table 3. For example, the fold increase in expression of the three genes in the tryptophanase operon (*tnaLAB*; 83.8 min) was 1.7 for *tnaL* and 8.1 for *tnaA* (averaged values), while *tnaB* was unclear; it gave background values in one experiment, but was clearly up-regulated in the other two experiments.

Differential expression of genes within *mar*-regulated operons could arise as result of other factors besides regulation of transcriptional initiation, e.g. differences in mRNA stability or the presence of regulatory secondary structures in the intercistronic regions of the operon. For example, the β -methylgalactoside (*mgl*) transport operon is composed of three Orfs, *mglBAC*. Northern analysis showed the presence of two transcripts, a polycistronic *mglBAC* mRNA and a smaller transcript which corresponds

- 101 -

to the first gene in the operon, *mglB* (Hogg, R. W., Voelker, C. & von Carlowitz, I. (1991) *Mol. Gen. Genet.* 229, 453-459). This finding was suggested to result from 3'-5' degradation of the larger mRNA, and from protection of the smaller transcript against nucleases by a repetitive extragenic palindrome sequence located at its 3' end. In agreement, these findings showed the smaller transcript at a much higher level than the larger one (Fig. 3). In Figure 3, eight genes up-regulated by *mar*: *acnA*, *gshB*, *hemB*, *mdaA*, *tpx*, *mglB*, *nfnB* and *yadG*, and 2 genes down-regulated by *mar*: *aceE* and *ndh*, were selected from those listed in Table 3. Samples were prepared and run in duplicate from *mar*-expressing (*mar*⁺) and *mar*-deleted (Δ *mar*) cells. RNA samples were transferred to nylon membranes and hybridized to ³²P-labeled PCR amplified probes of the genes in study

The only members of the *mar* regulon which appear to have a paralog in the *E. coli* genome are *acrA*, *pflB*, *ompF*, *marA* and *mtr* (<http://www.genetics.wisc.edu/>). However, with the possible exception of *mtr* vs. *tnaB*, none of the paralogs for these genes was identified as being regulated by *mar*, and therefore artifacts of cross-hybridization with other genes sharing substantial sequence homology (Richmond, C. S., Glasner, J. D., Mau, R., Jin, H. & Blattner, F. R. (1999) *Nucleic Acids Res.* 27, 3821-3835) do not appear to account for the observed findings.

Mar regulation of neighboring genes which are not part of previously documented operons was also observed (Tables 3 and Fig. 2). Up-regulation of *gshB* (min 66.6) expression by *mar* was routinely observed; moreover, *yggJ* whose function remains unknown, and is located immediately upstream from *gshB*, and the Orf downstream from *gshB*, *yqgE* (b2948), were also up-regulated by MarA. There are only 13 bp between the end of *yggJ* and the beginning of *gshB*, and 37 bp between *gshB* and *yqgE*, which does not allow for the presence of promoter sequences in the respective intergenic regions. These results support the annotation of these three genes as a "predicted operon" (Blattner, F. R., Plunkett, G. I. I., Bloch, C. A., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. M., G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997) *Science* 277, 1453-1462).

- 102 -

Transcription of the gene *ybjC*, a small Orf immediately upstream from , *nfsA*, also seems to be up-regulated by MarA. A promoter sequence internal to *ybjC* and near its start codon has been proposed for *nfsA* (44). Thus, *nfsA* could be transcribed independently from this promoter but the resulting transcript would hybridize to both
 5 genes in the array. On the other hand, the *E. coli* genome sequence suggests that these two genes may form an operon (Blattner, F. R., Plunkett, G. I. I., Bloch, C. A., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. M., G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997) *Science* 277, 1453-1462). The two genes downstream from *nfsA*, *rimK*
 10 and b0853, are also up regulated by MarA. A putative transcriptional terminator has been identified in the intergenic region of *nfsA* and *rimK* (Zenno, S., Koike, H., Kumar, A. N., Jayarman, R., Tanokura, M. & Saigo, K. (1996) *J. Bacteriol.* 178, 4508-4514). Nevertheless, a certain level of read-through transcription would explain the co-expression of this complex of genes.

15

Example 5. Relationship between the *mar* regulon and iron. Some of the genes regulated by MarA are associated with iron, e.g. *hemB*, *fumC*, *fecA*, *acnA*, *sodA*. The products of some of the genes contain iron-sulfur clusters, which play a major role in sensing O₂ and iron, and in regulatory functions (Beinert, H. & Kiley, P. J. (1999)
 20 *Curr. Opin. Chem. Biol.* 3, 152-157) (Ding, H. & Demple, B. (1998) *Biochemistry* 37, 17280-17286). Iron is an essential element for the bacterial cell (Earhart, C. F. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, eds. Neidhardt, F. C. (ASM Press, Washington, DC), pp. 1075-1090) and iron acquisition from the host is important in bacterial pathogenesis (Litwin, C. M. & Calderwood, S. B. (1993) *Clin.*
 25 *Microbiol. Rev.* 6, 137-149) (Mahan, M. J., Slauch, J. M. & Mekalanos, J. J. (1996) in *Escherichia coli and Salmonella. Cellular and Molecular Biology*, eds. Neidhardt, F. C. (ASM Press, Washington, DC), pp. 2803-2815). However, iron can also be harmful to the bacterial cell as it catalyzes the production of hydroxyl ions via the Fenton reaction, which may damage all cellular components and even lead to cell death (Zheng, M.,
 30 Doan, B., Schneider, T. D. & Storz, G. (1999) *J. Bacteriol.* 181, 4639-4643).

- 103 -

Some genes known to be regulated by Fur (ferric uptake regulator), are also responsive to SoxS, MarA and other regulators e.g. *acnA* and *sodA* (Cunningham, L., Gruer, M. J. & Guest, J. R. (1997) *Microbiology* 143, 3795-805) (Storz, G. & Imlay, J. A. (1999) *Curr. Opin. Microbiol.* 2, 188-194). This co-regulation would allow the cell to
 5 deal with the iron-associated oxidative stress and suggest a role for *mar* in bacterial pathogenesis.

Example 6. Northern blot analysis of selected genes. Ten newly identified *mar*-regulated genes, whose expression was either induced (*tpx*, *acnA*, *mglB*, *mdaA*,
 10 *gshB*, *hemB*, *yadG* and *nfnB*), or repressed (*aceE* and *ndh*) in the macroarrays were confirmed by Northern blot analysis. This showed changes in the expression of mono or polycistronic transcripts associated with the genes (Fig. 3). The magnitude of these changes, not unexpectedly, differed somewhat from that obtained for the macroarrays. Regulation of *gshB*, *mdaA* and *aceE* genes involved alteration in the levels of multiple
 15 transcripts as expected based on reported or predicted involvement of these genes in polycistronic elements (Blattner, F. R., Plunkett, G. I. I., Bloch, C. A., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K. M., G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997) *Science* 277, 1453-1462), (Spencer, M. E. & Guest, J. R. (1985) *Mol. Gen. Genet.*
 20 200, 145-154), (Quail, M. A., Ilaydon, D. J. & Guest, J. R. (1994) *Mol. Microbiol.* 12, 95-104).

The transcriptional activator MarA may control the expression of genes directly or indirectly. It could activate intermediate activator or inhibitor regulatory proteins which then could up- or down-regulate the expression of other genes in the regulon. A
 25 case in point is the *mar*-regulation of *ompF* mentioned earlier (Cohen, S. P., McMurry, L. M. & Levy, S. B. (1988) *J Bacteriol.* 170, 5416-5422). MarA activates *micF*, an antisense RNA which negatively affects the translation of *ompF*, leading to decreased outer membrane porin OmpF (Cohen, S. P., Hachler, H. & Levy, S. B. (1993) *J Bacteriol.* 175, 1484-1492, Cohen, S. P., McMurry, L. M., I-looper, D. C., Wolfson, J.
 30 S. & Levy, S. B. (1989) *Antimicrob. Agents Chemother.* 33, 1318-1325). Furthermore, transcriptional activators can act also as repressor proteins, depending on the position of

- 104 -

the regulator binding site at the exclusive zone of repression (Gralla, J. D. & Collado-Vides, J. (1996) in *Escherichia coli and Salmonella. Cellular and Molecular Biology*, eds. Neidhardt, F. C. (ASM Press, Washington, DC), pp. 1232-1245).

Only those genes whose expression trends were consistent in three experiments
5 are reported here. It is therefore likely that the size of the *mar* regulon is under estimated. Some of the genes containing putative marboxes in their promoter regions (Martin, R. G., Gillette, W. K., Rhee, S. & Rosner, J. L. (1999) *Mol. Microbiol.* 34, 431-441) were not shown under the conditions used here to be part of the *mar* regulon. Moreover, a large number of genes was expressed at background level or responded to *mar*
10 expression with small changes that were below the threshold applied in this study and therefore were not included. Under a different set of experimental conditions, such as using cells in a different stage of the growth phase, or grown in different media, it is possible that the magnitude of these changes will increase, or new genes will be affected, justifying inclusion in the *mar* regulon. Certainly small and transient changes in gene
15 expression could have important implications in the cell's response to external stresses. Differences observed in global expression analysis between experiments have been seen and extensively addressed by other authors (Richmond, C. S., Glasner, J. D., Mau, R., Jin, H. & Blattner, F. R. (1999) *Nucleic Acids Res.* 27, 3821-3835) (Tao, H., Bausch, C., Richmond, C., Blattner, F. R. & Conway, T. (1999) *J. Bacteriol.* 181, 6425-6440).
20 Among other factors the authors observed that the signal intensity of some genes was significantly different between experiments when using different batches of RNA. This problem was addressed in part by performing the study in triplicate. Trends detected by the gene array method must, therefore, be analyzed by other available molecular and biochemical techniques, such as Northern blot analysis and promoter fusion studies.

25

Table 3. Genes identified as part of the *mar* regulon using the *E. coli* Panorama gene arrays.

Gene name	Product*	MarA regulation
	Up-regulated genes	
<i>acnA</i>	Aconitate hydratase 1	2.7/5.9
<i>acrA</i>	Acridine efflux pump	1.9/2.3
<i>aldA</i>	Aldehyde dehydrogenase, NAD-linked	7.4/3.2
<i>b0447</i>	Putative LRP-like transcriptional regulator	3.5/4.4
<i>b0853</i>	Putative sensory transduction regulator	1.4/4.2
<i>b1448</i>	Putative resistance protein	1.8/2.3
<i>b2889</i>	Putative enzyme	2.5/5.6
<i>b2948</i>	Orf; hypothetical protein	1.4/2.5
<i>cobU</i>	Cobinamide kinase/cobinamide phosphate guanylyltransferase	1.6/2.2
<i>fumC</i>	Fumarate hydratase Class II; isoenzyme	2.5/2.9
<i>galK</i>	galactokinase	1.5/2.0
<i>galT</i>	Galactose-1-phosphate uridylyltransferase	2.5/2.4
<i>gatA</i>	Galactitol-specific enzyme IIA of phosphotransferase system	2.0/1.8
<i>gatC</i>	PTS system galactitol-specific enzyme IIC	3.4/1.6
<i>gltA</i>	Citrate synthase	2.1/1.9
<i>gshB</i>	Glutathione synthetase	3.5/5.7
<i>hemB</i>	5-aminolevulinate dehydratase=prophobilinogen synthase	5.7/5.1
<i>inaA</i>	pH-inducible protein involved in stress	5.0/20.2

- 106 -

	response	
<i>map</i>	Methionine aminopeptidase	1.7/2.1
<i>marA</i>	Multiple antibiotic resistance; transcriptional activator of defense systems	24.0/46.6
<i>marB</i>	Multiple antibiotic resistance protein	7.5/16.3
<i>marR</i>	Multiple antibiotic resistance protein; repressor of <i>mar</i> operon	15.9/46.3
<i>mdaA</i>	Modulator of drug resistance A	3.8/8.2
<i>mdaB</i>	Modulator of drug resistance B	5.5/8.2
<i>mglB</i>	Galactose-binding transport protein; receptor for galactose taxis	5.3/2.6
<i>mtr</i>	Tryptophan-specific transport protein	1.3/2.2
<i>nfnB</i>	Oxygen-insensitive NAD(P)H nitroreductase	12.4/20.1
<i>ompX</i>	Outer membrane protein X	1.6/2.1
<i>pflB</i>	Formate acetyltransferase 1	2.1/2.2
<i>pgi</i>	Glucose-6-phosphate isomerase	2.4/2.1
<i>ribA</i>	GTP cyclohydrolase II	1.1/2.2
<i>ribD</i>	Bifunctional pyrimidine deaminase/reductase in pathway of riboflavin synthesis	1.7/2.5
<i>rimK</i>	Ribosomal protein S6 modification protein	1.6/3.0
<i>sodA</i>	Superoxide dismutase, manganese	7.0/4.6
<i>slA_2</i>	PTS system, glucitol/sorbitol-specific IIB component and second of two IIC component	3.0/2.0
<i>tnaA</i>	Tryptophanase	7.9/8.4
<i>tnaL</i>	Tryptophanase leader peptide	1.3/2.1

- 107 -

<i>tolC</i>	Outer membrane channel; specific tolerance to colicin E1; segregation of daughter chromosomes	3.1/2.8
<i>tpx</i>	Thiol peroxidase	2.1/1.6
<i>yadG</i>	Putative ATP-binding component of a transport system	9.2/11.2
<i>yadH</i>	Orf; hypothetical protein	1.9/2.7
<i>ybjC</i>	Orf, hypothetical protein	6.7/17.4
<i>ydeA</i>	Putative resistance/regulatory protein	1.9/3.9
<i>yfaE</i>	Orf, hypothetical protein	2.5/5.9
<i>yggJ</i>	Orf, hypothetical protein	3.1/4.2
<i>yhbW</i>	Putative enzyme	10.6/6.5
<i>zwf</i>	Glucose-6-phosphate dehydrogenase	2.7/1.8
	Down-regulated genes	
<i>accB</i>	Acetyl-CoA carboxylase, BCCP subunit; carrier of biotin	2.2/2.0
<i>aceE</i>	Pyruvate dehydrogenase (decarboxylase component)	6.1/5.2
<i>aceF</i>	Pyruvate dehydrogenase (dihydro lipoltransacetylase component)	5.1/4.1
<i>ackA</i>	Acetate kinase	1.8/2.6
<i>b0357</i>	Putative alpha helix chain	3.2/2.2
<i>b2530</i>	Putative aminotransferase	1.2/2.3
<i>b3469</i>	Zinc-transporting ATPase	1.6/2.2
<i>fabB</i>	3-oxoacyl-[acyl-carrier-protein] synthase I	2.6/3.1
<i>fecA</i>	citrate-dependent iron transport , Outer membrane receptor	2.5/2.8
<i>glpD</i>	Sn-glycerol-3-phosphate	1.4/2.1

- 108 -

	dehydrogenase (aerobic)	
<i>guaB</i>	IMP dehydrogenase	2.9/2.3
<i>ndh</i>	Respiratory NADH dehydrogenase	5.8/3.8
<i>ompF</i>	Outer membrane protein 1a (Ia;b;F)	2.7/3.0
<i>purA</i>	Adenylosuccinate synthetase	2.1/2.1
<i>rplE</i>	50S ribosomal subunit protein L5	3.5/2.0

*Information about individual genes was obtained through the *E. coli* K-12 genome project Web page (<http://www.genetics.wisc.edu/>). *mar* regulation corresponds to ratios of gene expression between experimental and control samples for the up-regulated and the opposite for the down-regulated genes, obtained from two independent experiments.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, numerous equivalents to the specific polypeptides, nucleic acids, methods, assays and reagents described herein. Such equivalents are considered to be within the scope of this invention and are covered by the following claims.

- 1 -

SEQUENCE LISTING

<110> TRUSTEES OF TUFTS COLLEGE

<120> NIMR COMPOSITIONS AND THEIR METHODS OF USE

<130> PKZ-043PC

<140>

<141>

<150> 60/188,362

<151> 2000-03-10

<160> 98

<170> PatentIn Ver. 2.0

<210> 1

<211> 14000

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (9174)..(9644)

<400> 1

cactccttgt ctgctaaaaa tgtgactcaa aaaccctttg ccg gatggcg gccagcatc 60
tgtttacatc attacgacgt caaactgctc ctggttatag agcgggtcaa tttgtacttt 120
aacctgtttg ccaacgaaaa tttccacttc cgccagcgag tgtgactcct cgcctttcaa 180
ggcttcagct actgccggag aagcatagac caggaaacgg tcggagtcgt aagcatggtg 240
gacacgaaca atctcgcgca tgatttcata gcataccgtt tccacggttt tcaccgttcc 300
gcgaccgtgg cagggtgggc attcggtaca cagtacgtgc tcaatgcttt cgcgggtgcg 360
tttacgcgtc atctccacca gcccagcgc cgaaaaacca ttaacgctgg ttttcacccg 420
gtctttgctc aacgcctgct ccagcgagtg cagcactcgg cggcggtgat cttcattatt 480
catatcgatg aaatcaataa tgataatccc gccagatta cgcaaccgta actggcgagc 540
gatagcctgc gtcgcttcaa tattggtatt gaaaatggtg tcgtccagat tgcgatgacc 600
gacaaacgct ccggtattga tgtccacggt ggtcatcgct tcggtctggt cgataatgag 660
ataaccaccg gatttcagtt ctactttgcg ttccagcgct cgctggattt cgttttcgac 720
atcaaagaga tcgaaaatcg gctggcgctc tgtgtaatgc tccagcttgc ttgtcatctc 780
gggaatgtac tccgaggtga actcaagtaa cgcttcgtaa gtcaggcgtg agtcaacgcg 840
aatgcggtcc agttcggcat cggcgaaatc acgcagaaca cgctgcgcca gcgccagttc 900
gccgtacagc tgataacggg tctgcgggcg ttttttacgc tccataactt tgggtccagac 960

- 2 -

gcgtttcaga taagcggcat cggaggccag ttcagcctcg ccaacccctt ccgctgcggt 1020
acggatgata aaccgcgcct gctcgtcgca atactctgcg accacttttt tcaggcggttc 1080
acgttctgat tcgctttcaa tacgttggga aaccccaacg tgagaagccc ctggcataaa 1140
caccagatag cgagaaggga gcgtgatatc ggtggtcagg cgcgcacctt tagtgccaag 1200
cggatctttc accacctgca ccatcagatc ttgcccctga cgaaccagtt ccgagatgtc 1260
gcgcaccgtg aattgctttt gttcttcacc cgccacacat tcggtgtgcg gcatgatgtc 1320
ggatgcatga agaaacgcgg ctttatccag cccaatatct acaaaagccg cctgcatacc 1380
cggaagtaca cgacttacac gacccttgta gatattgcct actattccgc gtcgcgcctc 1440
acgttcaata tgaatttcct gcagaatacc gccatcaata tacgccactc gcgtttccga 1500
aggcgttacg tttactaaca attcagccgt catgtttatc ccttttctca cgcagtgcgt 1560
taaaattact taataattca tacgtttcaa ccagcggtaa gccgactacg gcgtgatagc 1620
tgccatttat ctctctgaca aaacagccac ccagcccctg aataccgtat gcacctgctt 1680
tatctaacgg ttcacgctg gcgacatagc ccgcatgtc ttctgtctgtt aacgttctga 1740
aagtcacatc ggtgaccacc aggcaatcga gaatgtgctg gctgtcggcc aacgccactg 1800
ctgtcatcac ctgatgggtc tgacccgata atttgcgcaa catctgcgcc gcatgctctg 1860
cgtcgcgcgg tttctccagc acttctccgt tcaggataac gatagtatcc gcaccagca 1920
ccgggagatc cttegcggtt tgcgcgacac ctgcccgtgc tttctcgcgc gccagacgca 1980
caacatactg ctgcgcgctc tcctgcggct gacgctgctc ctcaatgccc gtaacaatac 2040
gttcaaaggt cagccaagt tgcgcaagta actcctgacg acgcggagaa ccggaagcta 2100
aatacagaga agtcatagaa acctttattg cactgcaaac tgctgacgga ctttgcgcat 2160
cagcaagaaa atccacggcc agagcaccoc attgactaca ctactccaga aacttccgg 2220
tctgaaagag acgttaatca ctaaaaactc tgcccagaaa acaataatat ccaccaccag 2280
cgaaagcaac atgacgacca gcgcctgctg ccataatgcg aggttgcgga aaagctggta 2340
tttcagcgcc accaggtaag caatgatgct catcgccaat acgcgtacgc caagcgtcga 2400
gccgctgatc agatccagta tggcaccat caaaaacct gtgccacat ttacgcgatg 2460
aggcaaggcc aggatccaat acaacaagat gagtaacacc cagtttggcc ggaaaacaat 2520
caggttatcc ggccaggga tgatttgag caacagcgca atgaggaaag agagccagat 2580
taccagcgt ccctggctac gatagctcgc cactattgcc ctcccggcgc acgcgcaggc 2640
ggttgagcac cactttgcgg cggcgtagcc ctttgtggag agcgattagc agcaggctgt 2700

- 3 -

gtcggcgag caggcgag agctgcattt cctgtcgccg gttgctgcgg agtcggctga 2760
gcgatccccg ttgccgggtt aggttaacttt ggccccatcg cgtctggcga aggcaatacc 2820
tgcggcatca tctgcatcag acgttcatta gcaacacgat gcacctcttc cggcgatcatc 2880
gggttagcgc cgttacgatc tgccccccac agcagcagca gataacgcaa acgttgcagc 2940
cctgcagtcg gacgcgcctg aatcacagta taagcgcgct gggatatcgag ttttacggaa 3000
gagacaaccg cgaccggata gccttcgggg aaacgaccgc ccagaccgga agtcaccagc 3060
acatcaccaa cacgaatata cgtattcgcc ggcagatgct caagctgcaa atcatccgta 3120
caaccgttac cggctgcaat tacgcggata tcgttgcgca gcacctggat tggcagcgcg 3180
tgggtcgcgt cacaatcag cagcagcga ctggtcagtt tagcgacggc caccacctga 3240
ccaacaacac ctttgtcgct gatgaccggc tggccttcatt aaacgccatt aacgctacct 3300
ttatcgataa caacttgatc gctataagga tcgttaaccg tggagataac ctgagtcacc 3360
atcttctgct catcctgacg cagcggggaa cccagcagct cgcgcagacg cgcgttctcc 3420
tgtttgtatt gtccaagcat cagcagttca ctgtttttca gcaacagttc ctgacgtaac 3480
gcccggtttt caagttctaa ttggtcacgc gaggccagcg tctgcgatac gccatccagc 3540
aattcacgag gagcattgga aacaaagtag aaaggactga cggcggtatc catataagta 3600
cggatttgac tgaacgtccc caggcggctg tcggcaataa taatgccgag cgccaccagc 3660
accgccagaa taaggcgaat ctgtagcgac gggccacggc taaaaattgg cttcataagt 3720
tatgcgtatt ctcgtatcag accaggcagg gtaaacagac acttcccctg cctgcatccg 3780
attactcttc gctgaacagg tcgccgccgt gcatgtcgat catttccagc gctttgccgc 3840
caccgcgcgc cacacaggtc agcgggtctt cagcaacaac gactggaatg ccggtttctt 3900
ccattaacaa acggtcaagg ttacgcagca gtgcgccacc accggtgagc accatgccgc 3960
gctcggagat gtcggaagcc agttccggcg ggcactgttc cagtgcacc attaccgcgc 4020
tcacaatacc ggtcagcggg tcctgcagtg cttcgaggat ttcattggag ttcagggtaa 4080
aacgcgtgg aacaccttct gccaggttac ggccacgaac ttcgatttca cggacttcat 4140
cgcccgata agccgaaccg atttcgtgct tgatacgttc tgccgtggct tcaccgatca 4200
gagaaccgta attacgacgc acatagttga tgatagcttc gtcgaaacgg tcaccaccaa 4260
tgcgcacaga agaggagtaa accacaccgt tcaaggagat aacagcaact tcagtggtag 4320
caccaccgat atcaaccacc atagaaccgg tcgcttcaga aaccggcagg ccagcaccaa 4380
ttgcggcagc catcggttct tcaatcagga agacttcacg ggcaccagcg ccctgcgcgg 4440
attcacgaat tgcgcggcgt tcaacctggg tcgcgccaac cggcacacaa accagaacgc 4500

- 4 -

gcgggcttgg acgcataaag ctgttgctgt gcacttgttt gatgaagtgc tggagcattt 4560
tttcagtcac gaagaagtgc gcgataacgc cgtctttcat tgggcgaatg gcagcaatat 4620
tgcccggcgt acggcccagc atctgcttcg cgtcatgacc tactgcagct acgcttttcg 4680
gtgaaccggc acgatcctga cgaatggcca ccacggaagg ctcatccaat acgatgcctt 4740
gtccttttac ataaatgagg gtattcgcag taccaggtc aatggacaag tcattggaaa 4800
acatgccacg aaattttttc aacatactaa gggataatcc tgaaagctgg ggcggaaaag 4860
aaaatccgct tactttacca accacacgca gcagcgacaa ggcgcaaaaa tcacttgcta 4920
cggtgaaaat tagtgcagtt cgtttccttt gttacaaatc tctgcctgag tccagaaagg 4980
cttaatgcat cagcagcatt cctcgctgt ttgcaaccgc gtaaggatcat tcactgcat 5040
atgtgctgca acaatctggc gagcagacaa gcacactccc atgagacgca gcgcgcatta 5100
ttctacgtga aaacggatta aacggcaggt taaaccgagt atctttgtga atattttttc 5160
acgttagtat caagtggctg tgaggacgcg aaaaaatccc cttgcccgcg tgtaacaccg 5220
cgctgaatca gggctctgcca ctcgcttcgc gaacgcacgc cggtggcgta aacctgggtg 5280
ctgggtcccgg agcaggcttc caccaggctt tgaaccagca gctgggtctc cgttcgcttc 5340
tcaatgtttc tgaccagccc cggatggagc ttgagtaact caacattaag ttctttgatc 5400
caactggtac ttaccagcgt caaaccagcc tgggtgacgg ctaccgctac ccctaaagca 5460
ttcactaaac gaataacagg ttgtaaacga ctgatatgtt gacctacatc ggcctctgca 5520
agttcaataa ttatgcgttt tcgttgtgat ttttcacatt gcattaacgt atcgcgcagc 5580
caacgctgaa aacgcgggcg aatcagcgac tcaacggtaa cctgaatcgc cagattttcc 5640
tctggccagt aacgcaatag tggaataaga cggctgattt gcagacggtc atactcttcc 5700
gataagccaa actgcaagac catcggcata tactccgccc agctaacctc ttcatcaca 5760
tcgaagatgc ggcacatgag ttgcgatga tgaacctgac cttcgcgagt aaccgccggt 5820
ttttgataaa ggcgcgggccc gccgcgactg agcatttgct cgataagcgt acgccagcga 5880
acattaccgc gtcctttttc aggcaacgag tcactgtaaa tagcccagct attgccgccc 5940
tgcaatcccg cattacgcgt ggcagactct gcatgttcca ttacctgctc ggtatcctga 6000
ccactacgcc aggcgcagat accaatgtgg atcatatcgt cgcgatcagc cattttattg 6060
ttcggcaagg tatcaacggc ttgattaac tgaccggcga tgctctctgc ctcttttaac 6120
gtcgggtgcg gtaacagcgc agcaaaatca ctgcggtggg aacgcgccag cagtgcgcca 6180
gggtagcgca tcataaatgt cgacagcaga ttcgtcagag tgaagaactg ttcttcaacc 6240

- 5 -

tggctgtgcc cccaggtatc gctcaacata ttgaaatccg gcagacgaat catcatcacg 6300
atcccgtggg tacctacttt ctctgatct tccagtaacg ttgctaactg attatcgaaa 6360
aagagtcggt tattgaggcc ggttttcacg tcctgggcgg cataagagcg gatcagcgta 6420
tcaagacggc tgtgttggtc gcgtgcgttc tgaatttcac gaagcagcgt atccagcgca 6480
ctgctgggtc tgggcggcca ttcatagatg gttcccaaca cattagagcc acgctcaccg 6540
tttaagatac gagtagcccg ggtttccagc aattcttgcc cggcaagttg ccgttgtaac 6600
cagcgtaccg ccaggaagag cataacaatg ataaagccaa tcgccccgt gagcggcgcg 6660
gtggtcatca acgaatggaa atagttgcc atcggatcct gataaaccag acgcaacgac 6720
atccccgat gctttatcaa cggaacgctc agttcgcgaa acagatcgct ggagccaact 6780
ggacgataac taccatttct ggccagggtt taaacctgtt tatcaccatg gagcaggtct 6840
acacgaacga tatctgcga catcattaat tcggtaattt gtggccttaa tacgctgaag 6900
tcattcgaca caaggtgggt atcaatcgcc gtcgccaccg cctgaacgag atgactaaac 6960
ttatactgaa tggcgttgta gaaacttagc gaacagccca gcaaagtcac aaaaattgtt 7020
aaccgggtga gcagcgtaac aaaggccgaa aatttcgtcg ttaatctcat ccttggtgta 7080
actccgatag tgaggaagcg ggcatactag caaatcagat ttatctcgca atttattgag 7140
cttcateggc tttgcttttc cattagcgag tatagtcttc agaaattatt ttccaatcca 7200
tcatgcacat gaggaccact tatgcaggcg ttacttttag aacagcagga cggcaaaact 7260
ctcgcatcag tacagactct ggacgaaagt cgcctgccgg agggcgatgt cacggtcgat 7320
gttcactggc cgagcctgaa ctataaagat gcgctggcga ttaccggtaa gggaaaaatc 7380
atccgtaatt ttccgatgat tcctgggatc gatcttgccg gaactgtacg caccagcgaa 7440
gatccgcgtt ttcatgcgg tcaggagggtg ttactcactg gctggggcgt tggtgaaaac 7500
cactgggggtg ggctggcgga gcaggcgca gtgaaagggt actggctggc tgccatgccg 7560
caagggtg acgcgcgtaa agcaatgatt atcgggtactg ccggttttac cgccatgctg 7620
tgtgtgatgg cgctggaaga tgccgggtgtt cgccgcagg acggggagat tgtcgtgacg 7680
ggtgccagtg gtggcgctcg cagtaccgcc gtggcgctgc tgcataagtt gggttatcag 7740
gtcgttgccg tttccggctc cgaaagtacc catgaatata tgaaaagttt aggtgctagc 7800
cgtgttctcc ctctgatga gtttgccgaa tcccgctctc tggaaaaaca agtctgggct 7860
ggggcaattg acaccgttg cgacaaagt ctggcaaaag tgctggcgca aatgaattac 7920
ggcggctgcg tggcggcctg tggctctggcg ggtggtttta ctctgccaac cacggtcagt 7980
ccatttatc tgctgaatgt ccgtttgcaa ggggtggatt cagtaatgac gccaccagaa 8040

- 6 -

cgccgcgcac aagcctggca gcgactggtc gccgatttac cggaatcatt ctatacccag 8100
gcgggcaaaag agatatctct gtcagaggca ccgaactttg ccgaggccat cattaataac 8160
cagatccagg gtcgcacgct ggtgaagggt aactaaccat ttagcaggga ataataagag 8220
agggaaactca ttttgaaatc attattcaat cggttaacgg gaaaagcggg tagccggaca 8280
gctttcgtcg aacaccttgg tcaggaagtt atacaacatc atccaaactg gaaagtcattg 8340
atttcgactg accacaaatt gatgcgcatt gatactccac taaacagcta ttattgatac 8400
gcctccgtcg cctgttaggt ttatgttgct ttgcctgggc gacgctacgc ttagcccctt 8460
acttatttct ggtaccatgg ggtgaataat ctgattttgt ttgactacaa attaatacact 8520
cgaacctatt taatgctgag cattgtcaat cggttaattt tgcgtgcttt agcattcaca 8580
tctatccaga cgatgcagtg aaaattgggt aatccccagc aaccgctgcg taatgtcgtc 8640
tatcttgtcg cgatcctggc atccctacat tatttgtggg ctgtgaagat tatctcattg 8700
cagccccctca tcttcgcagg gctggctttt cagcttttca ccttacgtta taagaagttc 8760
cgtcgatgat ggcgctaatt tcgtgaattg tgcggcttgt tgcaaattac acggtgttga 8820
aggttattta catgttagct gttgattatc ttccctgata agaccagtat ttagctgcca 8880
attgctacga aatcggtata atgtgcgacc tcgtcctccc tgacgcagtt tttgcgctgc 8940
ggaaaagggtg acattggcgc aacgaaggta tattttgttt tttgccggag gatagcagca 9000
gatcgctgca caatgtccgt caagtctaac attgacactc tggggcaaaa tagaccggcg 9060
tcccggcctg ctggaattta tcgctatgca tacagctgtc ggggcatacg ctttacagac 9120
ggcgggtgaaa cgcctgtcac aatcacacta aacaaagagt acggaacca ctc atg 9176
Met
1

gat att cgt aag att aaa aaa ctg atc gag ctg gtt gaa gaa tca ggc 9224
Asp Ile Arg Lys Ile Lys Lys Leu Ile Glu Leu Val Glu Glu Ser Gly
5 10 15

atc tcc gaa ctg gaa att tct gaa ggc gaa gag tca gta cgc att agc 9272
Ile Ser Glu Leu Glu Ile Ser Glu Gly Glu Glu Ser Val Arg Ile Ser
20 25 30

cgt gca gct cct gcc gca agt ttc cct gtg atg caa caa gct tac gct 9320
Arg Ala Ala Pro Ala Ala Ser Phe Pro Val Met Gln Gln Ala Tyr Ala
35 40 45

gca cca atg atg cag cag cca gct caa tct aac gca gcc gct ccg gcg 9368
Ala Pro Met Met Gln Gln Pro Ala Gln Ser Asn Ala Ala Ala Pro Ala
50 55 60 65

acc gtt cct tcc atg gaa gcg cca gca gca gcg gaa atc agt ggt cac 9416
Thr Val Pro Ser Met Glu Ala Pro Ala Ala Ala Glu Ile Ser Gly His

- 7 -

70	75	80	
atc gta cgt tcc ccg atg gtt ggt act ttc tac cgc acc cca agc ccg			9464
Ile Val Arg Ser Pro Met Val Gly Thr Phe Tyr Arg Thr Pro Ser Pro			
85	90	95	
gac gca aaa gcg ttc atc gaa gtg ggt cag aaa gtc aac gtg ggc gat			9512
Asp Ala Lys Ala Phe Ile Glu Val Gly Gln Lys Val Asn Val Gly Asp			
100	105	110	
acc ctg tgc atc gtt gaa gcc atg aaa atg atg aac cag atc gaa gcg			9560
Thr Leu Cys Ile Val Glu Ala Met Lys Met Met Asn Gln Ile Glu Ala			
115	120	125	
gac aaa tcc ggt acc gtg aaa gca att ctg gtc gaa agt gga caa ccg			9608
Asp Lys Ser Gly Thr Val Lys Ala Ile Leu Val Glu Ser Gly Gln Pro			
130	135	140	145
gta gaa ttt gac gag ccg ctg gtc gtc atc gag taa cgaggcgaac			9654
Val Glu Phe Asp Glu Pro Leu Val Val Ile Glu			
150	155		
atgctggata aaattgttat tgccaaccgc ggcgagattg cattgcgtat tcttcgtgcc 9714			
tgtaaagaac tgggcatcaa gactgtcgct gtgcactcca gcgcggatcg cgatctaaaa 9774			
cacgtattac tggcagatga aacggtctgt attggccctg ctccgtcagt aaaaagttat 9834			
ctgaacatcc cggcaatcat cagcgccgct gaaatcaccg gcgcagtagc aatccatccg 9894			
ggttacggct tctctccga gaacgccaac tttgccgagc aggttgaacg ctccggcttt 9954			
atcttcattg gcccgaaagc agaaaccatt cgctgatgg gcgacaaagt atccgcaatc 10014			
gcggcgatga aaaaagcggg cgtcccttgc gtaccgggtt ctgacggccc gctgggcgac 10074			
gatatggata aaaaccgtgc cattgctaaa cgcattgggtt atccggtgat tatcaaagcc 10134			
tccggcgggc gcggcggtcg cggtatgcgc gtagtgcgcg gcgacgctga actggcacia 10194			
tccatctcca tgaccctgtc ggaagcgaaa gctgctttca gcaacgatat ggttttacatg 10254			
gagaaatacc tggaaaatcc tcgccacgtc gagattcagg tactggctga cggtcagggc 10314			
aacgctatct atctggcgga acgtgactgc tccatgcaac gccgccacca gaaagtggtc 10374			
gaagaagcgc cagcaccggg cattaccccg gaactgcgtc gctacatcg cgaacggttc 10434			
gctaaagcgt gtgttgatat cggctatcgc ggtgcaggta ctttcgagtt cctgttcgaa 10494			
aacggcgagt tctatttcat cgaaatgaac acccgattc aggtagaaca cccggttaca 10554			
gaaatgatca ccggcggttg cctgatcaaa gaacagctgc gtatcgctgc cgggtcaaccg 10614			
ctgtcgatca agcaagaaga agttcacgtt cgcggccatg cggtggaatg tcgtatcaac 10674			
gccgaagatc cgaacacctt cctgccaaagt ccgggcaaaa tcaccggttt ccacgcacct 10734			
ggcggttttg gcgtacgttg ggagtctcat atctacgcgg gctacaccgt accgccgtac 10794			

tatgactcaa tgatcggtaa gctgatttgc tacggtgaaa accgtgacgt ggcgattgcc 10854
cgcatgaaga atgcgctgca ggagctgac atcgacggta tcaaaaccaa cgttgatctg 10914
cagatccgca tcatgaatga cgagaacttc cagcatgggtg gcactaacat ccactatctg 10974
gagaaaaaac tcggtcttca ggaaaaataa gactgctaaa gcgtaaaaag gccggatttt 11034
ccggcctttt ttattactgg ggatcgacaa ccccataag gtacaatccc cgctttcttc 11094
acccatcagg gacaaaaaat ggacactcgt tttgttcagg ccataaaga ggcgcgctgg 11154
gcgctggggc tgaccctttt gtatctggca gtttggttag tagccgctta cttatctggc 11214
gttgcccccg gttttaccgg ctttcgcgc tggtttgaga tggcctgcat cctgacgccg 11274
ctgctgttta ttggactgtg ctgggcgatg gtgaaattta tctatcgca tatcccactg 11334
gaggatgacg atgcagcttg aagtaattct accgctggtc gcctatctgg tgggtggtgtt 11394
cggatatctg gtttatgca tgcgtaaacg gagcacccgc accttcctta atgagtattt 11454
cctcggcagc cgctctatgg gcggtattgt gctggcgatg acgctcacgc cgacctatat 11514
cagtgccagt tcgtttatcg gcgggccagg agctgcttat aaatacgggc tgggctgggt 11574
attgctggcg atgattcagc ttccctgcagt ctggctttca ctcggtattc tcggcaagaa 11634
gtttgcgatt cttgcgcgcc gctacaatgc agtgacgctg aacgatatgc tgtttgcccg 11694
ctaccagagt cgtcttctgg tgtggctggc gagtttgagt ttgctggttg cgttcgttgg 11754
tgcgatgacc gtgcagttta tcggcggtgc gcgcctgctg gaaaccgcgc cgggtattcc 11814
ttatgaaacc gggctgctga tttttggtat cagcattgca ttatataccg cttttggtgg 11874
ctttcgcgcc agcgtgctga acgacaccat gcaagggctt gtgatgctga ttggcacctg 11934
tgtgctgctt attggcgtag tacatgccgc tggcggtta agtaacgcag tacagacctt 11994
gcaaaccatc gatccgcaac tggttacgcc acaaggcgct gacgatattc tgcgcctgc 12054
ctttatgacg tcgttctggg tactggtgtg ttttggcgtg attggcctgc cgcatactgc 12114
ggtgcgctgt atctcttata aagacagcaa agccgtacat cgggggatca tcatcggtac 12174
gattgtggtc gcaattctga tgttcggtat gcacctggcc ggagcgttag gtcgggcggt 12234
gatccccgat ctaccgtac cggacctggt gatcccaacg ttaatggtaa aagtgctgcc 12294
accgtttgct gccgggatct tcctggctgc accgatggct gcgatcatgt cgacaattaa 12354
cgcccaactg ctgcaaagtt ccgctacgat cattaaagat ctctatctga atatccgtcc 12414
ggatcaaata caaacgaga cgcgtctgaa gcggatgtcg gcgtaatta cgttagttct 12474
cggcgcggtg ctgctgcttg ccgcctggaa gccgccagaa atgatcatct ggctgaattt 12534

- 9 -

gttggccttc ggtgggctgg aagccgtttt cctgtggccg ctggtgctgg gtctttactg 12594
 ggaacgcgcc aacgccaaag gcgcgctaag tgcatgatc gttggcggcg tgctgtatgc 12654
 cgtactcgcg acgctgaata ttcagtacct gggcttccac cctatcgtgc cctcgttact 12714
 actaagtttg ctggctttcc tggtcggaaa ccgtttcggg acatccgtcc cgcaagctac 12774
 cgttttgact actgataaat aaagagtttt gccatgcctt ggatccaact gaaactgaac 12834
 accaccggcg cgaacgcgga agatcttagc gatgcgctga tggaagcggg tgccgtttct 12894
 atcacttttc aggataccca cgatacgcca gtatttgaac cgctgccggg cgaaacgcgc 12954
 ctgtggggcg acaccgatgt gattggtctg ttcgacgctg aaaccgatat gaacgacgtg 13014
 gtggcgattc tggaaaacca tccgctgctc ggcgaggct tcgcgcataa aatcgaacaa 13074
 ctagaagata aagactggga gcgcgaatgg atggataatt tccaccgat gcgctttggt 13134
 gaacgactgt ggatctgcc tagctggcgt gatgtgccg acgaaaacgc cgtcaacgtg 13194
 atgttagatc cagggctggc gtttggtagc ggtacccatc caaccacctc tctgtgctg 13254
 caatggctcg acagcctcga ttttaaccggg aaaacagtca tcgactttgg ctgtggttcc 13314
 ggcattcttg cgatcgcggc gctgaaactg ggtgcagcaa aagccattgg tattgatata 13374
 gatccgcagg cgattcaggc cagccgcgat aacgccgaac gtaatggcgt ttctgaccgt 13434
 ctggaactct acttaccgaa agatcagcca gaagaaatga aagccgacgt ggtggctcgt 13494
 aacatccttg caggcccatt acgtgaactg gcaccgttaa tcagcgtcct gccggtttca 13554
 ggcggtttgc tgggcctttc cggatattctg gcaagccagg cagagagcgt ttgtgaagct 13614
 tatgccgata gcttcgcact ggacccggtc gtggaaaaag aagagtgggt ccgtattacc 13674
 ggtcgtaaga attaaccttc gcatcgccgt agggtgacgc ggggcgaagt gcgagcaagc 13734
 tcacaaaagg cacgtaaatt tgccgattat ttacgcaaatt ttgcgtgcca aaattttcat 13794
 tcataaagaa aaattgagaa cttactcaaa tttctttgag tgtaaatttt agtcactatt 13854
 ttctaatatg atgattttta tgagtaatta tcgcaccacg ctcattttta atgcaattct 13914
 ttgatccatc tcagaggatt ggtcaaagtt tggcctttca tctcgtgcaa aaaatgcgta 13974
 atatacgccg ccttgcagtc acagta 14000

<210> 2

<211> 156

<212> PRT

<213> Escherichia coli

<400> 2

Met Asp Ile Arg Lys Ile Lys Lys Leu Ile Glu Leu Val Glu Glu Ser

1

5

10

15

- 10 -

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

<210> 3
 <211> 10157
 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 <222> (5397)..(8060)

<220>
 <221> CDS
 <222> (8075)..(9967)

<400> 3
 aaacgggaaa gcagattccg aggtttttta tttcgttgca gcgaaagaca agaaatttgc 60
 gaggcgttac gaagaaagtt ggggaagggg agattatccg cccgcgatgg agcggataaa 120
 tctgtcaact attagcgaaa acgcattgaa aggtcgagtg cttgtacgtg tttagttagc 180
 gcaccgacgg agataaagtc cacgcccgtt tcggcaaatt cacgcagtgt tttgtcagtg 240
 acgttgccag acacttccag tagcgccttg ccgttggtgc gtttgacggc ttcgcgcata 300
 tgttctgttt cgaagttatc cagcatgatg atatcggctc ctgctttcag ggcttcatca 360
 agttcttcca gattctctac ttcgacttct actggcgcac ccgggtgcag ccaggacgct 420
 ttttcgaccg cctggcgcac tgagccggag gcaataatat ggttttcttt gatcaggaag 480

- 11 -

gcatcagaaa gccccagacg gtgattcgct ccgccgccgc aaagtaccgc gtatttcaga 540
gctgaacgca ggccgggtaa ggttttgcgc gtatccaaca actgcgtggt ggtgccttcc 600
agcaattcga catagtggcg taccttactg gcaactcctg aaagggtttg cacaaaatta 660
agcgcagtgc gttcgcccgt taacagcacg cgggatgggc cttcaagttc gaacaaggat 720
tgattggcat tgatgacatc gccgtcatcc acatgccaga ttatggtgac atcgtcgcct 780
gccagttgaa taaacacctc ttcaaccag cgtttgccgc aaaagacgcc attctcgcgg 840
gtgatcaccg tggcatgaga gcgagaattt tccggtaaaa gttttgccgt aatatcattg 900
ttggcatcga ctgttcgcc taaatcttcc cgcagcgcct gggccaccgc gccggggata 960
tcgagattaa tgcgttccag cagctcgtca cgtcgggtgt cagggttata gcggcgaggc 1020
ggcatgttaa aactccagat agctaacgaa tcataaggta gaaacatgct actctgaacc 1080
gggtattagc accacatata aggagatcct gcatgttggt agaacagggg tggctgggtg 1140
gcgcgcgccg cgttccctca ccacattacg attgccgcc ggatgacgaa acaccaccc 1200
tgctgggtggc gcacaatatt agcctgccgc caggcgagtt tggcgggtccg tggatcgacg 1260
cattattcac tggaactatt gatccgcagg cacatccttt ctttgctgag atcgcccatt 1320
tgcgcgtctc cgtcactgt ttgattcgcc gtgatgggtga aatagtccag tatgttcctt 1380
tcgataaacg tgcattggcat gcgggagtct ctcagtatca ggggcgcgaa cgctgcaatg 1440
atTTTTctat tgggattgag cttgaaggca ccgatacgct ggcgtatacc gatgcgcagt 1500
atcaacagct tgccggcggtt acgcgggcac tgattgattg ctatccggat atcgctaaaa 1560
acatgacggg ccattgtgat attgcgccgg atcggaaaac cgatcccggt cctgcatttg 1620
attgggcacg gtttcgtgtg ctggtcagca aggagacaac atgacgctat ttacaacctt 1680
actggtgtta attttcgagc gcctgtttta gttgggcgag cactggcagc ttgatcatcg 1740
tcttgaagcg ttctttcggc ggggtgaaaca tttttctctc gggcgcacgt taggcatgac 1800
cattattgcg atgggcgtga cttttttact gttacgcgca ttgcaggag tattgttcaa 1860
cgttcccacg ctactggtgt ggctgctgat tggtttgctg tgtattggcg caggtaaagt 1920
tcgtcttcat tatcatgctt atctgacagc tgcttcacgt aatgatagcc atgcccgtgc 1980
cacgatggct ggogaactca ccatgattca cggcgtcccg gcaggctgcg acgaacgtga 2040
gtatttgctg gagctgcaaa atgcattgct gtggattaac tttcgttttt atcttgcacc 2100
gctgttctgg ctgattgtgg ggggaacctg gggaccggt acgctgatgg ggtatgcgtt 2160
tttgcgtgca tggcaatact ggctggcgcg atatcagacg ccgcatcatc gtttacagtc 2220
cggcattgat gccgtgcttc atgtactgga ttgggtgccg gttcgtcttg cgggtgtggc 2280

- 12 -

atatgccttg atcgggtcatg gtgagaaagc gttaccggcc tggtttgctt cgctgggtga 2340
tttccatact tcgcagtatc aggtgttaac gcgtctggcg cagttctctc tggcgcgatga 2400
accgcatgtc gataaggtgg agacgccgaa ggcagcgggt tcaatggcga agaaaacctc 2460
gttcgtggtc gtgggtggatg ttgcactact gacgatttac ggggcggttg tgtaaagatt 2520
attgccctca ccctgtacgg gtgagggcgt agagagatta atgcgctttt acggcttttg 2580
cggttttctc tttaaacaga tagccgatac ctaacacgat cagccatacc gggatcaggt 2640
ataccgaaat cgccattcct ggggtcatca gcataatcac cagtaccgcc gccataaaca 2700
gcaggcagat ccagttaccc agcggataaa gcagagcagg gaagcgagtt accacgcctt 2760
gttcctgctt ggcgcgacgg aatttcatat gcgccaggct aatcatcgcc cagttgatta 2820
ccagtgcaga taccaccagc gccattaaca gtccgaaagc ggactctggg gcaaggtagt 2880
taatcagtac gcacaacgcc gttaccagtg cagacaccag aatggtatatt actggtacac 2940
cacgtttatc gacagacgcc agcgcttttg gcgcattacc ctgttggtgc agaccaaaca 3000
gcatacggct gttgcaatat acgcagctgt tgtacacgga gagcgccgca gtcagtacca 3060
cgatgttcag cgcattcgcc acaaaggat cgcctaactc gtggaagatc agcacaaacg 3120
gactgggtatc ggcggtaacg cgggtccacg gcatcagtga gagcagaacg gctaacgaac 3180
caatatagaa aatcaggatg cggtagataa cctgggttagt tgctttcggg atactttgct 3240
ccgggttatc agcttctgct gcgggtgatcc ccaccagttc cagaccaccg aacgagaaca 3300
tgataatcgc catcatcatc accagcccgg tgaagccgtg cggcaggaaa ccaccctgat 3360
cccacagggt gctaacggtc gcctgcgggc cgcggttgcc actgaatagc agccagccgc 3420
cgaagatgat catcgctacc accgcgataa ctttgataat ggcaaaccag aactccatct 3480
cgccaaacac tttaacgttg gtcagggtga tggcggttaat caccacaaag aatacggcgg 3540
cagaaacca ggtggggatt tccggatacc agaactgaat gtattttacc acggcagtc 3600
gctcagccat ggcaactaaa acgtacagta ccagtagtt ccagccagag gcgaaaccgg 3660
caaaactgcc ccagtattta taagcaaagt ggctaaagga gcctgcgaca ggttcttcga 3720
ccaccatttc acccagctga cgcagatca gaaaggcgat aaaaccagca atggcgtaac 3780
ccaggataat ccctggccct gcggactgta ttacggaggc gctaccagc aataaccgg 3840
tccctatcgc gccaccacgc gcgataagct gaatatggcg gtttttaagg ccgcgcttta 3900
gctgctcgcc gtgctgttga ccttccatca tgaaacctcg tgcgggtggt gtttttttga 3960
tctacgcagt gatgcgtgtg taagtttgca attccgcttg ttgtattaat ttgtttacat 4020

- 13 -

caaagaagtt tgaattgtta caaaaagact tccgtcagat caagaataat ggtatgcggc 4080
 agcgaatgca cccgctttat gcatgggtga agatgagttg cttaaaaaga aaccgtttgt 4140
 aaagctcagc ctcaaccctt ctcaatatgt agaatgaatt taaattcgtt ttaattgaat 4200
 taaaaatcac aaaattggta agtgaatcgg ttcaattcgg attttttatag ttttaataatc 4260
 gttaaaaaaac tcctttccta cgtaaagtct acatttgtgc atagttacaa ctttgaaacg 4320
 ttatatatgt caagttgtta aatgtgcac agtttcatga tttcaatcaa aacctgtatg 4380
 gacataaggt gaatactttg ttacttttagc gtcacagaca tgaaattggg aagaccaatt 4440
 gacttcggca agtggcctaa gacaggaact catggcctac agcaaaatcc gccaaccaaa 4500
 actctccgat gtgattgagc agcaactgga gtttttgatc ctccaaggca ctctccgccc 4560
 gggcgaaaaa ctcccaccgg aacgcgaact ggcaaaacag tttgacgtct cccgtccctc 4620
 cttgcgtgag gcgattcaac gtctcgaagc gaagggcttg ttgcttcgtc gccaggggtg 4680
 cggcactttt gtccagagca gcctatggca aagcttcagc gatccgctgg tggagctgct 4740
 ctccgaccat cctgagtcac agtatgactt gctcgaaaca cgacacgccc tggaagggtat 4800
 cgccgcttat tacgccgcgc tgcgtagtac cgatgaagac aaggaacgca tccgtgaact 4860
 ccaccacgcc atagagctgg cgcagcagtc tggcgatctg gacgcggaat caaacgccgt 4920
 actccagtat cagattgccg tcaccgaagc ggcccacaat gtggttctgc ttcáctctgct 4980
 aaggtgtatg gagccgatgt tggcccagaa tgtccgccag aacttcgaat tgctctattc 5040
 gcgtcgcgag atgctgccgc tggtagtagt tcaccgcacc cgcataattg aagcgattat 5100
 ggccggtaag ccggaagaag cgcgcgaagc atcgcacgc catctggcct ttatcgaaga 5160
 aattttgctc gacagaagtc gtgaagagag ccgccgtgag cgttctctgc gtcgtctgga 5220
 gcaacgaaag aattagtgat ttttctggta aaaattatcc agaagatggt gtaaatacaag 5280
 cgcataataa agcgcggcaa ctaaacgtag aacctgtctt attgagcttt ccggcgagag 5340
 ttcaatggga caggttccag aaaactcaac gttattagat agataaggaa taaccc atg 5399
 Met
 1

tca gaa cgt ttc cca aat gac gtg gat ccg atc gaa act cgc gac tgg 5447
 Ser Glu Arg Phe Pro Asn Asp Val Asp Pro Ile Glu Thr Arg Asp Trp
 5 10 15

ctc cag gcg atc gaa tcg gtc atc cgt gaa gaa ggt gtt gag cgt gct 5495
 Leu Gln Ala Ile Glu Ser Val Ile Arg Glu Glu Gly Val Glu Arg Ala
 20 25 30

cag tat ctg atc gac caa ctg ctt gct gaa gcc cgc aaa ggc ggt gta 5543
 Gln Tyr Leu Ile Asp Gln Leu Leu Ala Glu Ala Arg Lys Gly Gly Val
 35 40 45

- 14 -

aac gta gcc gca ggc aca ggt atc agc aac tac atc aac acc atc ccc	5591
Asn Val Ala Ala Gly Thr Gly Ile Ser Asn Tyr Ile Asn Thr Ile Pro	
50 55 60 65	
gtt gaa gaa caa ccg gag tat ccg ggt aat ctg gaa ctg gaa cgc cgt	5639
Val Glu Glu Gln Pro Glu Tyr Pro Gly Asn Leu Glu Leu Glu Arg Arg	
70 75 80	
att cgt tca gct atc cgc tgg aac gcc atc atg acg gtg ctg cgt gcg	5687
Ile Arg Ser Ala Ile Arg Trp Asn Ala Ile Met Thr Val Leu Arg Ala	
85 90 95	
tcg aaa aaa gac ctc gaa ctg ggc ggc cat atg gcg tcc ttc cag tct	5735
Ser Lys Lys Asp Leu Glu Leu Gly Gly His Met Ala Ser Phe Gln Ser	
100 105 110	
tcc gca acc att tat gat gtg tgc ttt aac cac ttc ttc cgt gca cgc	5783
Ser Ala Thr Ile Tyr Asp Val Cys Phe Asn His Phe Phe Arg Ala Arg	
115 120 125	
aac gag cag gat ggc ggc gac ctg gtt tac ttc cag ggc cac atc tcc	5831
Asn Glu Gln Asp Gly Gly Asp Leu Val Tyr Phe Gln Gly His Ile Ser	
130 135 140 145	
ccg ggc gtg tac gct cgt gct ttc ctg gaa ggt cgt ctg act cag gag	5879
Pro Gly Val Tyr Ala Arg Ala Phe Leu Glu Gly Arg Leu Thr Gln Glu	
150 155 160	
cag ctg gat aac ttc cgt cag gaa gtt cac ggc aat ggc ctc tct tcc	5927
Gln Leu Asp Asn Phe Arg Gln Glu Val His Gly Asn Gly Leu Ser Ser	
165 170 175	
tat ccg cac ccg aaa ctg atg ccg gaa ttc tgg cag ttc ccg acc gta	5975
Tyr Pro His Pro Lys Leu Met Pro Glu Phe Trp Gln Phe Pro Thr Val	
180 185 190	
tct atg ggt ctg ggt ccg att ggt gct att tac cag gct aaa ttc ctg	6023
Ser Met Gly Leu Gly Pro Ile Gly Ala Ile Tyr Gln Ala Lys Phe Leu	
195 200 205	
aaa tat ctg gaa cac cgt ggc ctg aaa gat acc tct aaa caa acc gtt	6071
Lys Tyr Leu Glu His Arg Gly Leu Lys Asp Thr Ser Lys Gln Thr Val	
210 215 220 225	
tac gcg ttc ctc ggt gac ggt gaa atg gac gaa ccg gaa tcc aaa ggt	6119
Tyr Ala Phe Leu Gly Asp Gly Glu Met Asp Glu Pro Glu Ser Lys Gly	
230 235 240	
gcg atc acc atc gct acc cgt gaa aaa ctg gat aac ctg gtc ttc gtt	6167
Ala Ile Thr Ile Ala Thr Arg Glu Lys Leu Asp Asn Leu Val Phe Val	
245 250 255	
atc aac tgt aac ctg cag cgt ctt gac ggc ccg gtc acc ggt aac ggc	6215
Ile Asn Cys Asn Leu Gln Arg Leu Asp Gly Pro Val Thr Gly Asn Gly	
260 265 270	
aag atc atc aac gaa ctg gaa ggc atc ttc gaa ggt gct ggc tgg aac	6263
Lys Ile Ile Asn Glu Leu Glu Gly Ile Phe Glu Gly Ala Gly Trp Asn	

- 15 -

275	280	285	
gtg atc aaa gtg atg tgg ggt agc cgt tgg gat gaa ctg ctg cgt aag Val Ile Lys Val Met Trp Gly Ser Arg Trp Asp Glu Leu Leu Arg Lys 290 295 300 305			6311
gat acc agc ggt aaa ctg atc cag ctg atg aac gaa acc gtt gac ggc Asp Thr Ser Gly Lys Leu Ile Gln Leu Met Asn Glu Thr Val Asp Gly 310 315 320			6359
gac tac cag acc ttc aaa tcg aaa gat ggt gcg tac gtt cgt gaa cac Asp Tyr Gln Thr Phe Lys Ser Lys Asp Gly Ala Tyr Val Arg Glu His 325 330 335			6407
ttc ttc ggt aaa tat cct gaa acc gca gca ctg gtt gca gac tgg act Phe Phe Gly Lys Tyr Pro Glu Thr Ala Ala Leu Val Ala Asp Trp Thr 340 345 350			6455
gac gag cag atc tgg gca ctg aac cgt ggt ggt cac gat ccg aag aaa Asp Glu Gln Ile Trp Ala Leu Asn Arg Gly Gly His Asp Pro Lys Lys 355 360 365			6503
atc tac gct gca ttc aag aaa gcg cag gaa acc aaa ggc aaa gcg aca Ile Tyr Ala Ala Phe Lys Lys Ala Gln Glu Thr Lys Gly Lys Ala Thr 370 375 380 385			6551
gta atc ctt gct cat acc att aaa ggt tac ggc atg ggc gac gcg gct Val Ile Leu Ala His Thr Ile Lys Gly Tyr Gly Met Gly Asp Ala Ala 390 395 400			6599
gaa ggt aaa aac atc gcg cac cag gtt aag aaa atg aac atg gac ggt Glu Gly Lys Asn Ile Ala His Gln Val Lys Lys Met Asn Met Asp Gly 405 410 415			6647
gtg cgt cat atc cgc gac cgt ttc aat gtg ccg gtg tct gat gca gat Val Arg His Ile Arg Asp Arg Phe Asn Val Pro Val Ser Asp Ala Asp 420 425 430			6695
atc gaa aaa ctg ccg tac atc acc ttc ccg gaa ggt tct gaa gag cat Ile Glu Lys Leu Pro Tyr Ile Thr Phe Pro Glu Gly Ser Glu Glu His 435 440 445			6743
acc tat ctg cac gct cag cgt cag aaa ctg cac ggt tat ctg cca agc Thr Tyr Leu His Ala Gln Arg Gln Lys Leu His Gly Tyr Leu Pro Ser 450 455 460 465			6791
cgt cag ccg aac ttc acc gag aag ctt gag ctg ccg agc ctg caa gac Arg Gln Pro Asn Phe Thr Glu Lys Leu Glu Leu Pro Ser Leu Gln Asp 470 475 480			6839
ttc ggc gcg ctg ttg gaa gag cag agc aaa gag atc tct acc act atc Phe Gly Ala Leu Leu Glu Glu Gln Ser Lys Glu Ile Ser Thr Thr Ile 485 490 495			6887
gct ttc gtt cgt gct ctg aac gtg atg ctg aag aac aag tcg atc aaa Ala Phe Val Arg Ala Leu Asn Val Met Leu Lys Asn Lys Ser Ile Lys 500 505 510			6935
gat cgt ctg gta ccg atc atc gcc gac gaa gcg cgt act ttc ggt atg			6983

- 16 -

Asp	Arg	Leu	Val	Pro	Ile	Ile	Ala	Asp	Glu	Ala	Arg	Thr	Phe	Gly	Met	
515						520					525					
gaa	ggt	ctg	ttc	cgt	cag	att	ggt	att	tac	agc	ccg	aac	ggt	cag	cag	7031
Glu	Gly	Leu	Phe	Arg	Gln	Ile	Gly	Ile	Tyr	Ser	Pro	Asn	Gly	Gln	Gln	
530					535					540					545	
tac	acc	ccg	cag	gac	cgc	gag	cag	gtt	gct	tac	tat	aaa	gaa	gac	gag	7079
Tyr	Thr	Pro	Gln	Asp	Arg	Glu	Gln	Val	Ala	Tyr	Tyr	Lys	Glu	Asp	Glu	
				550					555					560		
aaa	ggt	cag	att	ctg	cag	gaa	ggg	atc	aac	gag	ctg	ggc	gca	ggt	tgt	7127
Lys	Gly	Gln	Ile	Leu	Gln	Glu	Gly	Ile	Asn	Glu	Leu	Gly	Ala	Gly	Cys	
			565					570					575			
tcc	tgg	ctg	gca	gcg	gcg	acc	tct	tac	agc	acc	aac	aat	ctg	ccg	atg	7175
Ser	Trp	Leu	Ala	Ala	Ala	Thr	Ser	Tyr	Ser	Thr	Asn	Asn	Leu	Pro	Met	
		580					585					590				
atc	ccg	ttc	tac	atc	tat	tac	tcg	atg	ttc	ggc	ttc	cag	cgt	att	ggc	7223
Ile	Pro	Phe	Tyr	Ile	Tyr	Tyr	Ser	Met	Phe	Gly	Phe	Gln	Arg	Ile	Gly	
	595					600						605				
gat	ctg	tgc	tgg	gcg	gct	ggc	gac	cag	caa	gcg	cgt	ggc	ttc	ctg	atc	7271
Asp	Leu	Cys	Trp	Ala	Ala	Gly	Asp	Gln	Gln	Ala	Arg	Gly	Phe	Leu	Ile	
610					615					620					625	
ggc	ggt	act	tcc	ggt	cgt	acc	acc	ctg	aac	ggc	gaa	ggt	ctg	cag	cac	7319
Gly	Gly	Thr	Ser	Gly	Arg	Thr	Thr	Leu	Asn	Gly	Glu	Gly	Leu	Gln	His	
				630					635					640		
gaa	gat	ggt	cac	agc	cac	att	cag	tcg	ctg	act	atc	ccg	aac	tgt	atc	7367
Glu	Asp	Gly	His	Ser	His	Ile	Gln	Ser	Leu	Thr	Ile	Pro	Asn	Cys	Ile	
			645					650					655			
tct	tac	gac	ccg	gct	tac	gct	tac	gaa	gtt	gct	gtc	atc	atg	cat	gac	7415
Ser	Tyr	Asp	Pro	Ala	Tyr	Ala	Tyr	Glu	Val	Ala	Val	Ile	Met	His	Asp	
		660					665					670				
ggt	ctg	gag	cgt	atg	tac	ggt	gaa	aaa	caa	gag	aac	gtt	tac	tac	tac	7463
Gly	Leu	Glu	Arg	Met	Tyr	Gly	Glu	Lys	Gln	Glu	Asn	Val	Tyr	Tyr	Tyr	
	675					680					685					
atc	act	acg	ctg	aac	gaa	aac	tac	cac	atg	ccg	gca	atg	ccg	gaa	ggt	7511
Ile	Thr	Thr	Leu	Asn	Glu	Asn	Tyr	His	Met	Pro	Ala	Met	Pro	Glu	Gly	
690					695					700					705	
gct	gag	gaa	ggt	atc	cgt	aaa	ggt	atc	tac	aaa	ctc	gaa	act	att	gaa	7559
Ala	Glu	Glu	Gly	Ile	Arg	Lys	Gly	Ile	Tyr	Lys	Leu	Glu	Thr	Ile	Glu	
				710					715					720		
ggt	agc	aaa	ggt	aaa	gtt	cag	ctg	ctc	ggc	tcc	ggt	tct	atc	ctg	cgt	7607
Gly	Ser	Lys	Gly	Lys	Val	Gln	Leu	Leu	Gly	Ser	Gly	Ser	Ile	Leu	Arg	
			725					730					735			
cac	gtc	cgt	gaa	gca	gct	gag	atc	ctg	gcg	aaa	gat	tac	ggc	gta	ggt	7655
His	Val	Arg	Glu	Ala	Ala	Glu	Ile	Leu	Ala	Lys	Asp	Tyr	Gly	Val	Gly	
		740					745					750				

- 17 -

tct gac gtt tat agc gtg acc tcc ttc acc gag ctg gcg cgt gat ggt	7703
Ser Asp Val Tyr Ser Val Thr Ser Phe Thr Glu Leu Ala Arg Asp Gly	
755 760 765	
cag gat tgt gaa cgc tgg aac atg ctg cac ccg ctg gaa act ccg cgc	7751
Gln Asp Cys Glu Arg Trp Asn Met Leu His Pro Leu Glu Thr Pro Arg	
770 775 780 785	
gtt ccg tat atc gct cag gtg atg aac gac gct ccg gca gtg gca tct	7799
Val Pro Tyr Ile Ala Gln Val Met Asn Asp Ala Pro Ala Val Ala Ser	
790 795 800	
acc gac tat atg aaa ctg ttc gct gag cag gtc cgt act tac gta ccg	7847
Thr Asp Tyr Met Lys Leu Phe Ala Glu Gln Val Arg Thr Tyr Val Pro	
805 810 815	
gct gac gac tac cgc gta ctg ggt act gat ggc ttc ggt cgt tcc gac	7895
Ala Asp Asp Tyr Arg Val Leu Gly Thr Asp Gly Phe Gly Arg Ser Asp	
820 825 830	
agc cgt gag aac ctg cgt cac cac ttc gaa gtt gat gct tct tat gtc	7943
Ser Arg Glu Asn Leu Arg His His Phe Glu Val Asp Ala Ser Tyr Val	
835 840 845	
gtg gtt gcg gcg ctg ggc gaa ctg gct aaa cgt ggc gaa atc gat aag	7991
Val Val Ala Ala Leu Gly Glu Leu Ala Lys Arg Gly Glu Ile Asp Lys	
850 855 860 865	
aaa gtg gtt gct gac gca atc gcc aaa ttc aac atc gat gca gat aaa	8039
Lys Val Val Ala Asp Ala Ile Ala Lys Phe Asn Ile Asp Ala Asp Lys	
870 875 880	
gtt aac ccg cgt ctg gcg taa gaggtaaaag aata atg gct atc gaa atc	8089
Val Asn Pro Arg Leu Ala Met Ala Ile Glu Ile	
885 890	
aaa gta ccg gac atc ggg gct gat gaa gtt gaa atc acc gag atc ctg	8137
Lys Val Pro Asp Ile Gly Ala Asp Glu Val Glu Ile Thr Glu Ile Leu	
895 900 905	
gtc aaa gtg ggc gac aaa gtt gaa gcc gaa cag tcg ctg atc acc gta	8185
Val Lys Val Gly Asp Lys Val Glu Ala Glu Gln Ser Leu Ile Thr Val	
910 915 920 925	
gaa ggc gac aaa gcc tct atg gaa gtt ccg tct ccg cag gcg ggt atc	8233
Glu Gly Asp Lys Ala Ser Met Glu Val Pro Ser Pro Gln Ala Gly Ile	
930 935 940	
gtt aaa gag atc aaa gtc tct gtt ggc gat aaa acc cag acc ggc gca	8281
Val Lys Glu Ile Lys Val Ser Val Gly Asp Lys Thr Gln Thr Gly Ala	
945 950 955	
ctg att atg att ttc gat tcc gcc gac ggt gca gca gac gct gca cct	8329
Leu Ile Met Ile Phe Asp Ser Ala Asp Gly Ala Ala Asp Ala Ala Pro	
960 965 970	
gct cag gca gaa gag aag aaa gaa gca gct ccg gca gca gca cca gcg	8377
Ala Gln Ala Glu Glu Lys Lys Glu Ala Ala Pro Ala Ala Ala Pro Ala	
975 980 985	

- 18 -

gct gcg gcg gca aaa gac gtt aac gtt ccg gat atc ggc agc gac gaa	8425
Ala Ala Ala Ala Lys Asp Val Asn Val Pro Asp Ile Gly Ser Asp Glu	
990 995 1000 1005	
gtt gaa gtg acc gaa atc ctg gtg aaa gtt ggc gat aaa gtt gaa gct	8473
Val Glu Val Thr Glu Ile Leu Val Lys Val Gly Asp Lys Val Glu Ala	
1010 1015 1020	
gaa cag tcg ctg atc acc gta gaa ggc gac aag gct tct atg gaa gtt	8521
Glu Gln Ser Leu Ile Thr Val Glu Gly Asp Lys Ala Ser Met Glu Val	
1025 1030 1035	
ccg gct ccg ttt gct ggc acc gtg aaa gag atc aaa gtg aac gtg ggt	8569
Pro Ala Pro Phe Ala Gly Thr Val Lys Glu Ile Lys Val Asn Val Gly	
1040 1045 1050	
gac aaa gtg tct acc ggc tcg ctg att atg gtc ttc gaa gtc gcg ggt	8617
Asp Lys Val Ser Thr Gly Ser Leu Ile Met Val Phe Glu Val Ala Gly	
1055 1060 1065	
gaa gca ggc gcg gca gct ccg gcc gct aaa cag gaa gca gct ccg gca	8665
Glu Ala Gly Ala Ala Ala Pro Ala Ala Lys Gln Glu Ala Ala Pro Ala	
1070 1075 1080 1085	
gcg gcc cct gca cca gcg gct ggc gtg aaa gaa gtt aac gtt ccg gat	8713
Ala Ala Pro Ala Pro Ala Ala Gly Val Lys Glu Val Asn Val Pro Asp	
1090 1095 1100	
atc ggc ggt gac gaa gtt gaa gtg act gaa gtg atg gtg aaa gtg ggc	8761
Ile Gly Gly Asp Glu Val Glu Val Thr Glu Val Met Val Lys Val Gly	
1105 1110 1115	
gac aaa gtt gcc gct gaa cag tca ctg atc acc gta gaa ggc gac aaa	8809
Asp Lys Val Ala Ala Glu Gln Ser Leu Ile Thr Val Glu Gly Asp Lys	
1120 1125 1130	
gct tct atg gaa gtt ccg gcg ccg ttt gca ggc gtc gtg aag gaa ctg	8857
Ala Ser Met Glu Val Pro Ala Pro Phe Ala Gly Val Val Lys Glu Leu	
1135 1140 1145	
aaa gtc aac gtt ggc gat aaa gtg aaa act ggc tcg ctg att atg atc	8905
Lys Val Asn Val Gly Asp Lys Val Lys Thr Gly Ser Leu Ile Met Ile	
1150 1155 1160 1165	
ttc gaa gtt gaa ggc gca gcg cct gcg gca gct cct gcg aaa cag gaa	8953
Phe Glu Val Glu Gly Ala Ala Pro Ala Ala Pro Ala Lys Gln Glu	
1170 1175 1180	
gcg gca gcg ccg gca ccg gca gca aaa gct gaa gcc ccg gca gca gca	9001
Ala Ala Ala Pro Ala Pro Ala Ala Lys Ala Glu Ala Pro Ala Ala Ala	
1185 1190 1195	
cca gct gcg aaa gcg gaa ggc aaa tct gaa ttt gct gaa aac gac gct	9049
Pro Ala Ala Lys Ala Glu Gly Lys Ser Glu Phe Ala Glu Asn Asp Ala	
1200 1205 1210	
tat gtt cac gcg act ccg ctg atc cgc cgt ctg gca cgc gag ttt ggt	9097
Tyr Val His Ala Thr Pro Leu Ile Arg Arg Leu Ala Arg Glu Phe Gly	

- 19 -

1215	1220	1225	
gtt aac ctt gcg aaa gtg aag ggc act ggc cgt aaa ggt cgt atc ctg			9145
Val Asn Leu Ala Lys Val Lys Gly Thr Gly Arg Lys Gly Arg Ile Leu			
1230	1235	1240	1245
cgc gaa gac gtt cag gct tac gtg aaa gaa gct atc aaa cgt gca gaa			9193
Arg Glu Asp Val Gln Ala Tyr Val Lys Glu Ala Ile Lys Arg Ala Glu			
	1250	1255	1260
gca gct ccg gca gcg act ggc ggt ggt atc cct ggc atg ctg ccg tgg			9241
Ala Ala Pro Ala Ala Thr Gly Gly Gly Ile Pro Gly Met Leu Pro Trp			
	1265	1270	1275
ccg aag gtg gac ttc agc aag ttt ggt gaa atc gaa gaa gtg gaa ctg			9289
Pro Lys Val Asp Phe Ser Lys Phe Gly Glu Ile Glu Glu Val Glu Leu			
	1280	1285	1290
ggc cgc atc cag aaa atc tct ggt gcg aac ctg agc cgt aac tgg gta			9337
Gly Arg Ile Gln Lys Ile Ser Gly Ala Asn Leu Ser Arg Asn Trp Val			
	1295	1300	1305
atg atc ccg cat gtt act cac ttc gac aaa acc gat atc acc gag ttg			9385
Met Ile Pro His Val Thr His Phe Asp Lys Thr Asp Ile Thr Glu Leu			
	1310	1315	1320
gaa gcg ttc cgt aaa cag cag aac gaa gaa gcg gcg aaa cgt aag ctg			9433
Glu Ala Phe Arg Lys Gln Gln Asn Glu Glu Ala Ala Lys Arg Lys Leu			
	1330	1335	1340
gat gtg aag atc acc ccg gtt gtc ttc atc atg aaa gcc gtt gct gca			9481
Asp Val Lys Ile Thr Pro Val Val Phe Ile Met Lys Ala Val Ala Ala			
	1345	1350	1355
gct ctt gag cag atg cct cgc ttc aat agt tcg ctg tcg gaa gac ggt			9529
Ala Leu Glu Gln Met Pro Arg Phe Asn Ser Ser Leu Ser Glu Asp Gly			
	1360	1365	1370
cag cgt ctg acc ctg aag aaa tac atc aac atc ggt gtg gcg gtg gat			9577
Gln Arg Leu Thr Leu Lys Lys Tyr Ile Asn Ile Gly Val Ala Val Asp			
	1375	1380	1385
acc ccg aac ggt ctg gtt gtt ccg gta ttc aaa gac gtc aac aag aaa			9625
Thr Pro Asn Gly Leu Val Val Pro Val Phe Lys Asp Val Asn Lys Lys			
	1390	1395	1400
ggc atc atc gag ctg tct cgc gag ctg atg act att tct aag aaa gcg			9673
Gly Ile Ile Glu Leu Ser Arg Glu Leu Met Thr Ile Ser Lys Lys Ala			
	1410	1415	1420
cgt gac ggt aag ctg act gcg ggc gaa atg cag ggc ggt tgc ttc acc			9721
Arg Asp Gly Lys Leu Thr Ala Gly Glu Met Gln Gly Gly Cys Phe Thr			
	1425	1430	1435
atc tcc agc atc ggc ggc ctg ggt act acc cac ttc gcg ccg att gtg			9769
Ile Ser Ser Ile Gly Gly Leu Gly Thr Thr His Phe Ala Pro Ile Val			
	1440	1445	1450
aac gcg ccg gaa gtg gct atc ctc ggc gtt tcc aag tcc gcg atg gag			9817

- 20 -

Asn Ala Pro Glu Val Ala Ile Leu Gly Val Ser Lys Ser Ala Met Glu
 1455 1460 1465
 ccg gtg tgg aat ggt aaa gag ttc gtg ccg cgt ctg atg ctg ccg att 9865
 Pro Val Trp Asn Gly Lys Glu Phe Val Pro Arg Leu Met Leu Pro Ile
 1470 1475 1480 1485
 tct ctc tcc ttc gac cac cgc gtg atc gac ggt gct gat ggt gcc cgt 9913
 Ser Leu Ser Phe Asp His Arg Val Ile Asp Gly Ala Asp Gly Ala Arg
 1490 1495 1500
 ttc att acc atc att aac aac acg ctg tct gac att cgc cgt ctg gtg 9961
 Phe Ile Thr Ile Ile Asn Asn Thr Leu Ser Asp Ile Arg Arg Leu Val
 1505 1510 1515
 atg taa gtaaaagagc cggcccaacg gccggctttt ttctggtaat ctcatgaatg 10017
 Met
 tattgagggtt attagcgaat agacaaatcg gttgccgttt gttgttttaa aattgttaac 10077
 aattttgtaa aataccgacg gatagaacga cccggtgggtg gttagggtat tacttcacat 10137
 accctatgga tttctgggtg 10157

<210> 4
 <211> 887
 <212> PRT
 <213> Escherichia coli

<400> 4
 Met Ser Glu Arg Phe Pro Asn Asp Val Asp Pro Ile Glu Thr Arg Asp
 1 5 10 15
 Trp Leu Gln Ala Ile Glu Ser Val Ile Arg Glu Glu Gly Val Glu Arg
 20 25 30
 Ala Gln Tyr Leu Ile Asp Gln Leu Leu Ala Glu Ala Arg Lys Gly Gly
 35 40 45
 Val Asn Val Ala Ala Gly Thr Gly Ile Ser Asn Tyr Ile Asn Thr Ile
 50 55 60
 Pro Val Glu Glu Gln Pro Glu Tyr Pro Gly Asn Leu Glu Leu Glu Arg
 65 70 75 80
 Arg Ile Arg Ser Ala Ile Arg Trp Asn Ala Ile Met Thr Val Leu Arg
 85 90 95
 Ala Ser Lys Lys Asp Leu Glu Leu Gly Gly His Met Ala Ser Phe Gln
 100 105 110
 Ser Ser Ala Thr Ile Tyr Asp Val Cys Phe Asn His Phe Phe Arg Ala
 115 120 125
 Arg Asn Glu Gln Asp Gly Gly Asp Leu Val Tyr Phe Gln Gly His Ile
 130 135 140
 Ser Pro Gly Val Tyr Ala Arg Ala Phe Leu Glu Gly Arg Leu Thr Gln

- 21 -

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

- 22 -

Ser	Arg	Gln	Pro	Asn	Phe	Thr	Glu	Lys	Leu	Glu	Leu	Pro	Ser	Leu	Gln
465					470					475					480
Asp	Phe	Gly	Ala	Leu	Leu	Glu	Glu	Gln	Ser	Lys	Glu	Ile	Ser	Thr	Thr
				485					490					495	
Ile	Ala	Phe	Val	Arg	Ala	Leu	Asn	Val	Met	Leu	Lys	Asn	Lys	Ser	Ile
			500					505					510		
Lys	Asp	Arg	Leu	Val	Pro	Ile	Ile	Ala	Asp	Glu	Ala	Arg	Thr	Phe	Gly
		515					520					525			
Met	Glu	Gly	Leu	Phe	Arg	Gln	Ile	Gly	Ile	Tyr	Ser	Pro	Asn	Gly	Gln
	530					535					540				
Gln	Tyr	Thr	Pro	Gln	Asp	Arg	Glu	Gln	Val	Ala	Tyr	Tyr	Lys	Glu	Asp
545					550					555					560
Glu	Lys	Gly	Gln	Ile	Leu	Gln	Glu	Gly	Ile	Asn	Glu	Leu	Gly	Ala	Gly
				565					570					575	
Cys	Ser	Trp	Leu	Ala	Ala	Ala	Thr	Ser	Tyr	Ser	Thr	Asn	Asn	Leu	Pro
			580					585					590		
Met	Ile	Pro	Phe	Tyr	Ile	Tyr	Tyr	Ser	Met	Phe	Gly	Phe	Gln	Arg	Ile
		595					600					605			
Gly	Asp	Leu	Cys	Trp	Ala	Ala	Gly	Asp	Gln	Gln	Ala	Arg	Gly	Phe	Leu
	610					615					620				
Ile	Gly	Gly	Thr	Ser	Gly	Arg	Thr	Thr	Leu	Asn	Gly	Glu	Gly	Leu	Gln
625					630					635					640
His	Glu	Asp	Gly	His	Ser	His	Ile	Gln	Ser	Leu	Thr	Ile	Pro	Asn	Cys
				645					650					655	
Ile	Ser	Tyr	Asp	Pro	Ala	Tyr	Ala	Tyr	Glu	Val	Ala	Val	Ile	Met	His
			660					665					670		
Asp	Gly	Leu	Glu	Arg	Met	Tyr	Gly	Glu	Lys	Gln	Glu	Asn	Val	Tyr	Tyr
		675					680					685			
Tyr	Ile	Thr	Thr	Leu	Asn	Glu	Asn	Tyr	His	Met	Pro	Ala	Met	Pro	Glu
	690					695					700				
Gly	Ala	Glu	Glu	Gly	Ile	Arg	Lys	Gly	Ile	Tyr	Lys	Leu	Glu	Thr	Ile
705					710					715					720
Glu	Gly	Ser	Lys	Gly	Lys	Val	Gln	Leu	Leu	Gly	Ser	Gly	Ser	Ile	Leu
				725					730					735	
Arg	His	Val	Arg	Glu	Ala	Ala	Glu	Ile	Leu	Ala	Lys	Asp	Tyr	Gly	Val
			740					745					750		
Gly	Ser	Asp	Val	Tyr	Ser	Val	Thr	Ser	Phe	Thr	Glu	Leu	Ala	Arg	Asp
		755					760					765			
Gly	Gln	Asp	Cys	Glu	Arg	Trp	Asn	Met	Leu	His	Pro	Leu	Glu	Thr	Pro
	770					775					780				

- 23 -

Arg Val Pro Tyr Ile Ala Gln Val Met Asn Asp Ala Pro Ala Val Ala
 785 790 795 800
 Ser Thr Asp Tyr Met Lys Leu Phe Ala Glu Gln Val Arg Thr Tyr Val
 805 810 815
 Pro Ala Asp Asp Tyr Arg Val Leu Gly Thr Asp Gly Phe Gly Arg Ser
 820 825 830
 Asp Ser Arg Glu Asn Leu Arg His His Phe Glu Val Asp Ala Ser Tyr
 835 840 845
 Val Val Val Ala Ala Leu Gly Glu Leu Ala Lys Arg Gly Glu Ile Asp
 850 855 860
 Lys Lys Val Val Ala Asp Ala Ile Ala Lys Phe Asn Ile Asp Ala Asp
 865 870 875 880
 Lys Val Asn Pro Arg Leu Ala
 885

<210> 5

<211> 630

<212> PRT

<213> Escherichia coli

<400> 5

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

- 24 -

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

- 25 -

Ser Leu Ser Glu Asp Gly Gln Arg Leu Thr Leu Lys Lys Tyr Ile Asn
 480 485 490 495
 Ile Gly Val Ala Val Asp Thr Pro Asn Gly Leu Val Val Pro Val Phe
 500 505 510
 Lys Asp Val Asn Lys Lys Gly Ile Ile Glu Leu Ser Arg Glu Leu Met
 515 520 525
 Thr Ile Ser Lys Lys Ala Arg Asp Gly Lys Leu Thr Ala Gly Glu Met
 530 535 540
 Gln Gly Gly Cys Phe Thr Ile Ser Ser Ile Gly Gly Leu Gly Thr Thr
 545 550 555
 His Phe Ala Pro Ile Val Asn Ala Pro Glu Val Ala Ile Leu Gly Val
 560 565 570 575
 Ser Lys Ser Ala Met Glu Pro Val Trp Asn Gly Lys Glu Phe Val Pro
 580 585 590
 Arg Leu Met Leu Pro Ile Ser Leu Ser Phe Asp His Arg Val Ile Asp
 595 600 605
 Gly Ala Asp Gly Ala Arg Phe Ile Thr Ile Ile Asn Asn Thr Leu Ser
 610 615 620
 Asp Ile Arg Arg Leu Val Met
 625 630

<210> 6

<211> 10863

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (9596)..(10798)

<400> 6

gactgtcggc gcagtaagct tcgcgaaata gatttaggaa taagcgaaga aaaatccctt 60
 tgccgacagg cgcaaattaa atctcgtcag gtgtacgcag gttagttacg gcaatgcgtt 120
 caccgcgctt gcgttcgcgc tctgattgca tattggcgcg ataaacgccc tgatcgccaa 180
 ccacccagga gagcggacga cgttctttgc cgatagattc ctgcaacagc atcagtgcct 240
 gcatgtacgc ttcaggacgc ggcgggcagc ccgggatata cacatcaacc gggatgaatt 300
 tatcgacgcc ctgcacaacg gaataaatat cgtacatacc accagagttg gcacaggcac 360
 ccattgagat aaccattttt ggttccagca tctggtcata cagacgctga ataaccggtg 420
 ccatttttgt aaagcaggtt cctgcaacca ccatcaggtc agcctgacgc ggcgaagcac 480
 gcaatacttc tgcgcaaaa cgcgccacgt catgcaccgc ggtaaacgaa gtcaccatct 540

- 26 -

caacgtaaca gcaggaaaga ccgaagttat acggccaaat tgagttttta cgaccccagt 600
taaccatgtc attgagcttg cccataaaca cgtttttgtt aacttcttgc tccagaggggt 660
cggttacgat ctctgctttt tgcaggggggt aacggtcggt ctaccggtg ggatctatgc 720
gggtgagcgt ataatccatc ttaatgcctc gcggttagcg ttgacgatta gcgatactgt 780
tcgtttccgg gttcatacgc tcgcggcgtg aacgcgcggg cgtccagtcc agcgcgccaa 840
tacgcaccag ataaaccaga cctgccagta acacaaaaat aaaaattgca gcttccacaa 900
agcctacca gcgcctttcg cggatagagg ttgacatgc gaacagatac agcgcttcaa 960
cgtcgaagat aacgaagaac atggccacca gataaaactt ggcgacagc cgtaagcggg 1020
cggagccgac cgagtogata ccggtattca acggcacgtt tttcgacctc gcgcgtgcgc 1080
gaccgcctaa aaaccaaccg cctaccagca tcaggcaaca caggccaatg gcaacgataa 1140
gaaagatagc gaatgcccag tgatgagcga tgacttcagt ggatgttgac atactcattg 1200
cttactcatc aaaagtagcg ccagattcac tgctcttcac ggcagatgga cgccacatcg 1260
attcatgggg aggaataaaa aaaaccttac aatcactgta gaaattcttt tatacagcta 1320
attgatgtgg ttttttactc ctttctataa ctttttgtca actttaacaa aagtttcttc 1380
acattagttt acataatatc aacaccatta gcatttaatg ccctttcacc ccagatcctt 1440
gacgactcca ggataattag atgttggtga atcgtgtccg ttgtgaagca atggaaaaaa 1500
tacgggtcta ttttgacagg aatttggtgc gattcctccc cccaaaagag agtatcttct 1560
tgatctgtga cacgcttttg tcattccata acaaaaacgc agcaacaaat ttacgtattt 1620
tttaacatca ttgtagcagg tgattttttt caggcgatta tttgtgcgtt cgggacgtga 1680
atctctggtg gttgaaaaat gaacagtttt gtacgttctg cactatgcga tgaaggattt 1740
ttactaaaaa aaagccgctg gggtttaaaa cccccccagc ggctcgtttt ttacactatt 1800
gtctcaggaa ttatctatcg tccgtcgatt actcgatatc cttttcaatc aacaatgaat 1860
catccccttc cggagcagac attgggctgt attgccacgg attatggtag ctctccatcg 1920
cctgatagat cacctgggcc agctcattat tactggacgg atcatagcac agcaaattatt 1980
cggtatcagg tagcggcgggt aagccatcta ctccactcag cacacgcagg tccgggctca 2040
tcattctaac cggccttgcc gtcacgccaa gaccggcttt cactgccgca cgaacggccg 2100
gaagcgtcga ggcgacataa gccagtcgcc atggaatatc tgctttatta agcgtcgcca 2160
gcaccatatc gcgaaacggg ctaggatcat ccagcaatac aagagggatc ggctcgcctt 2220
tttgcaatat gtattccgct gcgcagtacc agtgtgttgg cgaggtagcg aggttcaatg 2280

- 27 -

ccttaaacgc tgaaggacga tgggtgggta ccatcaaato cacttcctgg gattccagca 2340
tttctgccat ataggcatta cgtttaacgc ggacatccag cgctaatttc ggataaacgc 2400
aactcacgcg atttaacagg aaaggtaaga tcgtatcggc agattcatct gaagcaccga 2460
tagttaatac gccctgaaga ttactgaaca ttaatgatga gcaggcctca tcattaaaac 2520
gcaggatttt cctggcgtaa ccaagaagtt gaatgccatg ttcagttaac agtttggtgc 2580
gaccgtgacg agcgaacagt tctttcccaa cgagttgttc cagacgctgc atttgctgac 2640
ttacggcgga ctgagtacga cacacagcgg cagctgcggc agcaaaagtg ttcagatcgg 2700
caacagcaac aaatgttctc agcagatcga ggtcgagggt aattatcgga cgatttgac 2760
ttatcatata ttatcactta ctggcggctc atactgagct ggttaatgct gtgcacacac 2820
aaacaagcaa ttccatttgt aatgtgcctc cctggcagtt tcatccggaa cccggacgaa 2880
agtaaaaatg catatgagtt gcactaaaaa agcgactcac attgttccgt tataatgcct 2940
gaagtagatc acagaatata tcttcaggga tcgcatatct attaagttac tcactctttt 3000
ctatttatga catgcgcgtg tttgtataaa tgtaaagtgt agtccttggt ccactctcgt 3060
gcagcatcgc tggtcatacg cgaacacgta ccaacagcaa tggtgaggcg gcatcaagag 3120
cagggatccg tcatttatcc gagcatttta ccccaaaaac cttttattta taagggtcat 3180
tgcgaattat ctgatgcaaa gttatgttat gttaggcaaa gtaatcttct attattaata 3240
agcacatcaa aacctttttg aatattaaat aataattaat tagcatcatc ctcatcatt 3300
aattccgctt aacaatagtt tcacaccttg ccccatctgc gacacctcgg aatcatcac 3360
gcagttaaac tctccgtaat gaggaaatat catcatcgcg tttcattagt gaattcttaa 3420
atgaggcatt ttcacacaat tatcttacag ataaaaaac cagacttaca attaagaatc 3480
agaacaattc actatataac attgcatgta aagcatatac acctcattat tttgtcatta 3540
ttaagttatt aacagcacia tcgagccttc ccctctggca aaatcttatt ctgcagacct 3600
tcaaaacacc gtcctggggg agtacattgt tctaagctga cttccacggc agggagtggc 3660
gataacagca aaaaagggtca agattcatgt ccccatctga aaaatccagc aaattagaga 3720
atgtctgtta tgacatccgt ggtccggtgc tgaaagaagc aaaacgcctg gaagaagaag 3780
gtaacaaggt actgaaactg aacatcggca acccagcccc gttcggtttt gacgcgccag 3840
atgaaatcct cgttgacgtg atacgcaacc tgcctaccgc tcaagggtat tgcgattcca 3900
aagggtctta ctccgcgcgt aaagccatca tgcagcacta ccaggctcgt ggcatgcgtg 3960
atgttaccgt ggaagatatt tacatcggca atgggtgtatc ggagcttata gttcaggcaa 4020
tgcaggcatt gctgaacagc ggggacgaaa tggtggttcc tgcaccagat taccactct 4080

- 28 -

ggaccgcggc ggtttcgctt tccagcggta aagcgggtgca ttatctttgc gatgaatcct 4140
ctgactgggt cccggacctc gatgatattc gcgctaaaat tacgcctcgt acgcgtggga 4200
tcgttattat caacccaa at aaccaaccg gcgcgggtata ttccaaagag cttttaatgg 4260
agattgtgga gattgcacgt cagcataatc tcattatctt cgccgatgaa atttatgaca 4320
aaattctcta cgacgacgct gagcatcact caattgcgcc gctggcacct gacctgctga 4380
ccattacctt taacggactg tcgaaaacgt accgcgttgc aggcttccgt caggggtgga 4440
tggtgttgaa cgggccgaaa aaacacgcca aaggctacat cgaaggctctg gaaatgctgg 4500
cttcaatgcg cctgtgtgct aacgttcctg cgcaacacgc cattcagacc gcgctagggtg 4560
gttatcagag catcagtga tttattacc ctggcgggtcg tctttatgag cagcgtaacc 4620
gcgcgtggga actgatcaac gatattccgg gcgtttcctg cgtgaaacct cgtgggtgcgc 4680
tgtatatgtt cccgaaaatc gacgccaac gctttaacat tcacgacgat cagaaaatgg 4740
tgttggattt cctgttgacg gaaaaagttc tgttgggtgca agggacggca ttcaactggc 4800
cgtggccgga tcacttccgc attgtcacgc taccgcgtgt cgatgatatc gagctgtctt 4860
tgagcaagtt cgcgcgtttc ctttctggtt atcatcagct gtaatcttaa tttcactgcc 4920
ggagattgca tccggcagcg ttatcccgcc acaatgacct gatgatgtca tcatacgtaa 4980
ggtcactatg aaacagagcc atttttttgc ccatctctcc cgcctgaaac tcattaaccg 5040
ctggccgctc atgcgcaacg tgcggacgga aaatgtgtcc gaacacagtt tgcaggtagc 5100
gatggtcgcc catgcgctgg cagctatcaa aaatcgaaaa tttggcggta atgtcaacgc 5160
cgaacgtatc gctttactgg cgatgtacca cgatgcctca gaagtgtca ccggcgatct 5220
ccctactccg gtgaaatact tcaattcgca aatcgctcag gagtacaagg ctattgaaaa 5280
aatcgctcag caaaaactgg tcgatatggt tccggaagag ctgcgggata tctttgcgcc 5340
gttaattgac gagcatgcat atagcgatga agaaaaatcg ctggtgaaac aggcagatgc 5400
actgtgtgca tatctgaaat gtctggaaga actcgcggcc ggaaataatg aattcttgct 5460
ggcaaaaacg cgactggaag cgacgcttga agcgcgctcg agccaggaga tggactactt 5520
catggaaata tttgttccca gcttccatct ttcgctcgat gagattagcc aggattcacc 5580
gctgtaagca gccggagtct gcgtcgcatc aggcaataag cgccggatgc gacatcaggc 5640
tcttgtcaaa acggaaacag catcgggatc atcaccacac aaaccgccat cacgataatg 5700
gtgaacggta cccccaactt cacaaagtca ctgaagctgt aatttcccg acctaaaacc 5760
agtgtgttaa cagggtgaaga aaccggtgtc ataaaggcgg cggatgctgc catcgccacg 5820

- 29 -

accatcgga atggataagg cgacaccccc atcgttttgg cagcagccag cgcaatcgga 5880
gccatcaaca ccgccgtcgc ggtattagag ataaatagcc caataaccgc cgacaagaca 5940
aacaacacc ccagcatcat atgtggcccc taaccaccgc caatgtccat cagccctttc 6000
accgccagcg cgacacctcc cgttttctgt aatgccacag caaacggcat catcccaacg 6060
atcaaaataa tgctcggcca gtgaatggat ttataggcgc tttcagcatc tatacagcgg 6120
aatttcccca tcagcaggca ggcgatgata gcggcaacag gattaggaat ttcattctgtc 6180
agcattaacg ccaccattag caccagacag aaaatggcat ggggtgcctg gctgtgcgcg 6240
ggtgatgctt cactcacctc ttccggtaag ttcagcgcga cgaagtcgcg gccctgtttg 6300
gccagcatac cgatcagttt ccagttaccc acaaccagga tgatatcgcc cagcagcaga 6360
ggctcatccg ccagcgaacc ttccagcgcg acgccattgc gcttttagccc caccacattc 6420
agtccgtagc gggtagcga accaatttcg cgcaccgatt taccaatcag ttctgactca 6480
ggaattaatg aaatctctgc catgcccaca tcaagggcct ggtcagaaaa atactcgccg 6540
cgcagtacca tcggctccag caattgctca ctacaaaatt gccggagatc gacatcagcc 6600
gcagacatat caataagcaa aacgtcacgc gcgcgaaatt cagaaacccc attaacgttc 6660
acgataacgc gacgaaaacg ccgccagcgt tcaacaccga tgacgttagc gccataacgc 6720
tcacgtaatt tgagatcatc cagccgttga ccaatcattg gcgatccggg gcgaatagcc 6780
agacgtcgcg cagccccgtt cagtcgatat tcacggataa gatcgcgaaa gggtcgcgcg 6840
gtccagcctt cgcgtcgcg ggtctgggta tcccccttca gcatgaaacg catcactaac 6900
atatacaaaa taccagcac cagcacgacc aggccaatag gtgttacgct aaagaaactg 6960
aagccgtgat agccttcacg cagcaattca ctgttgacta ccaggttcgg cggcgtcgcc 7020
accagcgtca tcatgccgct aatcagcccg gcaaaactta acggcatcat cagacgtgac 7080
ggcgacgttt gcatacgcg ggcaacgctt aacaccacgg gaataaagat agcgacaacg 7140
ccggttgaac tcataaacgc gccaaagccc gcgacgggtc gcatcaacaa aaccaacatt 7200
ttgatttcac tattgccgcg aactttgacc agccatgttc ccattacggg ggcaacaccg 7260
gtacggacca aaccatcgcc aataataaac aaggcggcaa tcaggacaac gttaggatca 7320
gaaaagccgg aaaatacttc tgggactgtc agcgttccgc ttaatgcaa cgcgacaata 7380
acaaacaaag cgaccgcac catacgcact ctgcccgtcg caaacaagac gatggcaacg 7440
gccagtaatg aaagaaccca aatcaattca ccgttcacaa cttatccttg ttaattgagg 7500
gggatgactt gattctgcca taaaaagcc ccgacgagac ggggctaaat catgatcagg 7560
tgtttactg aataataaca tcgccatttg gctgtttggg cacagtaatt tgctccagac 7620

tgtggaggac caaatcgacc tcattcaggc gcgggggtatc tgccggagcg ttaaccgcaa 7680
tgacatgaca acccgccgcc aggccagaaa gcacgccagc gggagcatct tccaccacca 7740
cacactcctg cggcgcaagc cccagcagct gcgcgcctaa cagatacgca tcaggttctg 7800
gttttccgcg cttcactcgc tcagcggtta caaacacctc tgggtgcggga agcccagcta 7860
ttttatggcg cgctcgcgct accggcatgg agccagaagt cacaatggcc cacggaatac 7920
ctgctttatt caaatgactg agtaaggcga ttgcccccg aagcgcggta ataccttcgg 7980
tttccgtggc ctcgatgtgc tccagacgcg taaactcggc ggcaatatca gcctcggatt 8040
tgcccgccat aaaatggcgc agagaggtga tcgcctgttt accgtgaatg aaagccagca 8100
cctcttcogg cgctaaccce tgacgtctgg ccagttgct ccacgcccgt tctaccgcag 8160
gcagggaatc caccagcgtt ccatcaagat caaacagaaa acctttgcac cgcacgcggg 8220
cctcctcagg cattgataat ttgattaatt tcgttggcgc tcaaattgga ctgacgcggg 8280
caggcatgcc acacattaag catgcgctga tatttttccc acattggcgt ctgggcgtta 8340
aaaccgtgag ttccggcgtc aaaatgggta tagcgccctt ccacattaac cataaagcgg 8400
acataaccga ggtaacgtgc ttcagtggcg gcgtcaaagc cgaggaaggt gacacgacgt 8460
tcategatgg attgctggtc ctgcaaatta gaccaggaaa catgcaacgc atgatacatc 8520
tccataatgt cgatgatagt gcggcagggt tcttctttca gctcgccaaa ctgcgcatcc 8580
aattcacgca tctgtaatcc gtaaccacgc tcaataattg tttgcaggcg acggtaacgt 8640
tcagcatttg ccgatcgag catagtcac atcttgtact gattagacaa aataagacgt 8700
tgcgcgttgg tcatttccat tgttgactcc tgtatcactc tactacggtg aaaaaaaga 8760
aggctgagta tgccttcttt tatatgcgta atcaggggtc aattacaaat catcaaggaa 8820
agttttatcc agttgtttga aggcgcgctt aagcgtgtca gctaattgcct ggtaatcagg 8880
cttgccttca acgggtgcc aacactgtcc agactcctgc aatttaccgc gaacttcata 8940
aaaccagtta aggattgcag ggggtaatgg cgtgacagaa cgcttgccca gccaccacaa 9000
tccctgcatg ggtaaaactta aggcgaacag ggcagtggca actgccggcc caagctgacc 9060
gccagggca atctgccagc agagagtaaa tacggcgatc ggcggcataa aacggatcgc 9120
ataacgcgtc atcttgataa cgcgattttc gacaaagacc ggggcaaggc gtttttccag 9180
cggccacgtc tttgagtaat gctgtccccg gcgaaacaag ctaaaaaaat taacagaacg 9240
attatccggc gttgacatgc ttcacctcaa cttcacatat aaagattcaa aaatttgtgc 9300
aaattcacia ctcagcggga caacgttcaa aacattttgt cttccatacc cactatcagg 9360

- 31 -

tatccttttag cagcctgaag gcctaagtag tacatattca ttgagtcgtc aaattcatat 9420

acattatgcc attggctgaa aattacgcaa aatggcatag actcaagata tttcttccat 9480

catgcaaaaa aaatttgcag tgcattgatgt taatcataaa tgtcgggtgtc atcatgctgt 9540

acgctctatg gctccctgac gtttttttttag ccacgtatca attataggta ctcc atg 9598
Met
1

tcg agt aag tta gta ctg gtt ctg aac tgc ggt agt tct tca ctg aaa 9646
Ser Ser Lys Leu Val Leu Val Leu Asn Cys Gly Ser Ser Ser Leu Lys
5 10 15

ttt gcc atc atc gat gca gta aat ggt gaa gag tac ctt tct ggt tta 9694
Phe Ala Ile Ile Asp Ala Val Asn Gly Glu Glu Tyr Leu Ser Gly Leu
20 25 30

gcc gaa tgt ttc cac ctg ccc gaa gca cgt atc aaa tgg aaa atg gac 9742
Ala Glu Cys Phe His Leu Pro Glu Ala Arg Ile Lys Trp Lys Met Asp
35 40 45

ggc aat aaa cag gaa gcg gct tta ggt gca ggc gcc gct cac agc gaa 9790
Gly Asn Lys Gln Glu Ala Ala Leu Gly Ala Gly Ala Ala His Ser Glu
50 55 60 65

gcg ctc aac ttt atc gtt aat act att ctg gca caa aaa cca gaa ctg 9838
Ala Leu Asn Phe Ile Val Asn Thr Ile Leu Ala Gln Lys Pro Glu Leu
70 75 80

tct gcg cag ctg act gct atc ggt cac cgt atc gta cac ggc ggc gaa 9886
Ser Ala Gln Leu Thr Ala Ile Gly His Arg Ile Val His Gly Gly Glu
85 90 95

aag tat acc agc tcc gta gtg atc gat gag tct gtt att cag ggt atc 9934
Lys Tyr Thr Ser Ser Val Val Ile Asp Glu Ser Val Ile Gln Gly Ile
100 105 110

aaa gat gca gct tct ttt gca ccg ctg cac aac ccg gct cac ctg atc 9982
Lys Asp Ala Ala Ser Phe Ala Pro Leu His Asn Pro Ala His Leu Ile
115 120 125

ggt atc gaa gaa gct ctg aaa tct ttc cca cag ctg aaa gac aaa aac 10030
Gly Ile Glu Glu Ala Leu Lys Ser Phe Pro Gln Leu Lys Asp Lys Asn
130 135 140 145

gtt gct gta ttt gac acc gcg ttc cac cag act atg ccg gaa gag tct 10078
Val Ala Val Phe Asp Thr Ala Phe His Gln Thr Met Pro Glu Glu Ser
150 155 160

tac ctc tac gcc ctg cct tac aac ctg tac aaa gag cac ggc atc cgt 10126
Tyr Leu Tyr Ala Leu Pro Tyr Asn Leu Tyr Lys Glu His Gly Ile Arg
165 170 175

cgt tac ggc gcg cac ggc acc agc cac ttc tat gta acc cag gaa gcg 10174
Arg Tyr Gly Ala His Gly Thr Ser His Phe Tyr Val Thr Gln Glu Ala
180 185 190

gca aaa atg ctg aac aaa ccg gta gaa gaa ctg aac atc atc acc tgc 10222

WO 01/70776

PCT/US01/07478

- 32 -

Ala	Lys	Met	Leu	Asn	Lys	Pro	Val	Glu	Glu	Leu	Asn	Ile	Ile	Thr	Cys	
195						200					205					
cac	ctg	ggc	aac	ggt	ggt	tcc	ggt	tct	gct	atc	cgc	aac	ggt	aaa	tgc	10270
His	Leu	Gly	Asn	Gly	Gly	Ser	Val	Ser	Ala	Ile	Arg	Asn	Gly	Lys	Cys	
210					215					220					225	
ggt	gac	acc	tct	atg	ggc	ctg	acc	cgc	ctg	gaa	ggt	ctg	gtc	atg	ggt	10318
Val	Asp	Thr	Ser	Met	Gly	Leu	Thr	Pro	Leu	Glu	Gly	Leu	Val	Met	Gly	
				230					235					240		
acc	cgt	tct	ggt	gat	atc	gat	cgc	gcg	atc	atc	ttc	cac	ctg	cac	gac	10366
Thr	Arg	Ser	Gly	Asp	Ile	Asp	Pro	Ala	Ile	Ile	Phe	His	Leu	His	Asp	
			245					250					255			
acc	ctg	ggc	atg	agc	ggt	gac	gca	atc	aac	aaa	ctg	ctg	acc	aaa	gag	10414
Thr	Leu	Gly	Met	Ser	Val	Asp	Ala	Ile	Asn	Lys	Leu	Leu	Thr	Lys	Glu	
		260					265					270				
tct	ggc	ctg	ctg	ggt	ctg	acc	gaa	gtg	acc	agc	gac	tgc	cgc	tat	ggt	10462
Ser	Gly	Leu	Leu	Gly	Leu	Thr	Glu	Val	Thr	Ser	Asp	Cys	Arg	Tyr	Val	
	275					280					285					
gaa	gac	aac	tac	gcg	acg	aaa	gaa	gac	gcg	aag	cgc	gca	atg	gac	ggt	10510
Glu	Asp	Asn	Tyr	Ala	Thr	Lys	Glu	Asp	Ala	Lys	Arg	Ala	Met	Asp	Val	
290					295					300					305	
tac	tgc	cac	cgc	ctg	gcg	aaa	tac	atc	ggt	gcc	tac	act	gcg	ctg	atg	10558
Tyr	Cys	His	Arg	Leu	Ala	Lys	Tyr	Ile	Gly	Ala	Tyr	Thr	Ala	Leu	Met	
				310					315					320		
gat	ggt	cgt	ctg	gac	gct	ggt	gta	ttc	act	ggt	ggt	atc	ggt	gaa	aat	10606
Asp	Gly	Arg	Leu	Asp	Ala	Val	Val	Phe	Thr	Gly	Gly	Ile	Gly	Glu	Asn	
			325					330					335			
gcc	gca	atg	ggt	cgt	gaa	ctg	tct	ctg	ggc	aaa	ctg	ggc	gtg	ctg	ggc	10654
Ala	Ala	Met	Val	Arg	Glu	Leu	Ser	Leu	Gly	Lys	Leu	Gly	Val	Leu	Gly	
		340					345					350				
ttt	gaa	ggt	gat	cat	gaa	cgc	aac	ctg	gct	gca	cgt	ttc	ggc	aaa	tct	10702
Phe	Glu	Val	Asp	His	Glu	Arg	Asn	Leu	Ala	Ala	Arg	Phe	Gly	Lys	Ser	
	355					360					365					
ggt	ttc	atc	aac	aaa	gaa	ggt	acc	cgt	cct	gcg	gtg	ggt	atc	cca	acc	10750
Gly	Phe	Ile	Asn	Lys	Glu	Gly	Thr	Arg	Pro	Ala	Val	Val	Ile	Pro	Thr	
370					375					380					385	
aac	gaa	gaa	ctg	ggt	atc	gcg	caa	gac	gcg	agc	cgc	ctg	act	gcc	tga	10798
Asn	Glu	Glu	Leu	Val	Ile	Ala	Gln	Asp	Ala	Ser	Arg	Leu	Thr	Ala		
				390					395					400		
tttcacaccg	ccagctcagc	tggcggtgct	gttttgtaac	ccgccaaatc	ggcggtaacg											10858
aaaga																10863

<210> 7
 <211> 400
 <212> PRT

- 33 -

<213> Escherichia coli

<400> 7

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

- 34 -

Val Tyr Cys His Arg Leu Ala Lys Tyr Ile Gly Ala Tyr Thr Ala Leu
 305 310 315 320

Met Asp Gly Arg Leu Asp Ala Val Val Phe Thr Gly Gly Ile Gly Glu
 325 330 335

Asn Ala Ala Met Val Arg Glu Leu Ser Leu Gly Lys Leu Gly Val Leu
 340 345 350

Gly Phe Glu Val Asp His Glu Arg Asn Leu Ala Ala Arg Phe Gly Lys
 355 360 365

Ser Gly Phe Ile Asn Lys Glu Gly Thr Arg Pro Ala Val Val Ile Pro
 370 375 380

Thr Asn Glu Glu Leu Val Ile Ala Gln Asp Ala Ser Arg Leu Thr Ala
 385 390 395 400

<210> 8
 <211> 11630
 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 <222> (1163)..(2602)

<400> 8
 aattctctgc tgcaagatga ataatgttta tctacagcat ttccttaaaa gatatgtcag 60
 gcttgccgag tggcgggttaa ggacatacga tttcctcctt tcagagtgc cgccttctca 120
 ctattatctc acgcagtatt ctttaagggaa cgataaggag gaaccatgaa cattaccccg 180
 tttccgacgc tttcgccggc aactatagat gccataaatg ttatcggaca gtggctggcg 240
 caggatgatt tctccggtga ggtgccgtat caggccgatt gcgtgatcct tgcaggcaat 300
 gcggttatgc cgactatcga tgcggcatgt aagattgccc gcgatcagca aattccttta 360
 ctgattagtg gtggtatcgg tcaactcgaca acttttttgt atagcgccat cgcacagcat 420
 ccgcactaca acactatccg caccactggc agagcagaag cgaccatcct ggcggatatc 480
 gctcatcagt tctggcacat tccgcatgaa aaaatctgga ttgaagacca gtcaacaaac 540
 tgcggtgaaa acgcacgctt tagcatcgcg ctattgaatc aggccgtaga acgagttcat 600
 acggctatcg ttgttcagga cccacccatg cagcggcgca cgatggcgac gttccgccgt 660
 atgactgggg acaatcccga tgcaccacgc tggttaagtt atcccggatt cgttcctcag 720
 ttaggaaata acgcagacag tgtaatcttt attaatacgt tacaaggatt atggccagtt 780
 gagcggtatc tctcactact cactggcgag ctgccgcgtt tacgcgatga tagcgatggc 840
 tacgggtccc gcgggcgaga ttttatcggt caggttgatt ttccggcaga agtcatccat 900

gcatggcaaa cgctgaaaca tgatgcggtg ctcatcgagg cgatggaaag tcgctcgtta															960
cgttaaaaat tgcccgtttg tgaaccactt gtttgcaaac gggcatgact cctgactttt															1020
atttctgcct tttattcctt ttacacttgt ttttatgaag cccttcacag aattgtcctt															1080
tcacgattcc gtctctctga tgattgatgt taattaacaa tgtattcacc gaaaacaaac															1140
atataaatca caggagtcgc cc atg tca gta ccc gtt caa cat cct atg tat															1192
Met Ser Val Pro Val Gln His Pro Met Tyr															
1 5 10															
atc gat gga cag ttt gtt acc tgg cgt gga gac gca tgg att gat gtg															1240
Ile Asp Gly Gln Phe Val Thr Trp Arg Gly Asp Ala Trp Ile Asp Val															
15 20 25															
gta aac cct gct aca gag gct gtc att tcc cgc ata ccc gat ggt cag															1288
Val Asn Pro Ala Thr Glu Ala Val Ile Ser Arg Ile Pro Asp Gly Gln															
30 35 40															
gcc gag gat gcc cgt aag gca atc gat gca gca gaa cgt gca caa cca															1336
Ala Glu Asp Ala Arg Lys Ala Ile Asp Ala Ala Glu Arg Ala Gln Pro															
45 50 55															
gaa tgg gaa gcg ttg cct gct att gaa cgc gcc agt tgg ttg cgc aaa															1384
Glu Trp Glu Ala Leu Pro Ala Ile Glu Arg Ala Ser Trp Leu Arg Lys															
60 65 70															
atc tcc gcc ggg atc cgc gaa cgc gcc agt gaa atc agt gcg ctg att															1432
Ile Ser Ala Gly Ile Arg Glu Arg Ala Ser Glu Ile Ser Ala Leu Ile															
75 80 85 90															
gtt gaa gaa ggg ggc aag atc cag cag ctg gct gaa gtc gaa gtg gct															1480
Val Glu Glu Gly Gly Lys Ile Gln Gln Leu Ala Glu Val Glu Val Ala															
95 100 105															
ttt act gcc gac tat atc gat tac atg gcg gag tgg gca cgg cgt tac															1528
Phe Thr Ala Asp Tyr Ile Asp Tyr Met Ala Glu Trp Ala Arg Arg Tyr															
110 115 120															
gag ggc gag att att caa agc gat cgt cca gga gaa aat att ctt ttg															1576
Glu Gly Glu Ile Ile Gln Ser Asp Arg Pro Gly Glu Asn Ile Leu Leu															
125 130 135															
ttt aaa cgt gcg ctt ggt gtg act acc ggc att ctg ccg tgg aac ttc															1624
Phe Lys Arg Ala Leu Gly Val Thr Thr Gly Ile Leu Pro Trp Asn Phe															
140 145 150															
ccg ttc ttc ctc att gcc cgc aaa atg gct ccc gct ctt ttg acc ggt															1672
Pro Phe Phe Leu Ile Ala Arg Lys Met Ala Pro Ala Leu Leu Thr Gly															
155 160 165 170															
aat acc atc gtc att aaa cct agt gaa ttt acg cca aac aat gcg att															1720
Asn Thr Ile Val Ile Lys Pro Ser Glu Phe Thr Pro Asn Asn Ala Ile															
175 180 185															
gca ttc gcc aaa atc gtc gat gaa ata ggc ctt ccg cgc ggc gtg ttt															1768
Ala Phe Ala Lys Ile Val Asp Glu Ile Gly Leu Pro Arg Gly Val Phe															

- 36 -

190						195						200						
aac	ctt	gta	ctg	ggg	cgt	ggt	gaa	acc	gtt	ggg	caa	gaa	ctg	gcg	ggt	1816		
Asn	Leu	Val	Leu	Gly	Arg	Gly	Glu	Thr	Val	Gly	Gln	Glu	Leu	Ala	Gly			
		205						210					215					
aac	cca	aag	gtc	gca	atg	gtc	agt	atg	aca	ggc	agc	gtc	tct	gca	ggt	1864		
Asn	Pro	Lys	Val	Ala	Met	Val	Ser	Met	Thr	Gly	Ser	Val	Ser	Ala	Gly			
		220						225					230					
gag	aag	atc	atg	gcg	act	gcg	gcg	aaa	aac	atc	acc	aaa	gtg	tgt	ctg	1912		
Glu	Lys	Ile	Met	Ala	Thr	Ala	Ala	Lys	Asn	Ile	Thr	Lys	Val	Cys	Leu			
					240					245					250			
gaa	ttg	ggg	ggt	aaa	gca	cca	gct	atc	gta	atg	gac	gat	gcc	gat	ctt	1960		
Glu	Leu	Gly	Gly	Lys	Ala	Pro	Ala	Ile	Val	Met	Asp	Asp	Ala	Asp	Leu			
				255					260					265				
gaa	ctg	gca	gtc	aaa	gcc	atc	gtt	gat	tca	cgc	gtc	att	aat	agt	ggg	2008		
Glu	Leu	Ala	Val	Lys	Ala	Ile	Val	Asp	Ser	Arg	Val	Ile	Asn	Ser	Gly			
			270					275					280					
caa	gtg	tgt	aac	tgt	gca	gaa	cgt	gtt	tat	gta	cag	aaa	ggc	att	tat	2056		
Gln	Val	Cys	Asn	Cys	Ala	Glu	Arg	Val	Tyr	Val	Gln	Lys	Gly	Ile	Tyr			
		285					290					295						
gat	cag	ttc	gtc	aat	cgg	ctg	ggt	gaa	gcg	atg	cag	gcg	gtt	caa	ttt	2104		
Asp	Gln	Phe	Val	Asn	Arg	Leu	Gly	Glu	Ala	Met	Gln	Ala	Val	Gln	Phe			
		300				305					310							
ggt	aac	ccc	gct	gaa	cgc	aac	gac	att	gcg	atg	ggg	ccg	ttg	att	aac	2152		
Gly	Asn	Pro	Ala	Glu	Arg	Asn	Asp	Ile	Ala	Met	Gly	Pro	Leu	Ile	Asn			
					320					325					330			
gcc	gcg	gcg	ctg	gaa	agg	gtc	gag	caa	aaa	gtg	gcg	cgc	gca	gta	gaa	2200		
Ala	Ala	Ala	Leu	Glu	Arg	Val	Glu	Gln	Lys	Val	Ala	Arg	Ala	Val	Glu			
				335					340					345				
gaa	ggg	gcg	aga	gtg	gcg	ttc	ggt	ggc	aaa	gcg	gta	gag	ggg	aaa	gga	2248		
Glu	Gly	Ala	Arg	Val	Ala	Phe	Gly	Gly	Lys	Ala	Val	Glu	Gly	Lys	Gly			
			350					355					360					
tat	tat	tat	ccg	ccg	aca	ttg	ctg	ctg	gat	gtt	cgc	cag	gaa	atg	tcg	2296		
Tyr	Tyr	Tyr	Pro	Pro	Thr	Leu	Leu	Leu	Asp	Val	Arg	Gln	Glu	Met	Ser			
			365				370					375						
att	atg	cat	gag	gaa	acc	ttt	ggc	ccg	gtg	ctg	cca	gtt	gtc	gca	ttt	2344		
Ile	Met	His	Glu	Glu	Thr	Phe	Gly	Pro	Val	Leu	Pro	Val	Val	Ala	Phe			
			380			385					390							
gac	acg	ctg	gaa	gat	gct	atc	tca	atg	gct	aat	gac	agt	gat	tac	ggc	2392		
Asp	Thr	Leu	Glu	Asp	Ala	Ile	Ser	Met	Ala	Asn	Asp	Ser	Asp	Tyr	Gly			
					400					405					410			
ctg	acc	tca	tca	atc	tat	acc	caa	aat	ctg	aac	gtc	gcg	atg	aaa	gcc	2440		
Leu	Thr	Ser	Ser	Ile	Tyr	Thr	Gln	Asn	Leu	Asn	Val	Ala	Met	Lys	Ala			
				415					420					425				
att	aaa	ggg	ctg	aag	ttt	ggt	gaa	act	tac	atc	aac	cgt	gaa	aac	ttc	2488		

Ile	Lys	Gly	Leu	Lys	Phe	Gly	Glu	Thr	Tyr	Ile	Asn	Arg	Glu	Asn	Phe
			430					435					440		
gaa gct atg caa ggc ttc cac gcc gga tgg cgt aaa tcc ggt att ggc	2536														
Glu Ala Met Gln Gly Phe His Ala Gly Trp Arg Lys Ser Gly Ile Gly															
	445						450					455			
ggc gca gat ggt aaa cat ggc ttg cat gaa tat ctg cag acc cag gtg	2584														
Gly Ala Asp Gly Lys His Gly Leu His Glu Tyr Leu Gln Thr Gln Val															
	460					465				470					
gtt tat tta cag tct taa tgagtgaag aggcggagggt tttttcctcc	2632														
Val Tyr Leu Gln Ser															
	475				480										
gcctgtgcgc gtcagagttt agcgaatttt tcgagggtgc gaataagctg tgtgacgaag	2692														
ccatattcgt tatcgtagcca ggcgaccggt ttcaccagtt gtaaatacgcc cacggcggtta	2752														
atttccgttt gcgtggcatc aaacaccgaa ccgaaatggc tgccaatgat atcggaagag	2812														
actatttctt catcggtata accaaatgac tcgttatttg tggttgcttg tttaagtgcg	2872														
ttattcacct cttcggcagt cacttttttc cgagaatcga taccagttca gtgaccgaac	2932														
ctgttttcac cggcacgcgt tgcgcatgac ctttcagttt gccgctcagt tccgggatca	2992														
ccagaccaat ggctttttgcc gcccccgtag tgtgggggaat gatattttct gccgctgcgc	3052														
gtgaagcacg taaatcttta ccacgcgggc catccaccag tgactgggtg ccagtatagg	3112														
catgaatggt cgtcatcgtg ccgacttcta tcccgaact gtcatgcaag gctttggcca	3172														
tcggcgcaag acagttagtg gtgcatgacg ccacggaaac aatgggtgtcg ttgccatcca	3232														
gagtgtcgtc attgacgtta taaacgatag ttttcatttc accggcaggg gcggaaatca	3292														
acaccttctt cgcaccagca tcaagatgcg cctgcgattt ctgggcggag gtataaaaagc	3352														
cagtacattc gacaatgatt tctgcacctt tcgctttcca cggaatatat ttagcctctt	3412														
tttcggcgta aaccgcgata cttttcccat caacgataag tgaatcttcc gtaaaatcaa	3472														
cgctccaggg gaatgggtccg tagtttgaat catgtttcag caggtaggcg agaatttatg	3532														
gggaagtgag atcataata gcgacaacgt ctatgttgct tttgacttca agtaatcgac	3592														
ccaacaccag tcgaccgata cgacaaaaac cgттаатacc aactttactc atggttttct	3652														
cctgtcagga acgttcggat gaaaattgat cttttccaag cttagaccag gatggcggga	3712														
tgggcaatct ccattctcac agtgaaacgt aacgtactga aaacgggtga acaatattta	3772														
atgaaatttt gagaaaagcc cgttatgtta acggaaaatt atgttaaagc aggaaatggt	3832														
atggaaaata aatattcaag gttacaaatc agcattcact ggctgggtctt tttactgggt	3892														
atcgacgcgt attgcgcaat ggagtttcgt ggtttcttcc cacgtagcga tcggccactc	3952														

- 38 -

atcaacatga ttcattgtttc ctgtggcatc tcaattctcg tgctgatggc cgttcgtctt 4012
ttgttaaggc tgaaataccc aaccccgccg attataccta agccaaaacc gatgatgacg 4072
ggactggcgc atttgggaca tttgggtgatt tatctcctgt ttattgctgt gccagtgatt 4132
ggtttggtga tgatgtataa ccggggcaac ccgtgggttg cgtttgggtt gacgatgcct 4192
tacgcttcag aggccaattt cgaacgggta gatagcttaa agtcgtggca tgaaacgctg 4252
gcgaatctgg gatattttgt catcgggttg cagctgctgg ctgcactggc acaccactat 4312
ttctggaagg acaacacact tctacgcatg atgccgcgta aacgttcctg aaggatattt 4372
aaagaaaacg cctgtactaa aaccgaccg tggtacaggc gaagaatacg ggtctacatc 4432
ggaagcgcct attatattta tttgtatgat aaataaaacc ataatccttg cccatacgtc 4492
catctggctt atttttaatc aattcaccg atctttgatc tcatcaacgg tatcaaaata 4552
aaccagcgta taacgttcaa atcgaacata agattcgata aaccatggag gttatatgaa 4612
aaaactggca cttattttgt ttatgggaac gcttggttcc ttttatgccg atgccgggcg 4672
caaaccctgt tctgggtcga aaggggggat ctacactgt acggcaggcg gcaaatttgt 4732
ctgtaatgat ggttctatta gtgcatcgaa aaaaacatgc actaactgaa gtgtaaaagg 4792
ggtgccatga gaaaatgaat ctgctagtca aatgcgcggg gaaaatcccc gcgcttgccc 4852
ttacctggac gtgcaggcca tgagcgcagc aacctcctta tcaccgtccc ggaaccgcag 4912
ttcgtagagc gtttgtcggg tcaggagtgt gaatgtcaga atcgtaatgc agataatgag 4972
cagacacacc accagagggg tgtgcttcat agcctttctc cttgccggat ggcgggtaag 5032
aggctaagat ctgaattgct aggttcattc gttggcctcg gttgatagaa atatcggtcg 5092
gggccttcgt ctttctgatt cccgggttagc ctgaaaacag aaagtctcag gcaccgcag 5152
gcatcctatg aggtttcctt agggacgaaa ataatcactt cacgaaattg cgtgctgttt 5212
tccagaattt ttcgtcattc gggttagcca gtttagccat tcgttactct cttcattcca 5272
atagcattaa ttttctatgc aataattgtt gtaaaaatgt gacgcaaaga ggtttttggt 5332
cataagtaat taccgtcaag tgccgatgac tttctatcag gagtaaacct ggacgagaga 5392
caacggtaat gaatacaact ccctcacagc gattagggtt tttgcatcac atcagggttg 5452
ttccgttatt tgcctgcatt ctaggcggta tcttagttct attcgcatca agttcagccc 5512
tggtgggcta tttcctctgg caggccgatc gcgatcagcg tgatgttact gcggagattg 5572
agattcggac cgggttagcg aacagttcag attttttgcg ttcagcccg atcaatatga 5632
ttcaggccgg ggctgcgagt cgtattgcgg aaatggaagc aatgaagcga aatattgcgc 5692
aagccgaatc ggagattaaa cagtcgcagc aaggttatcg tgcttatcag aatcgaccgg 5752

tgaaaacacc tgctgatgaa gccctcgaca ctgaattaaa tcaacgcttt caggcttata 5812
tcacgggtat gcaacctatg ttgaaatatg ccaaaaatgg catgtttgaa gcgattatca 5872
atcatgaaag tgagcagatc cgaccgctgg ataatgctta taccgatatt ttgaacaaag 5932
ccgttaagat acgtagcacc agagccaacc aactggcgga actggcccat cagcgcaccc 5992
gcctgggtgg gatgttcatg attggcgctt ttgtgcttgc cctgggtcatg acgtgataa 6052
catttatggg gctacgtcgg atcgtcattc gtccactgca acatgccgca caacggattg 6112
aaaaaatcgc cagtggcgat ctgacgatga atgatgaacc ggcggtcgt aatgaaatcg 6172
gtcgcttaag tcgtcattta cagcaaatac agcattcact ggggatgaca gtagggactg 6232
ttcgacaggg cgcggaagag atttatcgtg gcaccagcga aatttcagct ggcaatgcgg 6292
acctgtcatc tcgcaccgaa gaacaagcgg cggctatcga acaaactgcc gccagcatgg 6352
agcaactcac tgcgacgggtg aaacagaatg cggataacgc gcatcatgcc agcaaactgg 6412
cgcaagaggg ttctattaaa gccagcgatg gcgggcagac gggttccggg gtagtaaaaa 6472
cgatgggcgc tatctccacg agttcgaaga aaatttctga gatcaccgcc gtcataca 6532
gtattgcttt ccagacgaat attctggcac tgaatgctgc cgttgaagcc gcgcgagcgg 6592
gtgagcaagg gcgtggattt gccgttgtcg ccagcgaagt acggacactc gcaagtcgca 6652
gcgctcaggg ggcgaaagag attgaaggct tgatcagtga atcagtcagg ttaattgacc 6712
tggggtcgga tgaggtggca acggccggga aaaccatgag cactattgtt gatgccgtcg 6772
cgagtgtcac acatatcatg caggaaatcg ccgccgcctc ggatgaacaa agtagaggca 6832
taacgcaggt tagccaggcg atttctgaaa tggataaggt gacgcaacag aatgcttctc 6892
tggtagaaga ggcctcagcg gcggcggtgt cccttgaaga acaggcggca cgattaactg 6952
aggcgggtga tgtattccgt ctgcacaaac attctgtgtc ggcagaacct cgcgagcgg 7012
gtgaaccagt tagtttcgct acggtgtgaa aatgttcaag gagggatcga cagatccctt 7072
cacctttcag aacggcattg attttcgaat agcgttaatc atcaactggc aaccagaaga 7132
gaacgtcgca tctacgcggg tcagtattcc aatcggttcg cctgcaccat gtcccggaac 7192
aggcagggcc accagcgtgg catgacgcag gtcgtctttt acagcgccag aagggacaaa 7252
ccacacgtaa tcgtattcaa ccgtaagttg acgagatagc gaagcagaca gcgtttcgat 7312
acaacccgaa ggaattttac agccctggct ctgcactaat gcatctgaat gctggcgtgg 7372
cgcagtcct tctggtgata caacgaccgg ccattccagc acccggtta gcgttacgtt 7432
ctcctgaagt agcgggtgat tagggcggac aaccagcttc aacgattcaa gaaacagcag 7492

- 40 -

ttcgtaatta agcccgggtca tcagttcagg atctgacatc cgaccaatgc cgatatcgat 7552
ttccccgggtt tttaaaccgg ccagaatcat agggttactc attgtcgcaa cttgcaaggt 7612
cgtctctttt tgttgctgat gaaactgacc tataaccgaa ggtaatatcc ccagtgccgc 7672
agtaggtagt gcaccaaccc tgacgacatc attattaaga ccttctttac gatgaagcga 7732
ccgtccggca gtgttgatgg cgtcaagaac tctgactgca tgcgttaaaa attgttcgcc 7792
gggtaaggta agttgcgccc cctgacgacc acgctcaaac aagcgagcgc cagtcagctg 7852
ctccagttca ttcaatgtct tagagagcgc aggttgactc aaattaaggg tttcagccgc 7912
gcgccccaaa gttccttggt gtgcgacagc tacgaatgta tgaagggtggc gcaaacggat 7972
gcgctgacta aacagactat ttttttccat aagcgatggt aaaaacgaag cgggtgctgct 8032
gacaagtga gttgtttgat tatgataact tgattgcaaa atattattaa caattaaagc 8092
aattatgtta cagcaaaatg gataatattg atgttttcgc ggcgagatca cagtttgtaa 8152
attcttcccg caagagtga tgcgggtacc tacactccag attactgacc actggaggca 8212
gacactatgg cgaacagcat cacggcggat gagattcggg aacagttttc gcaggcaatg 8272
tcagccatgt accagcaaga agttccgcaa tatggcacgc tgctggaact ggtagctgat 8332
gtgaatctgg ctgtgctgga aaacaatcct caactgcacg aaaaaatggg aaatgcagac 8392
gagctggcgc gactgaatgt tgaacgtcat ggggcgattc gcgttgggac tgcacaagag 8452
cttgctactc ttcggcggat gtttgccatt atggggatgt acccggtgag ctattacgat 8512
ctctcgcagg caggggtgcc ggtacattcg acagcatttc ggccattga tgatgcttct 8572
ctggcgcgta atcccttccg cgtttttacc tccttactcc gccttgagct tatcgagaac 8632
gaaattttgc gccagaaagc ggcggagatt ctgcgtcagc gcgatatctt caccaccagt 8692
tgtcgacaac tgttagagga atatgagcag cagggcgggt ttaacgaaac acaggcacag 8752
gagtttgtgc aggaagccct ggaaacgttt cgctggcacc agtcagcaac ggtagatgaa 8812
gaaacctatc gcgcattgca caacgaacat cggttgattg ctgatgtggg ctgttttcct 8872
ggatgccata tcaaccacct gacgccacgt acgctggata ttgaccgggt gcagtcgatg 8932
atgcctgaat gcggaattga acccaaaatt ctcatcgagg ggccgcgcgc ccgcgaggta 8992
cogattttac tacgccagac cagctttaaa gcaactggaag agacggtggt gtttgcgggg 9052
cagaaacagg gcacgcatac cgcgcgcttt ggtgaaattg agcagcgtgg cgtggcatta 9112
acgccgaaag ggcgacaact gtatgatgat cttctgcgta acgctggaac cgggcaggat 9172
aatctcactc accaaatgca ttacaggaa accttccgca cttttcctga cagtgagttt 9232
ttaatgcgtc agcaagggtt ggcattggtc cggtaaccgtc tgacgccttc gggtagggcg 9292

catcgtcagg cgattcatcc tggagacgat ccacagccct taattgaacg tggttgggta 9352
gtggcgcaac ccatcaccta tgaagatttc ttgcccgtta gcgcggcggg gatcttccag 9412
tcaaactctgg gtaatgaaac gcagacacgc agtcacggta atgccagtcg cgaagcattt 9472
gagcaggcgt tgggttgtcc ggttttggat gagttccagc tttaccagga agcggagaa 9532
cgcagtaaac gtcgctgtgg tttgctttaa aatctgacca tccgcctttg caaaaatttg 9592
cctgatttta caaacgaatc aggctcatcc catcgacata aaaaaaatgc cgatttatgc 9652
atattctctc agttcaacaa ttggattatt aataaatatt gtctagagtg agcggtcata 9712
aataagcact ttcttgccgc tgaaaacgac cagcgcggga ccattcaca caccagaagg 9772
actcactttc aggtatggat cgtagacgat ttattaaagg ttcaatggct atggccgccg 9832
tgtgcggtac cagcggcatt gcttctcttt tttctcaggc ggcattcgcg gcagattctg 9892
atattgccga cgggcaaacc cagcgttttg acttctccat tctacagtca atggcgcacg 9952
acttagcgca aacagcgtgg cgtgggtgcgc ctgcgtccgtt acctgacacg ctggcgacaa 10012
tgacgccgca ggcttataac agtattcaat acgacgccga aaaatcgctc tggcataacg 10072
ttgagaaccg tcaactggac gctcagttct tccatatggg aatgggattc cgtcgccgcg 10132
ttcgtatgtt ttctgtagat ccagcaacac atctggcgcg tgaaattcac tttcgcccgg 10192
agttgttcaa atacaacgat gcaggtgttg atacaaaaca attagaaggg caaagcgatc 10252
tcggctttgc cggttttcgc gtgttttaaag cccccgaact ggcgcgccgt gatgtagtat 10312
catttctcgg cgcgagttat ttccgcgccg ttgatgatac atatcaatac ggtttgtcgg 10372
cccgcggcct ggcgatcgac acttacaccg acagtaaaga agagttcccc gactttaccg 10432
ccttctgggt tgatacggta aaaccggggg caactacctt taccgtttat gcgttgctcg 10492
atagcgccag cattactggg gcctataagt tcactatcca ttgtgagaaa agtcagggtga 10552
ttatggatgt ggaaaatcac ctgtatgcgc gcaaagacat taaacagctg ggcattgcgc 10612
cgatgaccag tatgttcagc tgcggtacta atgaacgtcg gatgtgcgat acaattcatc 10672
cgcaaattca tgactctgat cgtctgtcca tgtggcgggg caacggcgag tggatttgcc 10732
gtccgctgaa taatccgcaa aaattgcagt tcaatgctta caccgacaac aaccgaaag 10792
ggtttggttt attgcaactg gatcgtgact tctcccatta tcaggacatt atgggctggg 10852
ataacaaacg cccaagtctg tgggtggaac cgcgtaacaa gtggggtaag ggcaccatcg 10912
gcctgatgga aatcccaaca acgggcgaaa cgctggataa cattgtctgc ttctggcagc 10972
cagaaaaagc tgtaaaagca ggtgatgagt ttgcattcca gtatcgtctg tactggagtg 11032

- 42 -

cgcaaccgcc tgttcattgc ccattagcgc gcgttatggc gacgcgtacc ggcatgggcg 11092
 gtttctcgga aggttgggcg ccaggtgaac actatcccga aaaatgggcg cgtcgttttg 11152
 ccgtcgattht cgttggtggt gatctgaaag ctgccgcgcc aaaaggcatt gagccggtga 11212
 ttacgctttc cagtggggaa gcgaagcaaa tcgaaattct ctatatgaa cccatcgatg 11272
 gttatcgtat tcagtttgac tggatatccga cttcggactc cactgatccg gtcgatatgc 11332
 ggatgtatct acgttgtcag ggggacgcta tcagtgaaac atggctgtat cagtatttcc 11392
 cgccagcgcc ggataaacgt cagtatgttg acgaccgcgt gatgagttaa tcgttttttc 11452
 ttcggcacct tcttcgggag gtgccgtctg gttaaacacg atcccgcctc catttttccc 11512
 taagttaa at gagtaatctg atgggtgtgta tttcagatac accttgtcag ccactaacag 11572
 ggagtgcgta tgtttccaga ataccgagat ttaatatccc gtctgaaaaa cgaaaatc 11630

<210> 9

<211> 479

<212> PRT

<213> Escherichia coli

<400> 9

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

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

- 44 -

<210> 10

<211> 12677

<212> DNA

<213> Escherichia coli

<400> 10

gtcatttgct cctttaatca gctgtcgcgt tcccctgcc tataaaagga gggatatgcac 60
cacgatgggt cattacccaa taagattgaa agctcaccac ttgttgaaa ttgacagcaa 120
acaaacaaaa aaatgcatth caccctttga catcaccatg cactgccatt aatatgcgcc 180
ccgttcacac gattcctctg tagttcagtc ggtagaacgg cggactgtta atccgtatgt 240
cactggttcg agtccagtca gaggagccaa attcaaaaaa gcctgctttc tagcaggctt 300
tttgctttct aattaccaac gctcttaaaa catctgtctt gaaccagaac taatttgcac 360
aggcattccc gatcgacgtt gcaacgcagc atttgcgaga ttacatcaa cttcttgccc 420
gttgataaac gcccgcaaag atgggggttac cggcaatggc acttttcggt cagactcata 480
ttctgcacga ttgcgcgaca atgggtcatg aacttccagc cagttcgagc catctgggtc 540
agtgggtgtat ttactgggt ggtagataat ttgcacacgc gtcccaacag gaacattatc 600
aacagatat ttgatatcgt cattgcgcag acgaatacag ccctgactta cccggagccc 660
aatacaaaaa ttggcattgg taccatggat ggcatacaac ctgccaatat aaatcgcgta 720
cagcccatg ggattatcgg ggcccgagg aacaaatgcg ggcaaactct cccctcgttt 780
cgcatattcg cgccgagtgt tcggcgttgg cgtccagggt ggagcttctt gtttacgttc 840
aacggtagtc acccagttac gcgggggttc tcgcccagcc tggccgatac caataggaaa 900
gacttccaca gtattactgt ctgggtgggt gtaataaaga cgcatctcag cgacgttaac 960
aacaatccct ttacgaacag tgtcgggcaa aatcagttgc tgcggaatgg tgagttgcga 1020
gccagacttc ggcaaaaaaa catcagcgcc cgggttcgct tccagcatgt tacttaacct 1080
ttgcccgtat tgtgcggcaa aagtctccag cggctgggtt ttgtgatcag gaacagttac 1140
agtaaocgac tgccccacta aacggctacc ctctggaggt aatggataag ttaccgccag 1200
gctagtatgg ctggcaaaaa gcagagcaaa tgagcaaaga atatttacac gacgcatcat 1260
gtccctttcc tatgtcgcga aagctatccg ttaagtatag cttttatcag acttttcggt 1320
tttaactggt caaatcagaa gtctattcc ccggtagaac aatattactg gcagcaagtt 1380
cgcccatggt gttgtatatc gcacaggcag ctctgatgat gggcatcgcc agagctgcgc 1440
cactccctc acctaaacgc atctccatat tgagataagg ctccagcccc aaatgcgaga 1500
gcgtatacgc cgccctttt tctgccgaca ggtgagaagg aatgagatac ggtttgatcg 1560
caggagacat ctggcaggct gcgagcgag cagcataaga aagaaatcca tccagcaaca 1620

- 45 -

cgggtaaacc gcaggaagca gcacctaaca tcactccagc tattccgacc aaatcaaate 1680
caccactttt tgccaggaca tcaacaccat cctgaggatt tggttgattc aacgtaatec 1740
cccgacgcac aacatcaatt ttattagcca gtttatctgt cggcaggttt gcgccaatcc 1800
caaccacttc ttcaggatcc cggccagtga ttgtgctgac tattgctgct gccggtgtcg 1860
tgtttgccat cccagtttca cctacaccaa acagcgtgac accgtttttt gccagctcct 1920
gcgtataaca tatgacgtcc aaaagcaact tttcagcctg acggcgactc attgccggag 1980
ctgaagcaat attgccgcta cctcgtgcga cacgcatgtt gataagcccg gggataggct 2040
cagcagtatc aataccaaca tcaattacgt ggacgttagc gcccgcttgt tctgccagca 2100
cacacacgcc ggttgttcca cgggtcatat tttcagcctg tatggctgtc acttcttttg 2160
gagaaatagc gaccccttcc tcccagacgc cgtgatcggc acacataacc agtaccgctt 2220
ttttgcccac atgcggtatg ccattcaacc ccggcattcc tgccagttgt atggcaagca 2280
cctccagctt tcccaggcta ccaacagggt tgagtaacct gtcaatatgc cgttgtgcac 2340
gcgacatagc ggtagaatcg atggctggga tcgtattcag taaatcggca agtatattgca 2400
tctcacgtct ctttatagca gtgccagcag gaagaccagt tcaccaagct ctatggctgc 2460
tcccagegta tcgcccgttt gcccgcccag cgtacgtttg agaagctgac cgaggatgaa 2520
aatcgccacc atcgtcacca ccatagcagc cacaccatgc ataccggca ataatacggc 2580
agcaaaaata gcggctaagc cgagtgtgac gcaggtttgt cgcccatcaa ttttgccaat 2640
gaacacattg ccaagccctt ctccccgcgc ataacgatga cgatacatca ataatgcggc 2700
agtcccgcgg ctgaccgcac atgctgccgc cagcgatgca aggatcgatt cgccacgcag 2760
agccaactcg cttaacacca gaatctttgc cagtaccaca aaaatcaatg ccagaccgcc 2820
gtgggtgcct aaacggctat cacgcatgat ttccaacatg cgatcgcggc tacgtgcaga 2880
aaatacgcca tcacaggatc cagcgaggcc gtccagatgg aatccccggc tcatcagcac 2940
cagcacaagt acgctaaaca gtgccgccag tgggtgcgcca caccatgcct gcagcaccat 3000
gaagaccagc ccgctaatec cgccaagcaa taatccaate aaaggaaaag taataatacc 3060
gcgagaataa tgctcgaaat ccagtccttg ggaccagcga cgcggtacag gcagacgcgt 3120
aataaatgag agcatcgccc aaaataattt actcatttga tttttactcc aataccgaa 3180
accaccagcc atacttcatt tgccgcagcg gccaaactgt gatttaccgc cccggcaata 3240
tcacgaaaat gtcgtgccag acgactctcc ggcacaatcc ccattcccac ttcgttagtc 3300
actaatacaa cctttgcggg gcaacgttgg caggcagcaa tcaacgactg aatctcagca 3360

- 46 -

ttaatcgccct gttccatcgc ctgataatcc cattcatcag ggtctttatc gccgccataa 3420
tcaaacaaca gattagtcac cattgtggta acgcattcaa gcaacacaac ctcatcggg 3480
ttaatgtctg catgaattaa ttcataaga tgttgccagc gctccactgt gcgccagtgc 3540
tccggggcggc cttgccgatg atgttctatc cgtgcagcca tctcatcatc aaggatttgc 3600
gaggtagcga tatacagaac ctgtgaagag tccccataa gcgcctctgc gtggcgactc 3660
ttcccgtcc gtgcgcccc cgtcaccaa atcatcatc cgactcctga ttcaaagtga 3720
tgatgggtatt cattgaacag acaatggat cccccaccgt tgctgaccgc aaaacattta 3780
cgcgattaat actgcgcgta atataattaa aaatccagca ttctcaatcc attacgaaat 3840
aagatggcat tacggaatag ccagactttc tcgtgttaac gtccgacaaa atgcgtcatt 3900
gacagcgggt tctgtggatt gacacatttc attcagataa tgaattaatg ctactgccgg 3960
aacaaccagc aaacgggcat tgtgttctga aatccaggct attgattcaa cgtcagcgac 4020
agatctgcgc tgtaaaagaa cttgtaaaac aacgtaatag aattatccgg aatgggtggcg 4080
actatgcact aggaagggtg cgaataagtg gggaaattct tctcggctga ctcagtcatt 4140
tcatttcttc atgtttgagc cgattttttc tcccgtaaat gccttgaatc agcctattta 4200
gaccgtttct tcgccattta aggcgttatt cccagttttt agtgagatct ctccactga 4260
cgtatcattt ggtccgccc aaacagggtg gccagcgtga ataacatcgc cagttgggtta 4320
tcgtttttca gcaaccctt gtatctggct ttcacgaagc cgaactgtcg cttgatgatg 4380
cgaaatgggt gctccacctt ggcccggatg ctggctttca tgtattcgat gttgatggcc 4440
gttttgttct tgcgtggatg ctgtttcaag gttcttacct tgccggggcg ctccggcgatc 4500
agccagtcca catccacctc ggccagctcc tcgcgctgtg gcgccccttg gtagccggca 4560
tcggctgaga caaattgctc ctctccatgc agcagattac ccagctgatt gaggtcatgc 4620
tcgttggccg cgggtggtag taggctgtgg gtcaggccgc tcttggcatc gacaccaatg 4680
tgggccttca tgccaaagtg cactgattg ctttcttgg tctgatgcat ctccggatcg 4740
cgttgctgct ctttgttctt ggtcgagctg ggtgcctcaa tgatggtggc atcgaccaag 4800
gtgccttgag tcatcatgac gcctgcttcg gccagccagc gattgatggc cttgaacaat 4860
tggcgggcca gttgatgctg ctccagcagg tggcggaat tcatgatggc ggtgcggctc 4920
ggcaaggcgc tatccaggga taaccgggca aacagacgca tggaggcgat ttcgtacaga 4980
gcatcttcca tcgcgccatc gctcaggttg taccaatgct gcatgcagtg aatgcgtagc 5040
atggtttcca gcggataagg tcgcccggca ttaccagcct tggggtaaaa cggctcgatg 5100
acttccacca tgttttgcca tggcagaatc tgctccatgc gggacaagaa aatctctttt 5160

- 47 -

ctggtctgac ggcgcttact gctgaattca ctgtcggcga aggtaagttg atgactcatg 5220
atgaaccctg ttccatggct ccagatgaca aacatgatct catatcaggg acttgttcgc 5280
accttcccta gaacaccaca atttcgctct ctcggttaagg cactgtcaca gcattcaaca 5340
gaatgtgact tgccagattt attagcgcca ccagatgttt aaccgggtaa ccacgacca 5400
gtcggggata tattccatct tctcgctgac atcatcaaga ataaaaaggt ttatcacact 5460
cagagcaacg ttcttcagca ggcgggataa ttcttcaaca tttagaagat gtgttattaa 5520
aggcatacaa actggacgta ttgtattttc ttttggtggt tgacctagg tagcagttta 5580
tcctgatgcg ctgagatttc tgtaatatct acgtcaacat tctcctcgaa tagaaaatgc 5640
agccccggca agtgattcac attttttaac aaaacgttgt tagcgtgcca ttttcagaag 5700
ctctccaaga aagcgaattc ttttctcaaa ttctgcattc tcagacacag ctttctactc 5760
catggtaatg gcccaaatac gaagttgctc aggatcatta gtggtgctga gaagttaa 5820
aaccatgcgg agtaccactc gtcatcacgt atttatatat catgaagcta ttgttggtgtt 5880
ttgtacatca gtagatatat tctgatatac tccttttgct agacataacc tttcacctgc 5940
ttgcaaagct tctgtgttct gacattgcca aattgttgca attctgtatc cagccttctt 6000
tcagtcatag cttcggggcg cgataagact cactgatctg accctgattc ctcttgacaga 6060
ctttatagac caattaaaat gcagtttctg cagggtcaacg tctgaccatc attgtcatca 6120
ctctggccat tagagtaacc ttctgcattc atccttttgt aaaaagttaa tattagtatc 6180
agcaattaac cggacctgat actgatatga gtcttacgc atatacggtc aatttcagca 6240
attaattaca ttatccacgc caaagtattt gtcatcaca tgatgggtacc ttctttcaga 6300
caccatTTTT tcaactccgt tttccacgga ccgcactctt atgtcaagag tgcgggtccgt 6360
ggatacaacc agagaccgac tgacacgagt cagaggaaac gacggatatg ttcagtcgta 6420
aaatatctat caaaaaacat gattaaggtc aaaaatgttt gatatttaca atttatgaag 6480
atgacaataa ttatagatat atgagaacat aatgaaaat aattatcatt acagtaatca 6540
tttgtacttt gtattaatga gggatgaaat gttatataat ataccttgtc gaatttatat 6600
cctttccact ctgtcattat gcatttctgg gatagtttct actgcaaccg caacttcttc 6660
agaaacaaaa atcagcaacg aagagacgct cgtcgtgacc acgaatcgtt cggcaagcaa 6720
cctttgggaa agccccggca ctatacaggt tattgaccaa caaacattgc agaactccac 6780
caatgcctcc atagccgata atttgcagga catccccgga gtagagataa cagacaactc 6840
cttggcaggc cgtaaacaaa tccgcattcg tggctggatt tgcccctata tttccagaca 6900

- 48 -

tctgttatca ctttaacccat tacaagcccg ctgccgcaga tattcccgtg gcgagcgata 6960
accagcgca ctatgcggat gccattcggt ataatgctcg aacgcctctg caagggttctt 7020
tgctgccgtt aaccggtctg gtttgggcat gatactgatg tagtcacgct ttatcgtttt 7080
cacgaagctc tctgctattc cggtactctc cggactccgc accgccgtgt tcttcgggtc 7140
aagtcccaac atccggggcga actggcgtgt ttcattagcc cggtagcatg aaccattatc 7200
cgtcagccac tccactggag acgacggaag atcggtgccg aagcggcggt ccaccgctcc 7260
cagcatgacg tcctgtactg tttcactggt gaagccgccg gtagtgaccg cccagtgcag 7320
tgccctcacga tcacagcagt ccagcgcgaa cgtgacacgc agtctctctc cgttatcaca 7380
gcagaactcg aaccggtcag agcaccatcg ctgattgctt tctttcacgg ccactctgcc 7440
tgtatgtgcc cgtttcgatg gcggtacagc aggttttcgc tcaagcaaca gcgcattctg 7500
gcgcattgat cggtaaaccac gtttggcatt gatcgaggc ataccatcaa gttctgcctg 7560
tctgcgaagc agcggccata cccgacgata accatacgtt ggcagctctc cgataacatg 7620
gtgtatacgg agaagcacat ccgtatcatc agtgtgacga ctgcggcggc catccatcca 7680
gtcatcggtt cgtctgagaa tgacgtgcaa ctgcgcacgc gacacccgga gacaacggct 7740
gactaagctt actccccatc cccgggcaat aagggcgctg gcgctatcca cttttttgcc 7800
cgcccatatt caacggcttc tttgaggagt tcattttcca tcgttttctt gccgagcagg 7860
cgctggagtt ctttaactctg cttcatggcg gcagcaagtt cagaggcagg aacaacctgt 7920
tctccggcgg cgacagcagt aagacttcct tcctgggtatt gcttacgcca gagaaataac 7980
tggttggtg ctacaccatg ttgccgggca acgaaggaga ccgtcatccc cggttcaaag 8040
ctctgctgaa caattgcgat cttttcctgt gtggtacgcc gtctgcgttt ctccggccct 8100
aagacatcaa tcactctgttc tccaatgact agtctaaaaa ctagtattaa gactatcact 8160
tattttaagt atattgggtg tctggagatt cagggggcca gtctagtggc gaagcatcct 8220
cccgtgtttt aattctcatt gatggtcagg aggttaactta tcagcgcgcc ggagataatt 8280
atggtgtggg actgttgata gatgagtctg cgctggagcg tggtgaggta gtgaaaggtc 8340
catattccgt actgtacggc tcacaggcaa ttggcggtat tgttaacttc atcaccaaaa 8400
agggtgaatc tccggactcc ctatatcact taaattgata caacttttta gagtagtcat 8460
tagtgaacag ataattgata actcagaacc agtaaaatgc aaaagacgca ccacgctggg 8520
aaaatcgctg ctattcagca gagttttgta tcgagaatgg tggccagcca cttattgctc 8580
tcgcgtaagc gggtagcgtg acattctgcc tgaacttgac ctggtactgt ggctgattaa 8640
agccgatgac cgtgccctgt ctgtggatga gtatttctgg cgacacatcc tgcagtgcgg 8700

- 49 -

acatcagcag gtgctgtttg tggtagcgca ggccgacaaa acggagccct gccatgaatg 8760
ggatatggcc ggtattcagc cttctcctgc acaggcacag aatattcgcg aaaaaacgga 8820
ggcggtatte cgtctgttcc ggcccgta ca tccggttgtg gccgtatcgg cccgcaccgg 8880
ctgggaactg gatacgctgg tcagtgcgct catgacagcg cttcccgacc atgccgccag 8940
tcccctgatg acccgactgc aggacgagct ggcacaggaa tctgtccggc ctcaggcccc 9000
tgaacagttt accgggtgcgg tggaccggat atttgacacg gcggagagcg tctgtattgc 9060
ctctgttgca cgcacgggtc tgcgtgccgt ccgtgacacg gtggctctctg ttgcccgctgc 9120
ggtatggaac tggattttct tctgaacctg tcgtgactga tgccctccct gactctgagt 9180
ctgctcacia aagcactgtt ttcgttactg tctctcttgt ccgtgcaata gctcaataat 9240
agaataaaac gatcaatata tattttatcg atcgtttata tcgatcgata agctaataat 9300
aacctttgtc agtaacatgc acagatacgt acagaaagac attcaggga caacagaacc 9360
acaattcaga aactcccaca gccggacctc cggcactgta accctttacc tgccgggtatc 9420
cacgtttgtg ggtaccggct tttttattca ccctcaatct aaggaaaagc tgatgaaacg 9480
acatctgaat acctgctaca ggctgggtatg gaatcacatg acgggcgctt tcgtgggtgc 9540
ctccgaactg gcccgcgcac ggggtaaacg tggcgggtgtg gcgggtgcac tgtctcttgc 9600
cgcagtcacg tcaactcccg tgctgggtgc tgacatcggt gtgcaccccg gagaaaccgt 9660
gaacggcgga aactggcaa atcatgacaa ccagattgtc ttcgggtacga ccaacggaat 9720
gaccatcagt accgggctgg agtatgggcc ggataacgag gccaataccg gcgggcaatg 9780
ggtacaggat ggcggaacag ccaacaaaac gactgtcacc agtgggtggc ttcagagagt 9840
gaaccccggt ggaagtgtct cagacacggc tatcagtgcc ggaggcggac agagccttca 9900
gggacgggct gtgaacacca cgctgaatgg tggcgaacag tggatgcatg agggggcgat 9960
agccacagga accgtcatta atgataaggc ctggcaggct gtcaagcccgt gtacagtggc 10020
aacggatacc gttgttaata ccggggcgga agggggaccg gatgcagaaa acggtgatac 10080
cgggcagttt gttcgcgggg atgccgtacg cacaaccatc aataaaaacg gtcgccagat 10140
tgtgagagct gaaggaacgg caaataccac tgtggtttat gccggcggcg accagactgt 10200
acatggtcac gcactggata ccacgtgaa tgggggatac cagtatgtgc acaacggcgg 10260
tacagcgtct gacactgttg tgaacagtga cggctggcag attgtcaaaa acgggggtgt 10320
ggccgggaat accaccgtta atcagaaggc cagactgcag gtggacgccg gtggtacagc 10380
cacgaatgtc accctgaagc agggcggcgc actggttacc agtacggctg caaccgttac 10440

- 50 -

cggcataaac cgcctgggag cattctctgt tgtggagggt aaagctgata atgtcgtact 10500
 ggaaaatggc ggacgcctgg atgtgctgac cggacacaca gccactaata cccgcgtgga 10560
 tgatggcgga acgctggatg tccgcaacgg tggcaccgcc accaccgtat ccatgggaaa 10620
 tggcggtgta ctgctggccg attccggtgc cgctgtcagt ggtacccgga gcgacggaaa 10680
 ggcattcagt atcggaggcg gtcaggcgga tgccctgatg ctggaaaaag gcagttcatt 10740
 cacgctgaac gccggtgata cggccacgga taccacggta aatggcggac tgttcaccgc 10800
 caggggcggc aactggcg gaccaccac gctgaataac ggcgccatac ttaccctttc 10860
 cgggaagacg gtgaacaacg ataccctgac catccgtgaa ggcgatgcac tcctgcaggg 10920
 aggctctctc accggtaacg gcagcgtgga aaaatcagga agtggcacac tcaactgtcag 10980
 caacaccaca ctcaccaga aagccgtcaa cctgaatgaa ggcacgctga cgctgaacga 11040
 cagtaccgtc accacggatg tcattgctca gcgcggtaca gccctgaagc tgaccggcag 11100
 cactgtgctg aacggtgcca ttgacccac gaatgtcact ctgcctccg gtgccacctg 11160
 gaatatcccc gataacgcca cgggtgcagtc ggtggtggat gacctcagcc atgccggaca 11220
 gattcatttc acctccacc gcacagggaa gttcgtaccg gcaaccctga aagtgaaaaa 11280
 cctgaacgga cagaatggca ccatcagcct gcgtgtacgc ccggatatgg cacagaacaa 11340
 tgctgacaga ctggtcattg acggcggcag ggcaaccgga aaaaccatcc tgaacctggt 11400
 gaacgccggc aacagtgcgt cggggctggc gaccagcggc aagggtattc aggtggtgga 11460
 agccattaac ggtgccacca cggaggaagg ggcctttgtc caggggaaca ggctgcaggc 11520
 cgggtgccttt aactactccc tcaaccggga cagtgatgag agctggtatc tgcgcagtga 11580
 aaatgcttat cgtgcagaag tccccctgta tgctccatg ctgacacagg caatggacta 11640
 tgaccggatt gtggcaggct cccgcagcca tcagaccggt gtaaattggtg aaaacaacag 11700
 cgtccgtctc agcattcagg gcggtcatct cggtcacgat aacaatggcg gtattgcccg 11760
 tggggccacg ccggaaagca gcggcagcta tggattcgtc cgtctggagg gtgacctgat 11820
 gagaacagag gttgccggtg tgtctgtgac cgcgggggta tatggtgctg ctggccattc 11880
 ttccgttgat gttaaggatg atgacggctc ccgtgccggc acggtccggg atgatgccgg 11940
 ctgcctgggc ggatacctga atctggtaca cacgtcctcc ggcctgtggg ctgacattgt 12000
 ggcacaggga acccgccaca gcatgaaagc gtcacggac aataacgact tccgcgcccg 12060
 gggctggggc tggctgggct cactggaaac cggctctgcc ttcagtatca ctgacaacct 12120
 gatgctggag ccacaactgc agtatacctg gcagggactt tcctggatg acggtaagga 12180
 caacgccggg tatgtgaagt tcgggcatgg cagtgcacaa catgtgcgtg ccggtttccg 12240

- 51 -

tctgggcagc cacaacgata tgacctttgg cgaaggcacc tcatcccgtg cccccctgcg 12300
 tgacagtgca aaacacagtg tgagtgaatt accggtgaac tgggtgggtac agccttctgt 12360
 tatccgcacc ttcagctccc ggggagatat gcgtgtgggg acttccactg caggcagcgg 12420
 gatgacgttc tctccctcac agaatggcac atcactggac ctgcaggccg gactggaagc 12480
 ccgtgtccgg gaaaatatca ccctgggctg tcaggccggt tatgcccaca gcgtcagcgg 12540
 cagcagcgtc gaaggggtata acggtcaggc cacactgaat gtgaccttct gacagaacca 12600
 tcgcctctct gtggtcccgg tcatcatgac cgggaccgga accggcgcaa cggatcttca 12660
 acgccacatt cgctggc 12677

<210> 11

<211> 181

<212> PRT

<213> Escherichia coli

<220>

<223> complement of position 3153-3698 of seq id 10

<400> 11

Met Met Ile Leu Val Thr Gly Gly Ala Arg Ser Gly Lys Ser Arg His
 1 5 10 15

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

Ser Gln Ile Leu Asp Asp Glu Met Ala Ala Arg Ile Glu His His Arg
 35 40 45

Gln Gly Arg Pro Glu His Trp Arg Thr Val Glu Arg Trp Gln His Leu
 50 55 60

Asp Glu Leu Ile His Ala Asp Ile Asn Pro Asn Glu Val Val Leu Leu
 65 70 75 80

Glu Cys Val Thr Thr Met Val Thr Asn Leu Leu Phe Asp Tyr Gly Gly
 85 90 95

Asp Lys Asp Pro Asp Glu Trp Asp Tyr Gln Ala Met Glu Gln Ala Ile
 100 105 110

Asn Ala Glu Ile Gln Ser Leu Ile Ala Ala Cys Gln Arg Cys Pro Ala
 115 120 125

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

Ser Arg Leu Ala Arg His Phe Arg Asp Ile Ala Gly Arg Val Asn Gln
 145 150 155 160

Gln Leu Ala Ala Ala Ala Asn Glu Val Trp Leu Val Val Ser Gly Ile
 165 170 175

- 52 -

Gly Val Lys Ile Lys
180

<210> 12
<211> 12790
<212> DNA
<213> Escherichia coli

<400> 12
tcactataga tattgatcat taagttgatt agacccaaaa tcatgattag acctatcatt 60
taaattgatta atagattgca tggagatact gaatgaagcg tgaagaaatc gctgatctga 120
tggcggttgt cgtcggtgca gaggagcgta gcttcactcg tgcagcagcc cgcctgagca 180
tggcgcgatc agctttaagc cagatagtgc gtcgtataga agaacgattg ggattgcggc 240
ttctgacgcg aaccacgcg agcggtgttc caactgaagc gggcgagcat cttttgtctg 300
ttcttggccc tatgttgcat gacatagatt cagccatggc atccctgagc gatctgcaga 360
accgcccata cgggacaata cgtattacta ctgtagaaca tgcagcaaaa acgatattgt 420
taccagcaat gcgcacattc ctgaaatcgc atcctgaaat tgatattcag ctcaccattg 480
attatggttt gaccgatgtc gtttctgaac gttttgatgc aggcgctcgt ctgggtgggg 540
agatggataa agatatgac gccattcgaa tcggggcaga tataccaatg gctattgttg 600
gctcaccgga ttatttttct cgcgaagtg ttccaacgtc agtgtcaca ttaatagatc 660
atcaggcaat taatttgtat ctcccacat cgggtacagc aaatcgctgg agattaatac 720
gcgggtggacg tgaagttcgt gttcgcatgg aaggtcagct tttactgaat acgatagacc 780
tgatcattga tgctgcaatt gatgggcatg gattggcgta tctacctat gatcagggtg 840
agcgggctat taaagaaaaa aaactgatac gtgttttgga taaattcaca ccagatttac 900
ccggttatca cctgtactat ccacaccgtc gacatgctgg ctggcattc tcattattta 960
tagataggct gaagtataaa ggtgctgttt agcactactt gctgatacat taatttaatt 1020
cttctcttaa cgtattctca gttcctttca acgttttggt ctttttttat tcttcgtaca 1080
atggcgacag atgctgatta tgataccgaa aacgggtttg aacgtgcgaa gcccgaaacgt 1140
agtgttcgga gttctatgtg ctttaccgca ttttgagac tattatttac actaaatctg 1200
atttgatata ttgatactta aaacatttga tgcttccttt gtcacttttt tgatggaagt 1260
tgtttgcat tctttaaggc gaaacaaata attacgcatc aattttaatg tcggttagag 1320
ggaaacttat gaagcactct gtttcagtca cgtgttggtc gctgttggtc agcagcattt 1380
ctctttcgta tgctgcagaa gttccgagcg gcacagtact ggcagagaag caggagctgg 1440

- 53 -

tgcgccacat taaagatgag cctgcgctcg tggatcccgc taaagccgtg ggcctgccag 1500
 agattcaggt cattcgcgat ctgtttgaag gtctggtgaa tcagaacgaa aaaggggaga 1560
 ttgtccccgg cgttgcgact cagtggaaaa gtaatgacaa ccgtatctgg actttttacc 1620
 tgcgcgataa cgcaaaatgg gcggatggca caccggtaac ggcgcaagat tttgtctaca 1680
 gctggcaacg tctggtggac ccaaaaacat tgctgccatt tgcattggtt gccgcgctgg 1740
 cggaatcaa caacgcacag gcgattattg atggtaaagc tacgcctgac cagcttggcg 1800
 tcaccgcagt tgatgccccat actttgaaaa ttcagcttga taaaccgttg ccgtgggttg 1860
 tgaatttaac cgctaacttt gccttcttcc cggtgcaaaa agccaacgta gaaagcggta 1920
 aagagtggac gaaacccgga aatctgatcg gcaatggcgc ttatgttctt aaagagcgcg 1980
 tagtcaatga aaaactggtc gtggtaccga ataccatta ttgggataac gccaaaacgg 2040
 tactgcaaaa agtgaccttc ctgccaatta atcaggaatc cgcagccact aagcgttacc 2100
 tcgcggggga tattgatatc accgaatcct tcccgaaaaa tatgtatcag aagctgttga 2160
 aggatattcc ggggcagggt tatacgcgcg cgcagctcgg gacctattat tatgcgttta 2220
 acacgcaaaa agggccgacg gcagatcagc gcgttcgtct ggcattaagt atgacgatag 2280
 atcgcgcgct gatgaccgaa aaagtattag ggacgggcga aaagccagcg tggcatttta 2340
 caccagatgt taccgcggga ttacgcgcg aaccttcgcc gtttgaacaa atgagtcagg 2400
 aagaactgaa tgcgcaggca aaaactttgt tgagcgcagc tggttatggt ccgcaaaaac 2460
 cgctgaagct gacgcttttg tataacactt cagaaaacca tcaaaaaatt gcgattgctg 2520
 tagcatcgat gtggaaaaag aaccttggcg tagatgttaa attgcaaaat caggaatgga 2580
 aaacctatat cgatagccgt aacaccggca attttgatgt tatccgcgcc tcgtgggtgg 2640
 gggattataa tgaaccctcc actttcctga cattattaac gtcaacgcat tcaggaaata 2700
 tttcacgctt taacaatccg gcatatgaca aagttctggc ccaggcatcg acggaaaata 2760
 ccgttaaagc gcgtaatgcc gattacaacg cggcagaaaa aatcctcatg gagcaagcac 2820
 cgattgcacc aatttatcaa tataccaatg gacgattaat caagccgtgg ctgaaagggt 2880
 atcccattaa taatcctgaa gatgtggcgt acagtcggac tatgtatatt gtgaagcatt 2940
 gatgtgatgg gaactggcgt tacccttggt cataacgcca gtgatgtctg atttagcgac 3000
 cctgttccgg cggtgtaatg ttatccatat acagcgtctg gctggggaag gcaaagtccg 3060
 cgccgtgtga ctgtacaata tcgataatct tcaaataaac gtcttgctgt gcagcaagcc 3120
 attcagcca taccgtgggt ttggtaaagc aataaaccat aatattcaat gaagagtcag 3180
 caaactgggt gaaataaacc agtaagggtt gtcgctggtc gatggccggg tgatttttca 3240

- 54 -

gcattctcacg tacagcttcg acaataacgc ccacttttgc cgcattcctca taacgtaaac 3300
caatgggtcgt ggtaatgcgg cggttgggtca ttcgtcctgg gttttcttacg ctgatcgacg 3360
aaaacagcga gttcgggtacg tacaatggac gattatcaaa ggtcgtaatt ttggtaattc 3420
gccagccaat ttccgctact gtaccttcga tattttctgtc cggatgaacgg atccagtcgc 3480
caatactgaa aggacgggtcg aaatagagca taatcccga aaagaagtta ctcagaatat 3540
ctttaccggc cataccgaca gccagaccac caataccacc aaagggtcagc aagccagaaa 3600
ggctcatgcc gaaatgttcg ccataaagca gaacaagcac cacaataatg gtgattttga 3660
tgatacgcgga cataatccgc gcaactgggtga tatcgcgacc ttttttaate tgctgttttt 3720
caaactgatt aatcagcaga aatagcttaa tcgtcagaat aaccgcaatc agggacgtac 3780
agataaaatc gataacgcct ggggtgataa atttgagttt atagttttct ataacataat 3840
taataatgct accaacagca ctgataatta tgggtgtagat taaaaattgc accgcatgga 3900
ataaaaatcc ttttctttta cgatttccac ggcgaacca aaagctcatc agaatcaatg 3960
ctgcgcagct accgaaaata atgaccagat taagcgcatt atttgtaaac agttcagcga 4020
tcattgtttt atcagggtcc tccagataat tgtcgtcatg ccggaacccc ctggcggggc 4080
tattttaccg cgacaattca ttcagatcat caatagtcag ggaaggaagt agcaacatta 4140
gctaaggaag gtgcgaacaa gtccctgata tgagatcatg tttgtcatct ggagccatag 4200
aacagggttc atcatgagtc atcaacttac cttcgccgac agtgaattca gcagtaagcg 4260
ccgtcagacc agaaaagaga ttttcttgtc ccgcatggag cagattctgc catggcaaaa 4320
catggtggaa gtcacgagc cgttttacc ccaaggctgg aatggccggc gaccttatcc 4380
gctggaaacc atgctacgca ttcactgcat gcagcattgg tacaacctga gcgatggcgc 4440
gatggaagat gctctgtacg aaatcgctc catgcgtctg tttgcccggg tatccctgga 4500
tagcgccttg ccggaccgca ccaccatcat gaatttccgc cacctgctgg agcagcatca 4560
actggcccgc caattgttca agaccatcaa tcgctggctg gccgaagcag gcgtcatgat 4620
gactcaaggc accttgggtc atgccaccat cattgaggca cccagctcga ccaagaacaa 4680
agagcagcaa cgcgatccgg agatgcatca gaccaagaaa ggcaatcagt ggcacttttg 4740
catgaaggcc cacattgggtg tcgatgcaa gagtggcctg acccacagcc tggtcaccac 4800
cgcgccaac gagcatgacc tcaatcagct gggtaatctg ctgcatggag aggagcaatt 4860
tgtctcagcc gatgccggct accaaggggc gccacagcgc gaggagctgg ccgaggtgga 4920
tgtggactgg ctgatcgccg agcgccccgg caaggtaaga accttgaaac agcatccacg 4980

- 55 -

caagaacaaa acggccatca acatcgaata catgaaagcc agcatccggg ccagggtgga 5040
 gcacccattt cgcacatca agcgacagtt cggcttcgtg aaagccagat acaaggggtt 5100
 gctgaaaaac gataaccaac tggcgatggt attcacgctg gccaacctgt ttcgggcgga 5160
 ccaaatagata cgtcagtggg agagatctca ctaaaaactg gggataacgc cttaaatggc 5220
 gaagaaacgg tctaaatagg ctgattcaag gcatttacgg gagaaaaaat cggctcaaac 5280
 atgaagaaat gaaatgactg agtcagccga gaagaatttc ccgcttatt cgcaccttcc 5340
 ctaactaatc aatgcgttga ttgtaaatcc agctaagagg tgaggttttc agagcagaca 5400
 acggtgaaat gtcattggtat tggtacgttt aggtacaag aaatttgtct gcacaaggat 5460
 tacatcatga ttatggcgaa actgaagtca gcgaaaggga agaaatttct ctttggtttg 5520
 ttggcggttt tcattattgc ggcgtcgggt gtgactcgcg cgaccatcg cggcggttata 5580
 gaacagtaca atattccgct gtctgagtgg acgacatcaa tgtatgtgat tcagtcacg 5640
 atgatttttg tttatagcct ggtctttact gtgttgctgg caatcccggt gggaatttat 5700
 ttccttggcg gcgaagagca gtaagtaaaa aataggcccg ataactcggg ccttgtcagt 5760
 tattgaagag tcgttaatcg tcttcttcgt catccagttc aacgggtgtc tgatactgg 5820
 caggtttaat gaccagcagg tcgcagcgaa gatgatcaat cacctgttcc gccgtgttgc 5880
 cgaggaatgc tgctgaaata ccggtgcgtc ctaccgtgcc cagaaccaca atccccgcct 5940
 gtaagtgtc cgccaaatca ggaatcacct cttctggcag acctttttct acgtgcgtca 6000
 tgttttcatt aatgccgaat ttctgccgca gggctttcat tgccagcaaa tggtgcccac 6060
 gaatggcatc gttataaacg ctcgggtcaa attccggcag ttcaatcgcg atattaattg 6120
 gcgttaccgg ataagcgcca accagatgaa cttcgggtat gttgacttgt tctgccagtt 6180
 cgatcgtctc tttgaccagt ttttcattga gcgcattatg atacggctct tcaactggcg 6240
 gattcacgc caccagcgcc ttgcctcctt ccggccacgg ctggtctttc accatccaca 6300
 ccgggcttgg gcatttgcgt aacagatgcc agtccgttgg cgtaaaaatc accgcttcca 6360
 gacggtcatg ttggtgcgcc attttttagca ccaaatacgtg tccgccgctg atcaattcct 6420
 gaatgatggc ttcgaaagga cggttatgcc agaccacttt aatttcaatg ggaacgccag 6480
 cattgagata atattttgcc tgctcgtgga tccaggctgt acgctggctg atgacgccct 6540
 gacgcatagc ggtacgttcg tccggggaga gcagggtggg catttcgtat gagaagtcat 6600
 agatcggcaa aaaggcttta attttgccac caatccgttg atgtaaataa acagctcgcc 6660
 gcaatgctgg ttggtcgtcc tggtaggat cgataacaac gagcatgttc tgatacatag 6720
 ccatacaggg tctccttaca acaactgtca acgcagtttg taattaaaag attaaccat 6780

- 56 -

atctggtgaa tgaacacagt atgaaccttc tgccagatca ataaatcaga aaaattttaat 6840
gatatgacag aaggatagtg agttatgcgg aaaaatcagg caacgttacg cgtatgacca 6900
gcaagctggg ccagcgcacg gttattttcg atggtgatgt atttaccttt gactgccagc 6960
atgccgcttt tctggaagcg acccagcaga cggctgatgg tttctaccgt caggcccaga 7020
tagttaccga tatcgccacg agtcacgcgc aggcggaatt cacgagggga gaagccgcgt 7080
tgggcaaaac gacgggacag gttgtagatg aatgcagcca gacgttcctc ggcatTTTTT 7140
ttcgacaaca gcaggatcat gtcctgatcg cctttgattt caccgctcat cagacgcacg 7200
atctgctgac gcagattcgg catTTTaccg gacaaatcgt ccagcgtttc gaacgggatt 7260
tcacatacca tcgaggtttc cagcgcctgc gcgaagctcg ggtgatggcc gctgccgatg 7320
gcgtcaaata ccaccaggtc gcctgctaaa tggaaccag tgatttgctc gtcgccttgc 7380
tcagtgatgg tataactttt aatcgtaccg gagcggatgg cataaagcga tTTaagttca 7440
tcaccagcct taaacagcgt ctggcctttc tgaataggct tcttccgctc aatgatatta 7500
tcaagctgat caagctcatg ttcgttgagt gtgaacggga tgcaaagctg gctgatgctg 7560
caatcctggc aatggatagc acaaccgcca gactgaatgc gccgtataat tcgctTTTTc 7620
gggatcatag gtctgctcaa gccgtaattg atatttgtca attTTaacat cTTTTtaggg 7680
agcaagtaag tctaagcaaa ccttaacagc agagaattcc gatattagat gtaaataatat 7740
gtctatctat ttgaaaacc ttaagttggt aagggttaact ttacataaaa gtgtgaacaa 7800
gctggcacia attgtTTaat gTTtacagca aaagataacc ttcattggcg aataaccact 7860
cttttcgctg aactccgcct gcatatccgg tcatgggtgc gtttcggcca ataaccgat 7920
ggcaaggtag gacgatgctg atgggattcg atccgtttgc cgcaccaacg gcacgcgccg 7980
cgccaggacg gcccaattgc tcagccagtt ggccgtaatg cattacctgc ccgcagggga 8040
tagtgcgtag tgTTTTccag acttcgcgct gaaatggcgt ccccccgta gcagtgggaa 8100
gcgtatcaat aatgctaaga ttaccggcaa aatattcacg aagcttgctg cttaaaccgc 8160
ctggattggt ggcagaaatg cgctcatagc cttctttgcg ataatggatg tccagcagct 8220
gcaccatgcg ttcgctgtac tcttcccatt caaccgcccg caggcgaaat tgctcatcgc 8280
aaatcaccca cagtggacc agtggcgtgg caattTTTTt tTcaagtaat ctacgcatcc 8340
gttctctctt aagacaaacg tgggtaaata cccggacctc tcggcagacc gacaagatac 8400
cacgccagca acatcagcag ccataccacc aaaaagataa gtggataggg caagactaac 8460
gaatagtaag taccagttt cgcgtctggt ttgtagcggt gcaggaatcc aagaaacagt 8520

- 57 -

ggaacaaaag gagataccgg cgctaaaggc aatacggatg agtcggcaat acgaaagagg 8580
at ttg cgcaa atg ccgggtg aaagccaagt agcataaaca ttggtacgaa aatggggggcg 8640
agaatcgacc agattgcgga accgctggca ataaacatgc ataagaaaga ggaaagcaac 8700
gccagaccga caaacgccgg gatgccgcta agccctgaac tttccagtat atcggtcagc 8760
cccacggcga tgaatttccc catgttgctc cagttaaaca tggcgacaaa ttgggcgagg 8820
ggaaaaacca tcacgataaa tcccgccatc tctttcatcg gttcaatcat taaatgcggt 8880
aaatccgcct gacgtcgaat tgtgcgggta gcgatgcat aagccagcga gacaacaaag 8940
aaaaaaagaa tgatcagtgg cacgatacct ttaataaagg gtgatggcat cacggtgtga 9000
ttaatcggat cgcgcaatat cccgttttgc gggatcacca tcagcgcaat cgcagcaata 9060
aaaagtagcg atacgacacc tgctatgcgt aaaccaaacc gctgactttc ggtcaatgtc 9120
tgcagtttct catcgctgtt tccctgccat tgacctaac gtggctcgat gattttgtcg 9180
gttatcaggc cgccaacaat cgtcagtacg actacggagc tggccataaa ataccagtta 9240
tcaattacac tgacgtgcat ttgcggattg aacgcagctg ccgcttccgt gctgatcccc 9300
gacagcaaca cgtcggttgt gacaatcagt aaattagccg taaagccgca acctacgcct 9360
gcaatggcag ccagtaaacc tgcaactgga tgcctgcca ccgccagaaa aatcagcgca 9420
cccatcgggtg gcatgatcac taacgccgca tcggaagaaa tgtggctgaa aaaagcaata 9480
aacagcacca tataactggc gtagcgggca ttaacatgcg atgccat ttt aaccattagt 9540
gctggcagta agccgacgcg ctccgccaga ccggcaccta aaaccagcgc caggatcgca 9600
ccaagtggag caaaaccgct aaagt ttttta ataacattgg gtaaaaacca gtgtaatcct 9660
tccacactga gcagg ttttt caccacgacc ggcgtaccat cggtcgggtt tttcgcactg 9720
acgccaaagg ccgacaaaat tgccgtcgtc accatgagta cgataatcaa atagataaag 9780
agcagaaaag gatggggaac cttgttacca attctttcga ccagccata gagcttcccg 9840
gattgggagg acgacggtat ggatgacata ctcatgggca ttcctcgggtt gttgtgtttg 9900
cgttggtgtt gttat tttta aggtgacggt gtcacgt ttt tcgggatagg gcagtgatac 9960
ggttgcgtgt ccgttacttg ctgatgctct tgttggcatt cttgcaatag ccctgaatca 10020
aggaagagat tgactgtgggt tgctgccata gttttcgccg ccagcagcat tcctttatga 10080
gcaatagatg ttcgcccctg gctaaccagt tgccacgtat gtagcgggtg accgacggca 10140
aaacaggggc tgaaacactg ggcaacaggc agtttccaac tgacgtcgcc gacatcagtc 10200
gatgccgcaa gcacgttatc ggtggcggca tatggagcga cttcatcgc cagtaccgtt 10260
tcacgatgac gtagtgcaaa aaccttgccg ttttcgccac cggttgcggc gatattatc 10320

- 58 -

agactgtttt gccgatcggt ggaggtgagc gtagcctgaa tttgtttcgc aaaagccagt 10380
tcttcggagt tccattccgg ggtacaaaaa tgggataggg cctggtacat ggcattttct 10440
aaggtgcat tctggagata actggaacag gctttgtcga agcggcattc aaccgtgggt 10500
tcggtcatca atgccgcacc ttcggcgatt ttggcgaccc gatcataaat atgctgcacg 10560
tcggtcattt cgggggagcg gataagataa agcacttctg cctgcgcctg gaccacgttg 10620
ggcgagatcc cgccgctatt tgtgatggca tagtgtacgc gcgctttttc aataatatgt 10680
tcgttgagga agttggtgcc agtgggtcatc aacgttacgg catcaagggc gctgcgtccc 10740
aaatgagggg aattcgcggc atgtgctgag atccctttaa agcgccatga tgcctgaatg 10800
tttgccagcg tgcgggtatt gaacataccg gcaaaggctt ccgggtgcca ggtgagtgcc 10860
gcatccacat catcaaatac cccctcgcga accatgaacg ttttaccga gccgccttct 10920
tcgccaggac aaccataaaa gcgcaccgtg ccgccttgcc catattgttc cagccatttc 10980
ttgacggcta ttgcagcggc aaaggcgagc gttcccagca aattgtgtcc gcaaccgtga 11040
ccattttcac cgggcgtcac ggatgtaggt tgccgcgaac ctgcttgctg acttaaacct 11100
gccagggcgt catattctcc cagcagggcg ataaccggtt tgccttgacc aaacgaagca 11160
ataaaggcat ttgggatatt gcctacgttg cgggtaacgg tgaagcctgc agattccagc 11220
gccgaagcca gatgctccgc tgaccagaac tcttcaaaac gtgtttctgg atgatcccag 11280
atttgatcgg caatatcggg ataaccgttg cgatcggctt caatcgcac gtcgataaaa 11340
cgatagatth cctgcatcag atacctcgcg tccagggaaa attgagcgcg gtgcgcgcca 11400
gcgttttcgac ggcaatagcg agaacctgct cgtcaaaatc gaatttttcg ttgtgatgac 11460
ctgcccgcag ctgtgtgcca aacaccacgt aggaggcttg cccttgatgt tgctgcacgc 11520
gggccatcat taatgtggca tcttcggaac ccgcaggcgc ttcaacacgt tcaatggcct 11580
gattgacccc cgcgacctga gccgcctgac tttgcaacca tgcgacctat tgcggcgaag 11640
gagaactggc ggtagctgca cccatcagac gagtttcaac gccgacacca tacatgggtg 11700
ctgcgccttg aatcgcttgt tgtgcacggg caaaaacata ttgattaatg acgtcgctgg 11760
ccccgcgtgt ttccactttc agcaacgccg aggcaggaac aacgttacga ccgcttcctg 11820
cctgcataac gccacggttt actctggaag ctccctcgtg gtgcggggcg attgcatgca 11880
gtgcaagagt ggcttgtgct gccgccaaca aggcattgtg accgtcttct ggttttgccg 11940
ctgcgtgagc ggcggtaccg gtgaagtgcg cgtcaaattt ggtgggtgcc ataaaattat 12000
cactgccgca caccacgggt cccgcaggta cgccagtgcc aatgtgcacg gcagtaaaat 12060

- 59 -

aatcaacatc atctacgaca cctgcatcga ccacgcgccg cgcgccacgc gtaccttcct 12120
ctgcaggctg aaaaatcagt ttgatgacgc catgtagtcc ggactcgaac tgtttaaggg 12180
tatgcgccag cccaagccca atggcggtat gtccatcatg accacaggca tgcattc 12240
cggcggtaca tgacgcaaaa ccgtcgcggt aggggcgatg gctgacatcc tgctcttcac 12300
tgagatccag cgcgtccata tcgacacgga aagccatcac cggaccgggg cgaccggtat 12360
ccagggtggc gacgatacca gtgaaaccac cttcaaaagc cgcaatccat tgtgctagcg 12420
caccctgttg acgagcgcgc tcgaattcgc gttgtagagt gaattcatca ggtaatccca 12480
tccggctact ttcattaact acttcgcgac ccagcgcgcag tgaatagccg agctggtgca 12540
attcttccgc aacaagggtg gcagtgcgga attccacca gccagactct gcatagtgat 12600
gaaaatcacg tcgccagtgc gataattttg gggcaagcga attaacaaat tgattcaaag 12660
actccataac ctttcccgtc atcagtaaaa agtgtgaccc ggttcacgta gcgatagttt 12720
ttacttatca ctaactgatt tttcacagtt ttaaccgttc ataaattacc ctgacacaat 12780
catctgcatt 12790

<210> 13

<211> 294

<212> PRT

<213> Escherichia coli

<220>

<223> complement of position 3733-4953 of seq id 12

<400> 13

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

Ser Arg Gly Arg Ser Gly Thr Ser Lys Asp Ser Gly Met Arg Ser His
20 25 30

Val Trp Gly Asn Val Lys Asp Thr Thr Gly Asp Arg Lys Val Val Arg
35 40 45

Met Ser Asp Ala Ser Tyr Ala Ser Met Ala Ala Asp Ala Gly Ser Ala
50 55 60

Tyr Asn Asn Arg Val Gly Ala Gly Ser Gly Gly Gly Ser Arg Val Gly
65 70 75 80

Ala Asp Ala Met Arg Gly Arg Gly Lys Ala Val Gly Tyr Val Val Thr
85 90 95

Lys Ala Met Ala Ser Gly Val Ser Ala Cys Ala Thr Lys His Gly Val
100 105 110

Asn Tyr Ser Ser Ser Ala Cys Ala Thr Ser Ala His Cys Gly Asn Ala
115 120 125

- 60 -

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

<210> 14

<211> 10244

<212> DNA

<213> Escherichia coli

<400> 14

aaaacatccc ttaaaaccaa tctaattctt ttcttctct gtgtatttgt tcccatatg 60
 gcgtcataat ttacgctacg taatacggga tagttacgat acgcagcgat agcgctaagt 120
 tttagttaaa atcccccatc agcgggatat ggggatctct gggagggttca ggtttttagca 180
 tcgcgacttt aatacgacag acgctgggtga cgatgatgcc aagtagtgac tcttccgaat 240
 aaggttccca ctgaccaatt ggcgcgcgct tcgcaataaa atctcccttc gcccagtatt 300
 gggaaaagta gataattca aactgtgtac gctgtttcgt ctcacagttc acgacattgc 360
 tggccgatga gctgaccatt gggcctatcc gggctgtccc gatggatatt acgccgaagt 420
 atgtgggaat tgccagcgga ttgatgaacg ccggttccgc tgcgccgac attatttcac 480
 ctatcgcctt tggcattatt atcgacaaaa caggcaactg gagtttacca ttctacggtt 540

- 61 -

ccgtcgcatt gcttggttatt ggtatatattcc tgacgtttctt catgcgtcca gataagtcatt 600
 tgtaaaagtc gaatcagggc tgaagtggca cactgaattt ggccacctga acagaggtga 660
 tatgctcacc tcagaacaac acaggtgctc caatgaaaaa aagaaatttc agcgcagagt 720
 ttaaacgcga atccgctcaa ctgggtgttg accagaaata cacgggtggca gatgccgcga 780
 aagctatgga tggtggcctt tccacaatga caagatgggt caaacaactg cgtgatgagc 840
 gtcagggcaa aacaccaaaa gcctctccga taacaccaga acaaatcgaa atacgtaagc 900
 tgaggaaaaa gctacaacgg attgaaatgg agaatgaat attaaaaagg ctactgtaga 960
 ttcaatctgt caatgcaaca cccctttcaa ttatctcttt cgggtgttttg aacttcagt 1020
 tctttctcgg tctgttggtt agctgagcag caaccagatc tagttcatgt tgagtatatt 1080
 gggcaagaca tgtcttttta ggaaagtact gccgaattag cccatttggtg ttctcatttg 1140
 ttccccgctg ccaaggactc tgaggatcgc agaagtaaac tttaacgccg gtgctgacag 1200
 taaattctag atgtctggcc agttccattc ctctgtccca tgtcagtgat tttctgagtt 1260
 ctgacggtaa actcaggaat ttgtcggtaa gagcctgatt tactgagaca gaatctttgc 1320
 ccctgagtct aaggatgatc gtataacgtg attttcggtc tacaagtgtg gctatatgag 1380
 agttttttgt acctgagact aaatcgccct cccaatgcc tagagagcgt ctgttatcga 1440
 tatttcggga acgttcgtga attgggtgtc cgttcactat gttaatcgta cctctttcgc 1500
 ctttgccgggt atgacgcctg ccatggcgaa ggctatgcga ccgtcgcaga tgctgtatat 1560
 tcaggtggtg tagcgcttca cggctacgaa agtacagcgt tttataaatt gtctcaggtg 1620
 atattcgcag cgttttttga cgtgggtttg ttgccttaa ccctcctgat atttgctctg 1680
 gagaccattt catctccagc ttttcagaa caagctttcg caatggtaaa ttttgatcca 1740
 gtaagcacgg ttttggcctt ttgcctattc tggtggctcg gttattagca tcaacagctt 1800
 tgtaatagcg tctgccccga ttacgctgaa cttcacgtga gatcgtcgaa ggactgcgat 1860
 tcagcgcagt agctatcgca cgaatgctca ttttggctga caaaccagct cgtatctcct 1920
 cgcgctcaga cagtgtcagg tgagctacag cccgcttacg ctcatgggggt tttatgccgc 1980
 cagtatccct taacatagtg aagatcgttc cgggttttga acccaggata ttgcgtatatt 2040
 cactgaagcc tggtccgttc ttccatagtt caaaaacaga ggctttttcc tctgctgtaa 2100
 atgttcgtct cattcaaaaa acctccgcaa ccccatgttt tcacataact gttgcgttga 2160
 ccaattgaat ctacagttgc tttttttaat atatctcgt caaggcgagc ttcatttaac 2220
 gccttacgca gttgcagaat ttcagattcc agttcagcca ccgtgcggga accaggagta 2280
 ccgagccctt ttctggcggc ggtaacccat tgcctaaag tgccttcagg aagagataat 2340

- 62 -

cgggaagcgc cttcactgat cgaaagttga ttttcaagaa ccgttctgac agcttcggct 2400
ttgaactttt tagagtaacg ttgggttttt ctgctcatta ttagctcctt ctgatgccat 2460
tctatttcag gaaggagtgt ccgttaaact caggctacct caggaaaaac cgtcctgaaa 2520
aaccagacgg cagacgggct gtattacgca gtcaggctact tgagctacat ggcatcagcc 2580
acggttcggc cggagcaaga agcatcgcca caatggcaac ccggagaggc taccagatgg 2640
gacgctggct tgctggcagg ctcatgaaag agctggggct ggtcagctgt cagcagccga 2700
ctcaccggta taaacgtggg ggtcatgaac atgttgctat ccctaactac cttgaaaggc 2760
agttcgccgt gaccgagcca aatcagggtgt ggtgcgggtga tgtgacctat atctggacgg 2820
gtaagcgctg ggcgtacctc gccgttggtc tcgacctgtt cgcaagaaaa ccagtgggct 2880
gggccatgtc gttctcgccg gacagcaggc ttaccatgaa agcactggaa atggcatggg 2940
aaaccctggg taagcccgctc ggggtgatgt tccaagcgat caaggcagtc attatacgag 3000
caggcagttc cggcagttac tgtggcgata ccggatcagg cagagtatga gtcggcgtgg 3060
aaactgctgg gataacagcc caatggagcg cttcttcagg agtctgaaga acgaatgggt 3120
gccagcgacg ggctatgtaa gcttcagcga tgcagctcac gcaataacgg actatatcgt 3180
tggatattac agcgcaactaa gaccgcaoga atataatggg gggttaccac caaacgaatc 3240
agaaaaccga tactggaaaa actctaacgc ggaggccagt tttagttgac cacaacagac 3300
tacctgaagg gagccgcggg cgcttggcag ttgcagtagc aggagatcat ccagccgcag 3360
tacaggtcac gatgactctg gttaatgata ccggctttga ccccgatatt tccggctcta 3420
tcgctgaatc atggcgctcag cagccgtgca caccatccta ttgttgtgac tgggaggctg 3480
ccaccatgct tcgcgctttc cctctggcga aaaagggaga aggacgggcc cgtctgcctt 3540
cactttatgc cagcttcggg aagctgggtg agacaccgac tcatgaagat atcattgata 3600
acaatcgatc catcaactgg cctgtataac gtggctgccg gtgattaaga aagctgcacc 3660
tacctaagta gtagcaaacg cacacttttt agaaaaatcg atggtcagaa actggattag 3720
caattccgtt ccatggttgc ttttgattta cgttggcgct tgatcattga tttatcctca 3780
aaagcccaac ctcatgggta atgaaccagc tccgtgaatg tccgctctgg cacagagcga 3840
aattttttga tctccccctt gaaatctaaa cttagtcatg tcacgttttt gggtttctaa 3900
aattttaact tcgcgttttt cgttgccgta agggttatac agaaatgtcc gttaagcaga 3960
gttcaaaatt gattgccgtg atcacgactg gtttgaaagc cgcgcccaag cctgtacagc 4020
tctggtttgc gttgattatg aacctgtcag cctaaagcaa gcggatggac gatgagtatt 4080

- 63 -

ggtaatcttt cagagtccgg aaaagttcag cccagtcctg aacaggcttg ctggcgccag 4140
tccagtttca ttcagtcgtg gtttggttct tacggcctgt gcaatctacc tcattaggca 4200
catcggcctg ccagataccg gctcgggggtg tatttccgct tccacgctga atactgttct 4260
cagcaatcct ggggtcatca cctcttctgg tgtgccttgc gccataacat gtccgtttgc 4320
cattaccacc agttgatcgc agtaccggct agcctgatta aggtcgtgca gcacagcgac 4380
caccgttttc ccctgagtcg ggagttcgcc catcaaccgc atcagggtcca cctgggtgatt 4440
gatatcaaga taggtggttg gctcatcaag taatacaacg ggcgtattct gggccaggac 4500
cattgccaga aatgcgcgct ggcgctgacc gccggaaagc tcggttaacc gacgaacggc 4560
aagatgattg atccgggtct ggttcatggc gacattaact cgtgcattgt cttcagcgga 4620
gagacgcccc cagagtgaca gccagggatt acgaccatac gaaaccagct cctggactgt 4680
gatcccctct ggcgttaaata ggtgctgagg cagcagcgaa agcctgcggg ccaactggcg 4740
cgatgagagc atatttatgg gattatcgcc gagaaatacg gtgccagact gcggcattaa 4800
aagccgcgaa aaacagttta acagcgtcga tttcccgcaa ccgttaggac cgatcagggc 4860
ggtgatcttc cccgttggca gtgagagtga aacgtcgtta agtaccttgt ctgtcccgtg 4920
actgaccgtc agattttcag ttcgtaaagt catttatcgc attctcacia gcaaccagac 4980
aaaccacggc gcaccgataa tggcggtcag cacgccaacc gggagctcca gtgggggatg 5040
aataattctc gccagcagat cggcaaccac caacagcaac gcacctgtca gggccgaaac 5100
aggcagcagt ctgcggtgac gtccaccggt gatgctacgc atcatatgcg gcaccacgag 5160
accaataaag ctaatcgggc cgcaggcggc cacgcccgtg gatgtcatgg cgacagctag 5220
taacaaagcc cagaatcggg tatggggcac cgacacaccg agcgtgggtg cgcgcgcatc 5280
gccgagtgca aggaggtcga gatcgcgga aaaactcagg ctacagcgga gaaataaaat 5340
catcagcggg atggcaatct tcacaaagct ccagtcacgg ccccataagc tgccggtcag 5400
ccacagcagg gcgttggttca catcctgcgg gcgcgagagc atcagataat ccgtcaggct 5460
ggcccagcat gcagaaagcg ccacgcccgt gagcgccagc ttcacggct ggtgggtctt 5520
tgccagcatc ttcagcaata tcaaccccgc catgccgccc gcaaaggcca gcagcggcag 5580
caccatcacg ggcagtgacg gcataagaag tagagcccc acagaggcca ggctggcggc 5640
atggttaaca ccgagaatat ccggtgatgc cagaggggtg cgcacaatcc cctgtatcag 5700
cacgcccgc acggcgaggg ctgcaccgac aaacagtgc agcagcaagc gcggcagtcg 5760
gtactccatc aatacataat aatgctcgtg tccggcctgc cagtcggtca gcagcgcgcg 5820
ccacggcacg gggatcactc ccatatggag tgataacagc gcacagcccc ccagggaag 5880

- 64 -

ggtgatgaaa ataaccagcg caattttcat cctcgctcc tcacaagcca gacaaagcaa 5940
gggctgccaa tcagcgccag cactgcgct gcgggcagat ctccggggaa ggccagcgcg 6000
cgtgcgagta catctgccag cagcatcagc gtggccccc gcagcatgct cactggcagt 6060
acgttgcgct gatcgaagcc tgcccagaag cgcgccagat gtggcaccag cagaccgata 6120
aacgccaccg gacctgctac gctgacgcac gcaccaacca gaagcagcac taacatatgt 6180
atgaccaaac gtagcctcgt caggttcact cccagcgtat gggcggtgct gtcgctgagg 6240
ttgagcaggt tcagttgatt cgccagcagc aacacgacag ggactgcagt gaccaccacc 6300
ggcaagagct gccagacatc ctgccagcgg gcgtgggaca ctccgcctgc cagccagtaa 6360
aagatgccgt aagcatgata ttcggccagc agcaggggtga tgcgggtaag gcccatacaa 6420
aaggccgaca gcgcgatacc cgcgaggatc agtttgtttc tgtcatgggt atgacgaaat 6480
ccgcctcctg cggatcatgac cagcagccag ctacgcccgc ccccgcatgc cgcatgaat 6540
gacagagaat agcctgcaat cggcgctcga ctacgcgcgc tggtaagcgc catagccagc 6600
gccgcgcccgc tgttaatgcc gagcagtga ggagaggcca ttgggttggt ggtcagggtt 6660
tgcagcagcg tgcccgcgag cgccaggctt gcgcgatca gaacggcgac caggcttcgt 6720
ggcaaacgaa ggttttgac cagcgcttct ggtagcgttg gcgtgtgtcc aggcagcagg 6780
gcgcgggttg catctgctcc ggaaacagga atggccgagt agcaaacag actcagccag 6840
aaaataataa taagtgtgc aacgggaagc cccacagca gcaccgggtg ttttatcgcg 6900
gtcatttcac aacggtaagc ggctgatgg ggaagatttt taccgtgtca gcggcaatac 6960
gctctgcagc aaaaataacc cgcatccgcg cccagggtgt actgtcgacc gaagcaacct 7020
gctgcttctg cgcgggcgtt aacatctgcc agagcggatc ttgttgccag cgtttaacaa 7080
tgctctcttc gcgatagtgg gcaaccagca gccaggcagg attgaccgcc agcagttgct 7140
ccaggccgat ggacggcatg gacgcacccg ccatcgagc gggaacgttc agccccagag 7200
aggccagcac gctgccggtc caggctctct gagtatgcag gttgaattgc tgttcgcggtg 7260
atgtgccaaa ggccacgcgt gtcccttttg gaagctggct ggccactgc gccatcctct 7320
ctttatgttg ttccagacgt gcctgcatct ctgcttttt acccaccatt tcgccgatga 7380
tagccgcaga ttgcaaattt tcagcgtagg ttctgttgcg ggacttaagc agcagtaccg 7440
gcgcgatttg ctgcaaggcg atgtaaacc cgcgatggcg actgctgtcg gcaatgatca 7500
ggcttggttt cagagcggca atggcttcca ggctcggctg cgcgcgcggt cgcacggact 7560
gccacgggtt cagggtgcga cgcacttcgg gcaggatgcg ttttgcatcg ttatcgtcgg 7620

- 65 -

caataccgat cgggatgacg tccacggcgg ccagcgcacg ggcgaacgag agttccagca 7680
ccacaatccg ttgtggcggt ttttcgagtg taaacgtgcc gtgttcgtcc tgaaccgtgg 7740
cggcaaaggc gtggctgac accagcagca ggcctgcaaa aagaaaacgg ataaatgcca 7800
acataatcac attccagcta aaagcccggc aagccggggc ttaacacacg agaacttcaa 7860
cgacccctgc atatacagcg tgcgcggctg gcctgcatag atgcctttgt tgttgtcgtc 7920
ataagagcgg atgaagtagt cctggtcgaa gatgtttttc acaccgaatg ccaggttcag 7980
atctgccacg tgcggggccaa agtcatacgc cacgcgtgcg cccagagca tgaagccggg 8040
aatgcggccg gtactgccgt cggcgccttc tttcaccgta ttggcgttat ccgcaaactg 8100
gctggactgg aaatcgctgt tcagattgaa cgtccagttt cctggcttgt agtccacgcc 8160
cagcgtgcct ttatgtttcg gggagaatgg taccagattg ccgtaggtgt cgcctttctc 8220
gcggatttcc gcgttcacat acgcatagct ggcgtagatg gaaacgttat caagcgttgg 8280
cgttagcgta cccagatcgt aacgtgcctg cgtttccagc ccggtatggc gcgttttgcc 8340
acgtgcagtg acggtgtcgt tggctctggt ggagtcgtac tgattgttaa agttaatcag 8400
gaacagcccc atttccgccg tcagcgcgcc gtcgtcgtag cgggtaccga gttcccagg 8460
tcgcgccttt tccggttcaa cattgccgct ttgcacagcc ttgccaatct ggctgtactg 8520
tacggtgccg aacgagcctt cagtgtttgc ataaagattc cagctgtcag tcagggtgata 8580
gagcacgttc aacgccggaa gcggtgcgtt atagctcact tcttcgtgcg tgctgtgat 8640
ggcgttggtc tggtagtact cgatatgttc gaaacgcata cccggcgtga tgggtccagtt 8700
gccgatgtcg attttgtcat ccagatacca ggcgtgcgcc tcggtgccgg aacgcgtatc 8760
gcggtcgtaa ggccttgagc cggacggcaa ctgcccgtg ctggtggcgg tgtagtaacg 8820
catttcatgc gttgattcat tcaaatacgc atagcccacg cccacttcgt gcgcggaagg 8880
gccgatcata aagatctggc tgtagcgtgg ctcaataccg cgcacccagt agttacgcgg 8940
cgagaggggt atgcgtttgc cttgctccag gtagccgctg cgcagggttt ggggtgtagaa 9000
cccctgaatg ttgaatttat gctggctgtc tggctggaac tggtagccca ggctcgccag 9060
cttgcgacga cccagaagc ggtcatacgg gcgggtggat tgccagcgat cggcgtcgta 9120
atccgcgcga gacaggccac cgggcatgtc ggcttcaccg tcgtaatat gcagcaggct 9180
gttgaagggt tgcacctcat ccggcgcata tttgcttttc agcatcagggt cgtcgatgcg 9240
ggtggcgtg tgctcgcgcc agtcaactgcc gcgcgtgccg gagtagagca gcgcggtgcc 9300
aaaaccgttg tccgctgtgc cggccaccat caggttgtgc gtctctttcg gggtgttttg 9360
tgaagagggt gggctgagct gaccttcac gccgcctcg ataccaaagt cctgcggaat 9420

- 66 -

ggcacgggta acaaagttca ccacgccgcc cacgctctgc ggtccgtaac gcaccgcacc 9480
 accaccgcgt accacgtcaa tggcatccat gttgccgagc gaaacgggag ccagtgaaag 9540
 ctgcggctga ccgtaagggg cgaaggggac ggggatgccg tccatcagga cggtcgagcg 9600
 gctggcgagg cgcgggttca ggccccgat gccaaagtgc atcgccaggt cgtggctgcc 9660
 ggtgccgttg ttttccggcg cgctgacgcc agggatgcgg ttaagtacct cacgcatggg 9720
 ggttgccgcg gttttggcga aatcctcacg gcggatcacg tcacgcgcgc cagcatgttc 9780
 aaatacgtcg ttttcacgcg catcaccag ccagtcgccg accacgggtca gggcatcttc 9840
 ttttggtgcg ggcgcgggct ccagcgtcca gctgttatTTT cccagcgggtt ttacctgcag 9900
 tccgctgccg tccagcagtt gttgcaggcc gctctcgacg tcgtaatcgc cgtgcaggcc 9960
 gttgctctgc ttgccgcgcg tcaggctggc gtcaaccgag agggtaaate cgctgtgtgc 10020
 ggcatactga ttgagcgctt tatcgagcga tcccggtgcg atattaacct gtgcagcaaa 10080
 agcggaaaac gagagaccgg ccagcggcag caggctcagg cgaatgggtg taaccaaagg 10140
 tgttgtttta cgaaaaacgc gtaacggcgt cataccttcc ccatcatcat ttttgttgtg 10200
 ttcagctatg agtcgaacga gaagggtaaa aaggacaatc gaaa 10244

<210> 15

<211> 774

<212> PRT

<213> Escherichia coli

<220>

<223> complement of position 7849-10173 of seq id 14

<400> 15

Met	Thr	Pro	Leu	Arg	Val	Phe	Arg	Lys	Thr	Thr	Pro	Leu	Val	Asn	Thr
1				5					10					15	

Ile	Arg	Leu	Ser	Leu	Leu	Pro	Leu	Ala	Gly	Leu	Ser	Phe	Ser	Ala	Phe
		20						25					30		

Ala	Ala	Gln	Val	Asn	Ile	Ala	Pro	Gly	Ser	Leu	Asp	Lys	Ala	Leu	Asn
		35					40					45			

Gln	Tyr	Ala	Ala	His	Ser	Gly	Phe	Thr	Leu	Ser	Val	Asp	Ala	Ser	Leu
	50					55					60				

Thr	Arg	Gly	Lys	Gln	Ser	Asn	Gly	Leu	His	Gly	Asp	Tyr	Asp	Val	Glu
65					70					75					80

Ser	Gly	Leu	Gln	Gln	Leu	Leu	Asp	Gly	Ser	Gly	Leu	Gln	Val	Lys	Pro
			85						90					95	

Leu	Gly	Asn	Asn	Ser	Trp	Thr	Leu	Glu	Pro	Ala	Pro	Ala	Pro	Lys	Glu
		100						105						110	

- 67 -

Asp	Ala	Leu	Thr	Val	Val	Gly	Asp	Trp	Leu	Gly	Asp	Ala	Arg	Glu	Asn
		115					120					125			
Asp	Val	Phe	Glu	His	Ala	Gly	Ala	Arg	Asp	Val	Ile	Arg	Arg	Glu	Asp
	130					135					140				
Phe	Ala	Lys	Thr	Gly	Ala	Thr	Thr	Met	Arg	Glu	Val	Leu	Asn	Arg	Ile
145					150					155					160
Pro	Gly	Val	Ser	Ala	Pro	Glu	Asn	Asn	Gly	Thr	Gly	Ser	His	Asp	Leu
				165					170					175	
Ala	Met	Asn	Phe	Gly	Ile	Arg	Gly	Leu	Asn	Pro	Arg	Leu	Ala	Ser	Arg
			180					185					190		
Ser	Thr	Val	Leu	Met	Asp	Gly	Ile	Pro	Val	Pro	Phe	Ala	Pro	Tyr	Gly
		195					200					205			
Gln	Pro	Gln	Leu	Ser	Leu	Ala	Pro	Val	Ser	Leu	Gly	Asn	Met	Asp	Ala
	210					215					220				
Ile	Asp	Val	Val	Arg	Gly	Gly	Gly	Ala	Val	Arg	Tyr	Gly	Pro	Gln	Ser
225					230					235					240
Val	Gly	Gly	Val	Val	Asn	Phe	Val	Thr	Arg	Ala	Ile	Pro	Gln	Asp	Phe
				245					250					255	
Gly	Ile	Glu	Ala	Gly	Val	Glu	Gly	Gln	Leu	Ser	Pro	Thr	Ser	Ser	Gln
			260					265					270		
Asn	Asn	Pro	Lys	Glu	Thr	His	Asn	Leu	Met	Val	Gly	Gly	Thr	Ala	Asp
		275					280					285			
Asn	Gly	Phe	Gly	Thr	Ala	Leu	Leu	Tyr	Ser	Gly	Thr	Arg	Gly	Ser	Asp
	290					295					300				
Trp	Arg	Glu	His	Ser	Ala	Thr	Arg	Ile	Asp	Asp	Leu	Met	Leu	Lys	Ser
305					310					315					320
Lys	Tyr	Ala	Pro	Asp	Glu	Val	His	Thr	Phe	Asn	Ser	Leu	Leu	Gln	Tyr
				325					330					335	
Tyr	Asp	Gly	Glu	Ala	Asp	Met	Pro	Gly	Gly	Leu	Ser	Arg	Ala	Asp	Tyr
			340					345					350		
Asp	Ala	Asp	Arg	Trp	Gln	Ser	Thr	Arg	Pro	Tyr	Asp	Arg	Phe	Trp	Gly
		355					360					365			
Arg	Arg	Lys	Leu	Ala	Ser	Leu	Gly	Tyr	Gln	Phe	Gln	Pro	Asp	Ser	Gln
						375					380				
His	Lys	Phe	Asn	Ile	Gln	Gly	Phe	Tyr	Thr	Gln	Thr	Leu	Arg	Ser	Gly
385					390					395					400
Tyr	Leu	Glu	Gln	Gly	Lys	Arg	Ile	Thr	Leu	Ser	Pro	Arg	Asn	Tyr	Trp
				405					410					415	
Val	Arg	Gly	Ile	Glu	Pro	Arg	Tyr	Ser	Gln	Ile	Phe	Met	Ile	Gly	Pro

WO 01/70776

PCT/US01/07478

- 68 -

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

WO 01/70776

PCT/US01/07478

- 69 -

Gly Val Lys Asn Ile Phe Asp Gln Asp Tyr Phe Ile Arg Ser Tyr Asp
 740 745 750

Asp Asn Asn Lys Gly Ile Tyr Ala Gly Gln Pro Arg Thr Leu Tyr Met
 755 760 765

Gln Gly Ser Leu Lys Phe
 770

<210> 16

<211> 10183

<212> DNA

<213> Escherichia coli

<400> 16

ttaacttttat gacatcatatc tgctttttaga agtgaaaaat taaaagggag agactccgct 60
 ctcccattat tggctatttt gcagggttac tgcgtggtac cgtcggtttt ggtatcgaca 120
 tcattattga tgccatcacc ggtttgtacc tttttattga tatccggaca gcgaccatct 180
 ttgcacatgg tgttcttgtg ctcttcatct ttggtcattc cgtcattggt cattgaagaa 240
 ccatccgaat gcagcattgt gccgccagaa ccggtattta ccccgttatt gtcgacgtta 300
 tttggcgoga cattttcacg ggcgtcaggg gctacctggc ccgcatcagc tgcggcggtt 360
 gcctggccgt tattagtttg cgctccgcta tcggcagcca gtgcggcacc gctggcaagg 420
 cttagagtgg cagtcagaaa taatgtggcc agttttgtca ttttcatagg atgctcctgt 480
 tatggtcggt atgtcggata acctcttcca acagtgcatt tgcaggtgaa tataaggcat 540
 tggtttaaga tttcagccag gttatgaaac gcagcagaga atcttgaaat aattaacaaa 600
 caaaggagtt acagttagaa attgtaggag agatctcggt tttcgcgaca atctggcggt 660
 tttcttgcta attccaggat taatccgttc atagtgtaaa accccgttta cacattctga 720
 cggaagatat agattggaag tattgcattc actaagataa gtatggcaac actggaacag 780
 acatgaatta tcagaacgac gattttacgca tcaaagaaat caaagagtta cttcctcctg 840
 tcgcattgct ggaaaaattc cccgctactg aaaatgccgc gaatacgggt gcccatgccc 900
 gaaaagcgat ccataagatc ctgaaaggta atgatgatcg cctgttggtt gtgattggcc 960
 catgctcaat tcatgatcct gtcgcggcaa aagagtatgc cactcgcttg ctggcgctgc 1020
 gtgaagagct gaaagatgag ctggaaatcg taatgcgcgt ctattttgaa aagccgcgta 1080
 ccacgggtggg ctggaaaggg ctgattaacg atccgcatat ggataatagc ttccagatca 1140
 acgacggtct gcgtatagcc cgtaaattgc tgcttgatat taacgacagc ggtctgccag 1200
 cggcagggtga gtttctcgat atgatcacc cacaatatct cgctgacctg atgagctggg 1260
 gcgcaattgg cgcacgtacc accgaatcgc aggtgcaccg cgaactggca tcagggcttt 1320

- 70 -

cttgtccggt cggcttcaaa aatggcaccg acggtacgat taaagtggct atcgatgcca 1380
ttaatgccgc cgggtgcgccg cactgcttcc tgtccgtaac gaaatggggg cattcggcga 1440
ttgtgaatac cagcggtaac ggcgattgcc atatcattct gcgcggcggg aaagagccta 1500
actacagcgc gaagcacgtt gctgaagtga aagaagggct gaacaaagca ggcctgccag 1560
cacaggtgat gatcgatttc agccatgcta actcgtccaa acaattcaaa aagcagatgg 1620
atgtttgtgc tgacgtttgc cagcagattg ccggtggcga aaaggccatt attggcgtga 1680
tggtggaaag ccatctggtg gaaggcaatc agagcctcga gagcggggag ccgctggcct 1740
acggtaagag catcacgat gcctgcatcg gctgggaaga taccgatgct ctgttacgtc 1800
aactggcgaa tgcagtaaaa gcgcgtcgcg ggtaaggttt aattgtcgga tgcgccgtca 1860
gagtggcgta tccgatgaat caccacaggc ctgataagtc gcgcagcgtc gcatcaggca 1920
atgtgctcca ttgttagcaa caaaaaagcc gactcacttg cagtcggctt tctcatTTTA 1980
aacgaatgac gtttacttcg ctttaccctg gtttgcaacc gccgctgctt tcgctgcgat 2040
ctcgtcagca ttaccagat aatagcgttt cagcggtttg aaattctcgt cgaactcata 2100
caccagcggc acgccagtcg ggatattaag ctcaagaatc tcttcttcgc tcatgttata 2160
aagatatTtc accagcgcac gtaaagagtt accgtgtgca gcgatgatca cgcgctcacc 2220
gctcttcata cgcggcagaa tagtttcatt ccagtaaggg atcacgcggg caatggtcag 2280
cgccaggctt tccgtcagcg gcagttcttt ctcgctcagt ttccgcgtaac gcggatcgtg 2340
acccggataa cgctcatcat ctttagtcag ttccggcgga gtcactgcaa aaccacgacg 2400
ccactgtTtc acctgctcgt cgccatactt ttcagcagtt tccgctttgt tcagaccctg 2460
caacgcaccg tagtgacgtt cgttcagttt ccaggatttc tcaacgggca gccatgcctg 2520
atccagttcg tccagcacat tccacagggt atggatagcg cgtttcagca cagaagtgta 2580
agcaaagtca aagctgtaac cttcctcttt cagcagctta cctgctgctt ttgcttcgct 2640
tacgcctTtc tcagacagat ccacgtcgta ccaaccgggtg aaacggTTTT ctttgttcca 2700
ctgactTtcg ccatgacgaa ccagaaccag cttagttaca gccatatact tactcctcaa 2760
atcatctTtt aatgataata attctcatta tattgccgcg acgaagcaac agcaatgctt 2820
acgcataacc atagcgaaaa tagtggcgca gtgtaagggt gttgtgaata ttgagttgca 2880
aatatgtcgg tgtttgctgg tgatttgaac aatatgagat aaagccctca tgacgagggc 2940
gtaacattac tcagcaataa actgatattc cgtcaggctg gaatactctt cgccaggacg 3000
caggaagcag tccggttgcg gccattcagg gtggttcggg ctgtccggta gaaactcgct 3060

- 71 -

ttccagagcc agcccttgcc agtcggcgta aggttcgggt ccccgcgacg gtgtgccgcc 3120
gaggaagttg ccggagtaga attgcagagc cggagcgggt gtgtagacct tcagctgcaa 3180
tttttcatct gctgaccaga catgcgccgc cactttcttg ccatcgctt tggcctgtaa 3240
caagaatgcg tgatcgtaac ctttcacttt gcgctgatcg tcgtcggcaa gaaactcact 3300
ggcgatgatt ttggcgctgc ggaaatcaaa agacgttccg gcgacagatt tcaggccgtc 3360
gtgcggaatg ccgccttcat caaccggcag atattcgctc gccagaatct gcaacttggt 3420
attgcgcacg tcagactgct cgccgtcaag attgaaatag acgtgattag tcatattcac 3480
cgggcaaggt ttatcaactg tggcgcgata agtaatggag atacggttat cgtcggtcag 3540
acgatattgc accgtcgcgc cgagattacc cgggaagccc tgatcaccat catctgaact 3600
cagggcaaac agcacctgac gatcgttctg gttcacaatc tgccagcgac gtttgtcgaa 3660
cccttccggc ccgccgtgca gctgggtaac gccctgactt ggcgaaagcg tcacggtttc 3720
accgtcaaag gtataacggc tattggcgat acggttggca taacgaccaa tagaggcccc 3780
cagaaacgcg gcctgatcct gatagcattc cgggctggca cagccgagca gcgcctcgcg 3840
gacgctgcca tcggaaagcg gaatacgggc ggaaagtaaa gtgcacccc agtccatcag 3900
cgtgactacc atccctgcgt tgttacgcaa agttaacagt cggtaagggt gaccatcggg 3960
tgccagtgcg ggagtttcgt tcagcactgt cctgctcctt gtgatggttt acaaacgtaa 4020
aaagtctctt taatacctgt ttttgcttca tattgttcag cgacagcttg ctgtacggca 4080
ggcaccagct cttccgggat cagcgcgacg atacagccgc caaatccgcc gccggtcatt 4140
cgtacgccac ctttgtcgcc aatcacagct ttgacgattt ctaccagagt gtcaatttgc 4200
ggcacgggtga tttcgaaatc atcgcgcata gaggcattgag actccgccat caactcgccc 4260
atacgtttca ggtcgccttg ctccagcgcg ctggcagctt caacgggtgcg ggcgttttca 4320
gtcagtatat gacgcacgcg ttttgccacg atcgggtcca gttcatgcgc aacagcgctg 4380
aactcttcaa tgggtgacatc acgcagggtt ggctgctgga agaaacgcgc accggtttcg 4440
cactgttcac gacgggtggt gtattcgctg ccaaccaggg tacgtttgaa gttactgttg 4500
atgatgacga cagccacacc tttgggcatg gaaactgctt tgggtcccag tgagcggcaa 4560
tcgatcagca aggcattgat tttcttgccg agcgcggaat ttagctgatc catgatcccg 4620
cagttacagc ctacaaactg gttttctgct tcctgaccgt taagcgcgat ttgtgcgccg 4680
tccagcggca gatgataaag ctgctgcaat acggttccga ccgcgacttc cagtgaagcg 4740
gaagaactta acccggcacc ctgcggcaca ttgccgctga tcaccatgtc cacgccgccg 4800
aagctgttgt tacgcagttg cagatgtttc accacgccac gaacgtagtt agccattga 4860

- 72 -

tagttttcat gtgcgacaat gggcgcatcg agggaaaact cgtcgagctg attttcataa 4920
tcggctgcca tcacgcgaac ttacgggtca tcgcgtggtg cacaactgat cacggtttga 4980
taatcaatcg cgcagggcag aacgaaaccg tcgttgtagt cgggtgtgttc accaatcaaa 5040
ttcacgcggc caggcgctg aatgggtgtga gtggcagggt agccaaatgc gttggcaaac 5100
agagattgtg ttttttcttt cagactcatt tcttacactc cggattcgcg aaaatggata 5160
tcgctgactg cgcgcaaacg ctctgctgcc tgttctgcgg tcaggtctcg ctgggtctct 5220
gccagcattt cataaccaac cataaattta cgtacgggtg cggagcgcag cagaggcgga 5280
taaaagtgcg cgtgcagctg ccagtgttga ttctcttcgc cattaaatgg cgcgccgtgc 5340
cagcccatag agtaggggaa ggagcactgg aagaggttgt cataacgact ggtcagcttt 5400
ttcaacgcca gcgccagatc gctgcgctgg gcgtcgggtca aatcgggtgat ccgtaaaacg 5460
tgggcttttg gcagcagtag cgtttcgaac ggccaggcag ccagtaagg cacgacggct 5520
aaccagtgtt cggtttcgac aacgggtacgg ctaccgtctg ccagctcgcg ctgaacataa 5580
tccaccagca ttggtgattt ctgttcggca aaatatctt tttgcaggcg gtcttcgcgc 5640
tcagcttcgt taggcaggaa gctatttgcc caaatctgac cgtgcggatg cgggttagag 5700
cagcccatcg ccgcgccttt gttttcaaaa acctgcaccc atgggtacgt tttcccagt 5760
tctgcggttt gctcctgcca ggttttgacg atttccgtca atgctgcaac gctgagctct 5820
ggcagcgttt tactgtgatc cggtgaaaag cagatcaccc ggctgggtgcc gcgcgcgctc 5880
tggcaacgca tcagcggatc gtgactttct ggcgcatctg gcgtgtcaga catcaaagcc 5940
gcaaagtcac tagtgaaaac gtaagtcccg gtgtaatcgg ggtttttatc gcctgtcacc 6000
cgcacattac ctgcgcagag gaagcaatct ggatcgtgcg caggtaacac ctgtttggct 6060
ggcgtttcct gcgccccctg ccaggggcgc ttagegcgggt gcggtgaaac cagaatccat 6120
tgcccgggtga gcgggttgta gcggcgatgt ggatgatcaa cgggattaaa ttgcgtcatg 6180
gtcgttcctt aatcgggata tccctgtgga tggcgtgact gccagtgcc ggtgtcctgc 6240
gccatttcat cgagtgtgcg cgttacgcgc cagttcagtt cacggtcggc tttgctggcg 6300
tccgcccagt aggccggaag gtcgccctcg cgacgcgggtg caaaatgata attaaccggt 6360
ttgccgcagg ctttgctgaa ggcattaacc acgtccagca cgctgttgcc tacgccagcg 6420
ccgaggttgt agatgtgtac gcctggcttg ttcgccagtt tttccatcgc cacgacgtga 6480
ccgtccgcca gatccattac gtggatgtaa tcgcgtacgc cagtaccatc ttcggtcgga 6540
taatcgttac caaaaatcgc cagcgagtcg cgacggccta cagcaacctg ggcgatgtat 6600

- 73 -

ggcatcaggt tattcggaat gccttgcgga tcttcgcca tatcgcccga cggatgcgcg 6660
ccaaccgggt tgaagtagcg cagcagggca atgctccagt ccggctgggc tttttgcaga 6720
tcggtgagga tctgttcac catcagcttg cttttgccgt aagggctttg cgggtgtccg 6780
gtcgggaagc tttcaacgta tggaattttg ggctgatcgc cataaacggt ggcggaggag 6840
ctaaaaataa agtttttgac gtttagcggcg cgcattggcg taatcaggcg cagagtgccg 6900
ttgacattgt tgtcgtaata ttccagcgggt ttttgtaccg attcgcccac ggctttcagc 6960
ccggcgaagt ggatcacggt gtcgatagcg tgatcgtgca ggatctcgggt catcaacgct 7020
tcgttacgaa tatcgcttc aacaaacggt ggatgtttgc cgcctaaacg ctcgataaca 7080
ggcagtacgc tgcgcttact gttacagagg ttatcaagaa tgatgacatc atgaccgttt 7140
tgcagtaatt gcacacaggt atgacttcca atgtaaccgc taccaccggt aaccagaact 7200
ctcataattc gctccattag gcttatggta tgaaataacc atagcataac aaagatgcga 7260
aaagtgtgac atggaataaa ttagtggaat cgtttacaca agaatttagc cgttttttat 7320
gcgcgattaa gtgattataa aacagagggt ttatgaatga ttgcgctttt tatctgaaaa 7380
aagacgcggt ttcattgctg catgcgtcga accgttggcc ggagagggtg ctaaggccgc 7440
ctccggcaag gtcagcacta ccgactcaat atatttttgt cagcacatag cgatagagtc 7500
caccgtccgg cacgaactca agacgatggg taatacaggc aggcgcattc tcagcgtggt 7560
gcgaaacaaa caacaattgc gtttcacctt cgctaatacag cacatcaaca aaacggcgga 7620
taagctggcg attcagcggg tcaagcccct gtagtggttc atcgagaata agcaacgtcg 7680
gatgtttcac cagtgcgcgg acaatcagcg ccagacgctg ctgtccccag gaaagactat 7740
ggaacggagc gtcagccgtg cgtttatcaa tgccgagaat atccagccac tgctgcacca 7800
gtttttgctg gcgatccgaa acggcctgat aaatgccaat cgaatcaaaa tagccagaaa 7860
gaatcacatt acgcacggta gtgctgacc ccgtaatccag atgcaacta ctgctgacgt 7920
aaccgatatg ctttttgata tcccagatgg ttctgccgct gccgcgacgt cgtccgaaaa 7980
gcgtcaaata gttgctgtaa ccttgcggat gatcgccagt aaccaggctt aataacgtcg 8040
attttcctgc accatttggc ccgacaattt gccagtgttc gcctggattc acctgccagc 8100
taaggttatt aagaatgggg cgatcgttat aagaaaccac gccattgttc agcacaatgc 8160
gcggttcggt ggcgggtaag gcgtgacgtg ctgaagggtc atccggctcc ggcagttgca 8220
caccttcaag ctgttcacta tgccgaggtt gcgcgacgag tgcttggttg agcagttcct 8280
ctttagcgcc agtttccgct aacgtgcaat ccgccagcac gccagcaaac tggacaaact 8340
ccgggatctc atcgaagcga ttgagcacca gtaccagagt aataccggac tgatgtaacg 8400

- 74 -

aggcgagtcg ctcagccagc tgctgacgtg aggcaacatc caggccatcg aacggctcat 8460
caagaatcaa caagtcaggc tccgacatca gcgcctgaca cagcagggtt ttctgcgtct 8520
cgccagtggg aaggtattta aagcgtcggc cgaggagggc ggtaataccg aactgctgcg 8580
ccagttgcat gcaacgcggc gcatccttta cttcatcctg aatgatctca gccgtagtgc 8640
gtccggtgtc atcttcgcca gggccgagca tatcgggtgtt attccgctgc cattcgctgc 8700
tgacgagttt ttgcaattgc tcgaaggaga gacgagtgat gtgggaaaac tggctttgcc 8760
gttcaccttt caaaagcggg agttcccccg ccagcgcgcg ggccagggcc gatttccgcg 8820
ttccattcga accgacaaac gcccaactat caccgcggtt taacgttagc tgaggcaatt 8880
gcagcgtttt tgtgtcgcta agacgaaacg tgccttgcaa aatttgcaac gatgacattt 8940
tatatccac tttgtgcagc gattactgac aggatacgt gtttcataac aaattgtcaa 9000
cacgcttagc acagcgtggc gataatcacg ctgtcggcat taaagtaggc cgtgacattc 9060
tgtccttgct gaagagaagt cgcttcattt accggcactg tggcgcacag tgtttgcccg 9120
tcgggtagcg ccattaatac ttgcgactgc tctgcgcgcg gctcaatatg actaataata 9180
cccggttaatt ggttgtcagc gttttgcgcg accgcctcgt cctgagtaat acctaccac 9240
ggcgccttta gcaatatcaa cacttctttg ccttcatcca gcccagacg cgcgccgctt 9300
tgtgcggtaa ttgcgacttt caggcgtgtt tttccgtcag ccagtaagac atcaacatgc 9360
tgttgaacgt catcatgatc gcgggcgggtg atggtaccga accactgggtt acgggcgctg 9420
gtttgcagtg aaaaacgtga gatcgcggcc agcaggctgt tcagcggcag ggcgtcatcg 9480
tcacttaaca catcaaaggc tttttgctgg atttgcgcca gtaagtcata gagctgaatc 9540
agtcgctgac catagcgggt cagtactgcg ccgcccgcac ctttaccgcc tgttgcgcgc 9600
tcgaccagaa tatgctcact taactgattc atctcggtta tggcatccca ggcgccttta 9660
tagctaatac cggcatcttt cgctccctgg ctaatggaac cggaaagcgc aatgtgtttt 9720
agtagcgaag tgccggcgcg gtcggcgaat aatttttggt ggagcttaag ggtgagaagg 9780
atttcggcct gcataacaat gtcctggcaa aagtcttatt gtgacggaag acgaacgcca 9840
cgcaaagctg accgcacaaa aggggagtg c ttttctgtgc ttagcgggtta gaatagtctc 9900
atgactatat ctggagttga ccatgttaga gttattaaaa agtctggtat tcgccgtaat 9960
catggtacct gtcgtgatgg ccatcatcct gggctctgatt tacggctctg gtgaagtatt 10020
caacatcttt tctggtgttg gtaaaaaaga ccagcccggg caaaatcatt gattccctga 10080
atgcccgctt agtcgggcat tttctttttc tcaacttcct gcttttcctg ccgatatttt 10140


```

<400> 17
Met Ser Leu Lys Glu Lys Thr Gln Ser Leu Phe Ala Asn Ala Phe Gly
  1          5          10          15
Tyr Pro Ala Thr His Thr Ile Gln Ala Pro Gly Arg Val Asn Leu Ile
          20          25          30
Gly Glu His Thr Asp Tyr Asn Asp Gly Phe Val Leu Pro Cys Ala Ile
          35          40          45
Asp Tyr Gln Thr Val Ile Ser Cys Ala Pro Arg Asp Asp Arg Lys Val
  50          55          60
Arg Val Met Ala Ala Asp Tyr Glu Asn Gln Leu Asp Glu Phe Ser Leu
  65          70          75          80
Asp Ala Pro Ile Val Ala His Glu Asn Tyr Gln Trp Ala Asn Tyr Val
          85          90          95
Arg Gly Val Val Lys His Leu Gln Leu Arg Asn Asn Ser Phe Gly Gly
          100          105          110
Val Asp Met Val Ile Ser Gly Asn Val Pro Gln Gly Ala Gly Leu Ser
  115          120          125
Ser Ser Ala Ser Leu Glu Val Ala Val Gly Thr Val Leu Gln Gln Leu
  130          135          140
Tyr His Leu Pro Leu Asp Gly Ala Gln Ile Ala Leu Asn Gly Gln Glu
  145          150          155          160
Ala Glu Asn Gln Phe Val Gly Cys Asn Cys Gly Ile Met Asp Gln Leu
          165          170          175
Ile Ser Ala Leu Gly Lys Lys Asp His Ala Leu Leu Ile Asp Cys Arg
          180          185          190
Ser Leu Gly Thr Lys Ala Val Ser Met Pro Lys Gly Val Ala Val Val
  195          200          205
Ile Ile Asn Ser Asn Phe Lys Arg Thr Leu Val Gly Ser Glu Tyr Asn
  210          215          220
Thr Arg Arg Glu Gln Cys Glu Thr Gly Ala Arg Phe Phe Gln Gln Pro
  225          230          235          240
Ala Leu Arg Asp Val Thr Ile Glu Glu Phe Asn Ala Val Ala His Glu
          245          250          255

```


- 76 -

Leu Asp Pro Ile Val Ala Lys Arg Val Arg His Ile Leu Thr Glu Asn
 260 265 270
 Ala Arg Thr Val Glu Ala Ala Ser Ala Leu Glu Gln Gly Asp Leu Lys
 275 280 285
 Arg Met Gly Glu Leu Met Ala Glu Ser His Ala Ser Met Arg Asp Asp
 290 295 300
 Phe Glu Ile Thr Val Pro Gln Ile Asp Thr Leu Val Glu Ile Val Lys
 305 310 315 320
 Ala Val Ile Gly Asp Lys Gly Gly Val Arg Met Thr Gly Gly Gly Phe
 325 330 335
 Gly Gly Cys Ile Val Ala Leu Ile Pro Glu Glu Leu Val Pro Ala Val
 340 345 350
 Gln Gln Ala Val Ala Glu Gln Tyr Glu Ala Lys Thr Gly Ile Lys Glu
 355 360 365
 Thr Phe Tyr Val Cys Lys Pro Ser Gln Gly Ala Gly Gln Cys
 370 375 380

<210> 18

<211> 348

<212> PRT

<213> Escherichia coli

<220>

<223> complement of position 5133-6179 of seq id 16

<400> 18

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

- 77 -

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

<210> 19

<211> 12343

<212> DNA

<213> Escherichia coli

<400> 19

tccggcaaca ttatcccacg catgggtcagc aaactgacat tatctctccg ccctggcagg 60
 catcagaaaa tcgctcatatc tttaatcggt aaacagcacc tttagatgct gttttcgata 120
 cacaattttca atcaaggagt cattatgggt ggttggtttg aactcagtaa gagcagtgat 180
 aatcagttcc ggtttgtgct aaaagcgggc aatgggtgaga ctatcctcac cagcgagctt 240
 tatacctcaa aaacctctgc ggaaaagggc atcgcgctcgg tgcgtagcaa cagcccgcga 300

- 78 -

gaagaacgct atgagaaaaa aacggcaagt aacggcaaat tctatttcaa tctgaaagcc 360
gctaatacatc aaattatcgg ctccagccag atgtacgcca ccgcgcaatc tctgaaacc 420
ggaattgcct ccgttaaagc caatggcaca agccagacgg tgaaagacaa tacgtaatta 480
cgatgccggg cgtgggttgt tggcgtccgg cgctttacca gagccccca cagcgctaca 540
atgcccggcc ttaaagtggg ggcactcccc taaccgcttc atcaggtgaa gcggatctga 600
cctgtcatca gaacgagaga attatgttta aaccggaact cctttccccg gcgggaacgc 660
tgaaaaatat gcgttacgct ttcgcttatg gcgcagatgc tgtttatgcg ggccagccgc 720
gttattccct gcgtgtgcg aacaacgaat tcaaccacga aaatcttcag ctggcatca 780
atgaagccca cgcgctgggg aaaaagtttt atgtcgtggt caacattgca ccgcacaacg 840
ccaagctgaa aacctttatc cgtgacctga aaccggtggt ggaaatgggg ccggatgcg 900
tgattatgtc cgatccaggg ctgattatgc tgggtgcgtga gcacttcctt gaaatgccga 960
tccaccttcc ggtgcaggct aacgccgtga actgggcgac ggtgaaattc tggcagcaaa 1020
tgggcctgac ccgcgtgatc ctctctcgcg agctgtcgtt ggaagagatt gaagagatcc 1080
gcaatcaggt gccggatatg gagatcgaga tcttcgttca cggcgcgctg tgcatggcct 1140
actccggtcg ctgcctgctc tctggctata tcaacaagcg cgacccgaac cagggcacct 1200
gcaccaacgc ctgccgctgg gagtacaacg tccaggaagg gaaagaagat gatggttgga 1260
acatcgta caagtacgag ccgattccgg tgcaaaatgt tgagccgacg ctgggtatcg 1320
gcgcaccaac cgacaaagtg tttatgatcg aagaggccca gcgtccgggc gagtatatga 1380
ccgcgtttga agatgagcac ggcacttaca tcatgaactc gaaagatctg cgcgccatcg 1440
cccatgtaga acgcctgacc aaaatgggcg tgcattcgtt gaaaatcgaa ggtcgtacca 1500
aatctttcta ctattgtgca cgcaccgcac aggtttaccg caaagctatc gatgacgccg 1560
ctgcgggaaa accgttcgat accagcctgc tggaaactct ggaaggctct gcgcacgtg 1620
gctataccga aggtttcctg cgtcgtcata ctcacgacga ttatcagaac tacgaatacg 1680
gttattcagt ttctgaccgc cagcagtttg ttggtgagtt taccggtgag cgcaaggggg 1740
acctcgcggc ggtagcgggtg aaaaataaat tctccgttgg cgacagcctt gagctgatga 1800
cgccgcaagg caacattaat tttacccttg agcacatgga aaacgccaaa ggccaagcta 1860
tgccgatagc accaggcgat gggtatactg tgtggctccc ggtcccgcag gatcttgagc 1920
tcaattacgc gctgctgatg cgtaatttct ccggggaaac cagcgtaat cccacggta 1980
agtgattaat ttcgattatt tttcccgat ggaaaattct tagaaaccga tcacatacag 2040
ctgcatttat taaggttatc atccgtttcg ctgaaaaaca taaccataa aatgctagct 2100

- 79 -

gtaccaggaa ccacctcctt agcctgtgta atctccctta cacgggctta ttttttacgc 2160
gtaatacaat gaaataaaaag gattttatttc tggtcacgtc cacacattga ccacatcgac 2220
aaaaaagccc ctcgactgag gggcttcctg tttctaatta catccacata atttgctgcc 2280
ctgacggcaa cgggtgcggc cttacgacgt ggactttctcc cggcttcaca atgtatcgct 2340
gtaccgactc ataagtgatg aacgtggcgc tgcaattcac gttctgacac tggtgataac 2400
gctcttttgt cgtgtcagtg atatagcggc ttgtacgcgc atgtgcggca tgctggcata 2460
aaggacaatg aaacatcgcg agcacctctt ccggttttgt tgatgggtgcc attttagtta 2520
atttaccctt acaaaacaaa cagataaaat aaaaacatca ctattatct tctgtttcgt 2580
accccacatc agaaagcctg acctcaagct ctaatgacgt cgtgaagccg ctattattca 2640
gaaaatgtgt caccttagtg attgtccagt cctgctcgtc tatgacgcgc ttaaagccag 2700
acactttgac cgggtgtttcc gtgtaaatat ctgcccgacc ggtagccagg ctgatggaga 2760
actctgctac acgggcttat ttttttatgc ataagcccta tctctggtaa ccgtcttcca 2820
ttgaccacat cgatagaatc ctcttcata gcacgatgcc tttcacttat cggcatcgtg 2880
ctcccacagg ttccggctac gcacagccag aacgcgcata tttgacgctt accaaaaaat 2940
attctcactc tccacatttg aatgtcagac gagcgacacc atgtaatcct acaccttctg 3000
tcttcagctc aactatttgc atttttttgc cctgagtaac acagaaatga gttgcatcat 3060
tttttactat attttctgca ccagatatcc taccctggc taaagaagct tcggcttcgg 3120
tgtagtattg gttatcgagt ttacgctgaa tattactttt atatgcaaga ccaaatttac 3180
cgatacttgt ctcatcatta tgcacagcac aaccagacat aataaaaaata ctaattaatg 3240
atatagcagc tatctttttc atctcacctt ccccatataa ataccaacga cactctctag 3300
tgtttaaata taataatggc atgattatta taattgaata ggattataat aaatgttctg 3360
tacaacattt cctacataag taggaattac ggacattgag gcccttcagg gtaactccat 3420
gggggcttta atatattata ttgaagatgc cactgttttag ttgaatatta ggtatatgct 3480
cttttttgaa atttatcggg ggcagccgtt agtattcgct gtccccattg caagctcctg 3540
gtggtaacca ctgaatcctc catacttgaa ctgacttttt atcctccgac tttcatcctg 3600
ttctgactcc accttttggt ttctgctcta cactatctac agaccaatca taaaggcaca 3660
tacgatcatg gcagaatttc ccgccagctt actgattctt aatggcaaaa gtactgacaa 3720
tctacccttg cgcgaagcaa ttatgctgtt gcgtgaggaa ggaatgacga tccatgtgcg 3780
ggtcacctgg gagaaaggcg atgccgcacg atatgtagag gaggcccgga agtttggcgt 3840

- 80 -

cgcaacggtg attgccggtg gtggcgatgg caccattaat gaagtttcta cggcgttgat 3900
tcagtgtgag ggggatgaca taccgcgctt gggaattttg ccattaggaa ccgccaatga 3960
ttttgccacc agtgtaggga ttcctgaggg actggataag gcgctgaaac tggcaattgc 4020
cggtgacgcc attgcatag atatggcgca ggtcaacaaa caaacctgtt ttattaatat 4080
ggcgacaggg ggatttggga cgcgtattac cacagaaacg ccggaaaaat taaaagccgc 4140
gctgggtagc gtctcttaca tcattcatgg cttaatgcgt atggatactc tgcaaccgga 4200
ccgttgatga atccgcggtg aaaactttca ctggcaaggc gacgccctgg tcattggtat 4260
tggtaacggg cgtcaggccg gtggcggtca gcaattgtgt ccgaacgcgt taattaacga 4320
tggcttgctg caactgcgca tttttaccgg cgatgaaata cttccggctc tcgtatcaac 4380
cttaaaatct gacgaagata accogaatat tatcgaaggc gcttcgctgt ggtttgatat 4440
tcaggcacca caccagatca cctttaatat tgatggcgaa ccgttgagtg ggcaaaattt 4500
tcatattgaa atacttccgg cagcgttgcg ttgtcgatta ccaccagatt gtccactgtt 4560
gcgttaatca aactattctc tcttgaacca gggagaatcc gtaatgccgg atgcggtgta 4620
aacaccttat ccggcataca gaacaatact acaccgtaat taattttacc cgcgtcttct 4680
gcaatgccag tttatccacc tctgacacag catcatcggc aataataatg tcaatttttt 4740
cgattggtag tacctgatta aaaccgcgac ggttaaattt cgatgaatcg agaacggcaa 4800
tcactttatt agcggctgtg accatcgac cgctaataga ataacctca ttaaaggctc 4860
taataccatt aaccgcatca ataccatcag caccgacaaa cattaaatcg gcattaatat 4920
cctgtaatga acgctcggca atggaaccat gcatcgagcg cgttttatgg cggaccgtgc 4980
caccacagac gaccagagta atgtctttat tttcggaaag ggcaaacgct gccgggagac 5040
tgttggtaat caccgtgata tttttggcgg tcattaatcc ttcggcgata agcattgtcg 5100
tacttccgct gtcgagaata acggtctgat cttaccagc aatagtggac acgcggtcaa 5160
gtgagtaaac tctcagtcag aggtgactca catgacaaaa acagtatcaa ccagtaaaaa 5220
accccgtaaa cagcattcgc ctgaatttcg cagtgaagcc ctgaagcttg ctgaacgcat 5280
cgggtgttact gccgcagccc gtgaactcag cctgtatgaa tcacaactct acaactggcg 5340
cagtaaacag caaatcagc agacgtcttc tgaacgtgaa ctggagatgt ctaccgagat 5400
tgcacgtctc aaacgccagc tggcagaacg ggatgaagag ctggctatcc tccaaaaggc 5460
cgcgacatac ttcgcgaagc gcctgaaatg aagtatgtct ttattgaaaa acatcaggct 5520
gagttcagca tcaaagcaat gtgccgcgtg ctccgggtgg cccgcagcgg ctggtatacg 5580
tggtgtcagc ggcggacaag gataagcagc cgtcagcagt tccgccaaca ctgcgacagc 5640

- 81 -

gttgctcctcg cggctttttac ccggtcaaaa cagcgttacg gtgccccacg cctgacggat 5700
gaactgcgtg ctgaggggtta ccccttttaac gtaaaaaccg tggcggcaag cctgcgccgt 5760
cagggactga gggcaaaggc ctcccgggaag ttcagcccgg tcagctaccg cgcacacggc 5820
ctgcctgtgt cagaaaatct gttggagcag gattttttacg ccagtggccc gaaccagaag 5880
tgggcaggag acatcacgta cttacgtaca gatgaaggct ggctgtatct ggcagtggtc 5940
attgacctgt ggtcacgtgc cgttattggc tgggtcaatgt cgccacgcat gacggcgcaa 6000
ctggcctgcg atgccctgca gatggcgctg tggcggcgta agaggccccg gaacgttatc 6060
gttcacacgg accgtggagg ccagtactgt tcagcagatt atcaggcgca actgaagcgg 6120
cataatctgc gtggaagtat gagcgcaaaa ggttgctgct acgataatgc ctgctggaa 6180
agcttctttc attcgctgaa agtggaatgt atccatggag aacactttat cagccgggaa 6240
ataatgcggg caacggtgtt taattatatc gaatgtgatt acaatcgggtg gcggcggcac 6300
agttgggtgtg gcggcctcag tccggaacaa tttgaaaaca agaacctcgc ttaggcctgt 6360
gtccatatta cgtgggtagg atcagtcac ccttcgtgga tcattttgac tgccgcctgc 6420
gctattctgt tttttggcgc gctggcgagc tgaaatcgct ctttaaaccg gacttcctgg 6480
gtctcggttt cactattacc agacattatt ttgcgcgcac cgccatgaaa gcgcgtaacc 6540
acgccttttt gtctgagaaa gcgcaaactc gcacggattg tcgcttccga ggcagcaaact 6600
actcccgcca gatcctgaac aagcacgggt cctgtttcat tcactaattg gatgatctta 6660
ttccttcgct cgaatgaatt catcgtcatt taccatatt caattgtggc tagtgtaaac 6720
gaagggggac cattggtgaa tgatctgttt tgcacaatac cgggccagca tcacgctggc 6780
ccgcgggttc agggaatgag caacactttg cccggcatag cattacgagc gatgtcacgc 6840
accgcctggg cgaagctttc aaagcttcca cggtgagcga ttaatggctc caggcttaac 6900
ttacgttctg tcagcaaccg gctcgccgtt tcccactcct gccccggcca agggctggag 6960
tagttcatcc aactgccgat aaccgtcagc tcttttacgca atattttgcc aaacgttgcc 7020
gatgttaaact gcagatcctg atgcaacgtg cccaccagcg ccagttgggc atgaggaccg 7080
gcaatctcta ccgccagttc gacagtttgc ggtacgccag ccgtctcgag gataagctga 7140
ttaaagcgca gttcgcgtaa aacgctctgc atttgcggcg cgctcatttc gctactgtta 7200
aatgtttgca tcgcaccgaa agattttgcc agtgccagtt tttctgaact aatgtcgatc 7260
gccgtcacac tctttgctcc cagcgcgaca gcgcactgaa tggccagcag gccaatgggt 7320
ccggcaccaa taataataac gtttttattc tcacaacctt gcgctaaatg aaaagcatgc 7380

- 82 -

agaccaacgg taatcggctc aataaaagcc ccatacctcaa taggcatatc cgtgggtaga 7440
 gcaaagacat tttttcgctt aacgacaata tattcagcaa atccaccatc acgccgcgag 7500
 ccaataaaat cataattttgc gcactgggaa taaaaccctt tcagacactc tggacaagta 7560
 aaacagggta ataacggcac acaggcaacc gcatacgccag ggtgtaaate atcaacaccg 7620
 gatcccaccg catcaatata gccgctaaat tcatggccta acgttatttg ataataatgt 7680
 gcaccatttt taaatatcct gggtaaatec gaaccacata agcccgagct ggcaattttt 7740
 acccgcacct catcctgatg ttttaatttca ggaatgacgc tttctgcaac gcgcacgata 7800
 ccatcagtat cattcaccac tgatttcata aaaactcctg attgttaagg gggataacct 7860
 cccctcagg taaaaattat tctgcgagaa cgactttctc ttgtttaata aagccacgcg 7920
 ctctacgcca ggtcatgaaa ataccggtca gataaattgc accgataata atgaaaccgg 7980
 gaatatatttg cggggagaaa acctgaatca gtaaccagggt aatgggagaa ccgccctgat 8040
 ccattgaagc caccataccc ccggctttta acgccccagc attagccgcc agttgggtgt 8100
 gcaaaccaat agtttgcgtc gcgatccaca gggatgatgct cataatgatg acacccgaga 8160
 ttaagggtgcg gaacagattt ccacgatgca cggcgaccgc catcgccacg aagaagccga 8220
 tgggtggcaag atcgccaaac ggcagcacct gattaccggg cacacaaaca gcaattaaaa 8280
 tggatgagtg gataaaaatc aggcttgccg ataccaccgc cgtatgtccc agcagcaacg 8340
 ccgatcaag gccaatcagg aactcctgac cgccgaactt cgctgtaaa cgactacgag 8400
 cctgcttagc gatgggcgtt aaaccatcca tgatgggttt aatcaccggt ggcattagca 8460
 gcatactgc cgccgttttt accgccagct gcaatacacc tttgacatcg taaccgcgca 8520
 ggatgccgat aatcagcccc atcacaaaac ccacggtgac aggctcgcca aatggaccaa 8580
 atttgcgctg aatatcgctg gcgctaaatt taattcggtt aacgcctggg attttctcga 8640
 tgatagcatc gaccagcacc gcaatcggcc ccataacgc cgacgtaccg tgcggaatag 8700
 caataccttc cagctcaaag aaatttcggg tatcgcgggc aaaccagtcg ccgagcttat 8760
 aaacaaacgc cgctgaatt accacacctg ccatacctat catccatgaa ccggttgcca 8820
 gatgcagcaa cgcgccggtg aaggatcatat gccagatatt ccagatatca acatttacca 8880
 cccgctcat acgggtcagt aacatcgcca cgtaaccag aatcgcaatc ggaatcgcca 8940
 ccagcgcaat ttgcgacgcc caggatcattg gtgaagagcc cggccagcca acatcgacca 9000
 catgcagatt caggctgaaa ttttccgcca tcgctttcgc cgccgggcca atggaatcca 9060
 gcattaagcc aatcacaagg ccaatgcca caaagccaat cccgatatgc agaccgctt 9120
 taaagcaatc gcctgccttc atgcctaata ttttagaaaa aataatgatg acaatcggca 9180

- 83 -

gcatcacctg agggccgagg tcgagaatat aacgcatgac ttctgaaaac ataggtcacc 9240
cctgtaagat agtcagaatt ttatTTTTgta atgcttcgat accgatacca gaaataaaaag 9300
gcatgccgtg aactaacgga atatcgccaa aactacgata cactttggca gtggtgcata 9360
tcaaattgcac accatccata taggtttcta tttcattaac ccgacactgg attaattcaa 9420
caggaatatt atgattctga cacaactctt taattttctt cgcgcgcatc gtagaggctc 9480
caaccgcgcc tccgcaagcg acaataatct tgcgtttcat aaatagtcct ttttattggt 9540
acagagggat ttattcgtct gaattctaaa catactttgt gaagtattcc ttttaactggg 9600
tttcaggaag agtgattagt gtctcgacga tatcgggctg ttgtaacttg ccaaataaac 9660
agcgtaaaaag tttcaattgc tgctgcggat tttccacaat caacgcaata accaacgata 9720
ccgccacgtc gttatcatca tccgcttgct gaaaatgaac tttattttgt ggccttaaca 9780
gataaatggc tgacgactta gcatgaatcg cctcacaatg cggatttgca atagcgtgct 9840
gctcaagcat tatcccggtg gggaattctg cttctctggc aattaatgcc tgtggccagg 9900
tatcatgaac cacacctttg gcgagcatct cattaccgat atgggttaaa acttcgctac 9960
gatcgacaaa agaaattccg ctacgaacaa acagggttagt catataccgt ccttattccg 10020
cacagccgta gcgataggcg cgcaaaacat cataaatttt atccataatc agctgatgcg 10080
gtattgctga taattccccg gactgaatgc gttcaaattg tttgggaaga tactgactaa 10140
tcatgcccag tgggatgtcc acgccttcaa gattcaccat catcgtttcg acgctatttt 10200
taatccgact atgcggccag taataacgaa tacgatccga caggctgtaa cgaatatcca 10260
gtaatgaatc gttaaaaccc gtacgataat attttttcca gtattgcggt tcgtcgagca 10320
tcacttcttc aattaccgcc aggcaaccgc tgcgattttc aggggcgata agttcctgct 10380
caatttgctg cagtgc aaat atcgccctcg gtaaagcaaa ggttaatgcg ggaccgactt 10440
tcaatattgc aaagtgatcg cggactaatt cccaataagc cgtccgggtc tggtaatcgg 10500
tagaatgtgc ttcataaacc attcgggtgt tttctatcca ttgcgccagc ggctgcgctt 10560
cgtgcggctg ataattggata atattgctgt gatcaaattc cacaccggc tgcaccacga 10620
tggcaatcac acgtgttaac gcctctgtca gccacgggc aataaaggcc ttttgatgcg 10680
tacgtaaagt attggcgga tcttcaacat gggatgatgt tactgactga atggcgctgg 10740
cctcaccgcc cggaaccgga acttcggtgc caatgacata gctcagttgc tcacgctggc 10800
aatctgtcgc cacactttcc gcagcaaagc aaagcacagc agctcgttcc gcaaccgttt 10860
ctggtgctaa cggatggga tccccgcgc aggacattga cgcataaga tgaattttac 10920

- 84 -

tgaagccggc acgaacatat tcctttacca gctcgacgga tttttccatc gccgcatccg 10980
 cattttcttg ctgccagcag tttggcccca gatgatcgcc gccgagaata atgcgttcgc 11040
 gtgcaaacc aactttatcg gcaatcgtaa aaacaaattc gcgaaagtct gccggtgtca 11100
 ttccggtata accgccaaat tgattgacct ggtttgacgt tgcttcaatc agcactttgc 11160
 gcgtgctggt gcgatcaaat gccagcgccg cttcgataac caacggatgg gcagaacaga 11220
 ctgaacatat gccgatatgt tcaccagctt tatgccgggc aattaacggt ttcatgtttt 11280
 ttccttgtaa aatggcagggt gcgttatgcc ctgccctcgc agccacaatc ggcaatcact 11340
 ttgctcacca catcgcgcat tgcggattta gccgactgca aataatcccg gggatcggtc 11400
 gcttcagggt gctcggtcag gtaatttttt aacgcctgcg agaaggcatt tttcagctcc 11460
 gttgcaacgt tgattttgca tatccccagt ttgatggttt gctgaatatc cttagtcgat 11520
 aaccctgacg cgccatgcag caccagcggg aagttcaccc actggcgaat gttctccagt 11580
 ctagaaaaat caagcgccgg tgcgctggca tacatcccat gagccgtgcc gatggcgacc 11640
 gccagggat caattccggt tgcctcggca aattcacgcg cctgagcggg gttggtgtac 11700
 aacgcacggt cttcattgac ttgcacatca tcttcctggc cgccaagttg ccccagctcc 11760
 gcctcgacgc tgacatcaaa gcgatggcaa aaatccacca cctctttgac ccgtgaaata 11820
 ttttgcgcaa aaggcaaatg cgaggcgtca atcatgactg agcgcacgcc agaacgaacc 11880
 ttctgagcga tatcgtcaaa tttcgtgtga tggtcgagat gaattgccag tggatggtga 11940
 tattgcttcg ccatcgcgct gaccagcgcc aacagatttt ctgtaccagc atgagtaaata 12000
 gtgccaggcg ttccggcgat gatgaccggc gcatgcaggt tggcagcggg ttctaccacc 12060
 acttgcatcg tttcgagatt gtgaatattg aatgccggaa ccgcataacc gccgcgctgt 12120
 gcgttggtca gcatctgctt tgcgataacc acgtacattt tcatatcctg tcgtttgttt 12180
 tcgatttcaa aatataatga aattatttgt ttttaaataat cgagataacg atcacaaaaa 12240
 cgacaatatg aaaattattc gaggagtgaaggaggcaaaaaa acggcctccc gatagggaag 12300
 ccgtagcaaa gtgcgcgtgt ttttatgccg gatgcggtgt aaa 12343

<210> 20
 <211> 150
 <212> PRT
 <213> Escherichia coli

<220>
 <223> complement of position 9551-10003 of seq id 19

<400> 20
 Met Thr Asn Leu Phe Val Arg Ser Gly Ile Ser Phe Val Asp Arg Ser

WO 01/70776

PCT/US01/07478

- 85 -

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

<210> 21

<211> 451

<212> PRT

<213> Escherichia coli

<220>

<223> complement of position 7877-9232 of seq id 19

<400> 21

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

- 86 -

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

- 87 -

420 425 430

Met Thr Trp Arg Arg Ala Arg Gly Phe Ile Lys Gln Glu Lys Val Val
 435 440 445

Leu Ala Glu
 450

<210> 22
 <211> 10776
 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 <222> (9140)..(10645)

<400> 22
 gaacataagg gaaaccagta ttcacgctgg atcagcgtcg ttttaggtga gttgttaata 60
 aagatttgga attgtgacac agtgcaaatt cagacacata aaaaaacgtc atcgcttgca 120
 ttagaaaggt ttctggccga ccttataacc attaatatcg aagcgcaaaa aaaataatat 180
 ttcctcattt tccacagtga agtgattaac tatgctgatt ccgtcaaaac taagtcgtcc 240
 ggttcgactc gaccataccg tggttcgtga ggcctgctg gctaaacttt ccggcgcgaa 300
 caacttcogg ctggcgctga tcacgagtcc tgccgggctac ggaaagacca ccttcatttc 360
 ccagtgggcg gcaggcaaaa acgatatcgg ctggtactcg ctggatgaag gtgataacca 420
 gcaagagcgt ttgccagct atctcattgc cgccgtgcag caggcaacca acggtcactg 480
 tgcgatatgt gagacgatgg cgcaaaaacg gcaatatgcc agcctgacgt cactcttcgc 540
 ccagcttttc attgagctgg cggaatggca tagcccaact tatctggtca tcgatgacta 600
 tcatctgac actaatccag tgatccacga gtcaatgcgc ttctttattc gccatcaacc 660
 agaaaatctc accctggtgg tgttgtcacg caaccttcgg caactgggca ttgccaatct 720
 gcgtgttcgt gatcaactgc tggaaattgg cagtcagcaa ctggcattta cccatcagga 780
 agcgaagcag ttttttgatt gccgtctgtc atcgccgatt gaagccgcag aaagcagtcg 840
 gatttgcgat gacgtttccg gttgggacgac ggcaactacag ctaatcgccc tctccgcccg 900
 gcagaatacc cactcagccc ataagtcggc acgcccctg gcgggaatca atgccagcca 960
 tcttttcggat tatctggtcg atgaggtttt ggataacgtc gatctcgcaa cgcgccattt 1020
 tctgttgaaa agcgccattt tgcgctcaat gaacgatgcc ctcatcacc gtgtgaccgg 1080
 cgaagaaaac gggcaaattgc gcctcgaaga gattgagcgt caggggctgt ttttacagcg 1140
 gatggatgat accggcgagt ggttctgcta tcacccgctg tttggtaact tcctgcgcca 1200

- 88 -

gcgctgccag tgggaactgg cggcggagct gccggaaatc caccgtgccg ccgcagaaag 1260
ctggatggcc cagggatttc ccagcgaagc aattcatcat gcgctggcgg caggcgatgc 1320
gctgatgctg cgcgatattc tgcttaatca cgcctggagt ctgttcaacc atagcgaact 1380
gtcgctgctg gaagagtcgc ttaaggccct gccgtgggac agcttgctgg aaaatccgca 1440
gttggtgtta ttgcaggcgt ggctgatgca aagccaacat cgctacggcg aagttaacac 1500
cctgctagcc cgtgctgaac atgaaatcaa ggacatcaga gaagacacca tgcacgcaga 1560
atttaacgct ctgcgcgccc aggtggcgat taacgatggg aatccggatg aagcggaacg 1620
gctggcaaaa ctggcactgg aagagctgcc gccgggctgg ttctatagcc gcattgtggc 1680
aacctcgggtg ctgggtgaag tgctgcactg caaaggcgaa ttgaccgct cactggcgct 1740
aatgcagcaa accgaacaga tggcacgcca gcacgatgtc tggcactacg ctttgtggag 1800
tttaatccag caaagtgaag ttctgtttgc ccaagggttc ctgcaaaccg cgtgggaaac 1860
gcaggaaaaa gcattccagc tgatcaacga gcagcatctg gaacagctgc caatgcatga 1920
gtttctgggtg cgcattcgtg cgcagctgtt atgggcctgg gcgcggctgg atgaagccga 1980
agcgtcggcg cgtagcggga ttgaagtctt gtcgtcttat cagccacagc aacagcttca 2040
gtgcctggca atgttgattc aatgctcgtt ggcccgtggg gatttagata acgcccgtag 2100
ccagctgaac cgtctggaaa acctgctggg gaatggcaaa tatcacagcg actggatctc 2160
taacgccaac aaagtccggg tgatttactg gcaaatgacc ggcgataaag ccgcccgtgc 2220
caactggttg cgtcatagcg ctaaaccaga gtttgcgaaac aaccacttcc tgcaagggtca 2280
atggcgcaac attgcccgtg cacaaatctt gctgggcgag tttgaaccgg cagaaattgt 2340
tctcgaagaa ctcaatgaaa atgcccggag tctgcggttg atgagcgatc tcaaccgtaa 2400
cctgttgctg cttaatcaac tgtactggca ggccggacgt aaaagtgcgc cccagcgcgt 2460
gttgctggac gcattaaaac tggcgaatcg caccggattt atcagccatt ttgtcatcga 2520
aggcgaagcg atggcgcaac aactgcgtca gctgattcag cttaatacgc tgccggaact 2580
ggaacagcat cgcgcgcagc gtattctgcg agaaatcaat caacatcatc ggcataaatt 2640
cgcccatttc gatgagaatt tcgttgaacg tctgctaaat catcctgaag tacctgaact 2700
gatccgcacc agcccgtga cgcaacgtga atggcaggta ctggggctga tctactctgg 2760
ttacagcaat gagcaaattg ccggagaact ggaagtcgcg gcaaccacca tcaaaacgca 2820
tatccgcaat ctgtatcaga aactcggcgt ggcccatcgc caggatgcgg tacaacacgc 2880
ccagcaattg ctgaagatga tggggctacgg cgtgtaagtt tagccggata acgcgccaga 2940
tccggcttac atctctgcat cattcaatgc tcaccgcgt tacgccatct gtttctatca 3000

- 89 -

aactaaaccg caccggcaag aaacgctcca ccaccgcgat attggtcagc agatggcatg 3060
agggatgggc gaccgtaaatt tccccgcgc cgcgcagcgc catcggtagc accagttggt 3120
cagcaagata ttccccacc gccgcgtgc ttgccaggta gcgtttcacc tctttcacca 3180
actgtgccgc aaccacctcg gcactgacgc gcttttcacc gacgacaaa aagcgttcgg 3240
tgatatatttc actttcgact tcaagcgaga cggtattacc cggcccctgg tcgcgcggca 3300
ggttatgaat attctgttca tgcagggaaa aactccccgc cagtgtagcg atttcacgct 3360
cagcaacatg gcgcggcaca ccagctaata gaacttctcc acgcatctgc acaatgttcc 3420
cgcgctcgcc aagttgcaag gtgttaaacg atgccacggg cgagacttcc gttgccacca 3480
caccgcctcc ggcagggtaa aaacogtggc gtaacagcgt ggtttgctga tgaattccta 3540
ttttcgccag cagcggctcc agcaccggc ggataaaatc cgaggcggcg ccgacgggtt 3600
atcggtcgcc gcgctcactt caacacgcga aggtccatcg gcaaaccaca gcgcgggcag 3660
caccgtttgc agcaccagcg tacaacttcc ggcgctaccg atagcaaagc ggtaatcgcc 3720
gccgcgcacg gtgccggggc ggaagagcag acgctgcgac ccagctccg cgccttccac 3780
ggttgctccg caaatctccg tcgcgcctt taccgcggtc agatgctggc gcaacagccc 3840
cggtttcgcc cgccgggcac gaatgctggt gatggtaaatt ggctggccgg ttatcatcga 3900
caggctcagc gccgagcgca ggatctgccc gccaccttcg ccctgtgcgc catccagcgc 3960
aatcatcctt ttcattcattt atcctttttac gcacaccacc tgacgcaggg tatagataac 4020
ttccaccaga tcgctttgtg ccgccatcac cgcattcaata tctttatacg ccatcgggat 4080
ttcgtcgatc acttcggcat ctttacggca ttccacatgc gcggtggcac gaatttgatc 4140
ttccacgctg aacagttttt tcgcttttagt tcggctcatt acccgcccg caccgtggct 4200
gcacgaacag aacgactctt catttcccag ccacggacg ataaagcttt ttgctcccat 4260
cgaaccggga ataattccat attgaccagc acgcgcagac accgcgcctt tacgcgtcac 4320
gtagatctct tcaccaaagt gctgttcttt ttgcacatag ttgtggtgac agttgatctc 4380
ttccatcgcc agcgtttgtg gctgtctgac cgttttctgc gtaatgctct gcaatgccgt 4440
taccacgttt tccatcatcg catcgcggtt aaggctggca aaaagctgcg ccagggccac 4500
ggctttcagg taatcatcaa agtattccgt accttcata aagtacgcca gatcacgcga 4560
cggcaacgctc tcaagcgttt cctgcatctc tttttgtgcc agatcgataa agtaagtccc 4620
gatggcggtta ccaattccgc gtgaaccgga gtgcagcata atccacacct ggtccgactc 4680
atcaaggcag atttcaataa agtggttacc ggttcccagc gttcccaggt gtttatagtt 4740

- 90 -

attggtattc aggaacggg gatatttttg cgttaaccac tgataaccgg cttcaagctc 4800
agcccattha gcatcgacgt taacaggtgg attttcccag gcacctttat cacgtttaca 4860
acggccagtg gtacgcccgt gcggcacggc cgtttcaatc gcctgacgca gctctgccag 4920
gttttcaggc aggtcttccg ccgttaacgc ggtacgcagc gcgttcattc cacagccaat 4980
atccacgccc accgcccgcg gaataatcgc ccctttggtc gggatcacgc taccaatggg 5040
ggaaccttta ccaggtgta catcaggcat taccgcaata tgtttgaaaa taaacggcat 5100
cttcgcccga ttaataagtt gctgacgcgc atcggcctct accggcacgc ctttgggtcca 5160
catttttacc ggggcatttt cagtggtcag taattcgtaa ttcattttgt ttttctcttt 5220
tcgttgtttg ctgtcctgat aattgcaacc gtctgtccag aaaattaaat aagcagctta 5280
attttttaat tcattgtttt aaaaaagatt atatctttac gtccgtaacc ggagatttcc 5340
cgcaaagcca atttaccgat aatgaaatat cgtcttttat aaggatatct aagatgcgta 5400
aaacagtggc ttttggcttt gtccgtaccg tactggatta tgccgggcgc ggcagtcagc 5460
gctgggtcaa atggcgtccg acactctgtt tatgccagca agaatcgttg gtcacgcatc 5520
gactggaatt gttgcacgac gcccgctcgc gctcgtatt tgaaacgctt aaacgcgata 5580
tcgccagcgt ttcgccagaa acagaagtgg tgagcgttga gattgaactg cataaccctg 5640
gggatttcga agaggtctac gcctgcctgc atgatttcgc ccgtggttac gagtttcagc 5700
cagaaaaaga agactattta attcacatca ccaccggtac ccacgtcgcg cagatttgct 5760
ggtttctgct ggcagaagcg cgttacctgc ccgccggct gatacaatct tcacctcgc 5820
gcaaaaaaga acagccgcgc ggcccagggtg aagtaacgat tatcgatctc gatttaagcc 5880
gttataacgc catcgccagc cgctttgccg aggaacgcca gcaaacgctt gattttctta 5940
agtccggcat tgccacgcgt aacccccact tcaaccgcat gattgagcag atcgaaaaag 6000
tggcgatcaa atcccgcgcg ccgattctgc ttaacggtec aaccggcgcg ggcaagtcac 6060
ttctggcgcg acgcatctta gagttaaaac aggcgcggca tcagttagc ggcgcgtttg 6120
tggaagtga ctgcgccacc ctgcgcggcg ataccgcat gtcgacgctg tttgggtcatg 6180
taaaaggcgc gtttaccggg gcgcgggaat ctctgaagg tttattacgc agcgccaacg 6240
gcggaatgtt gtttcttgat gagattggcg aactgggcgc agacgaacag gcaatgctgc 6300
tgaaagccat tgaagagaaa accttttacc cgtttggcag cgatcgccag gtgagcagtg 6360
attttcagct tatcgccgga acggtgcgcg atttgcgcca gctggttgcc gaaggcaaata 6420
ttcggaaga tctgtacgcg cggtatcaatc tctggacctt caccctgccg ggtctacgcc 6480
agcgccagga agatattgaa ccgaacctgg attatgaagt ggagcgccac gcctcactca 6540

- 91 -

ctggcgacag cgtgcgtttt aacaccgaag cgcggcgcgc ctggttggct tttgcgacct 6600
ctccccaggc aacatggcgc ggtaactttc gcgaactttc tgccagcgtc acgcggatgg 6660
ccacctttgc cactagcgga cgcatactc tggacgtggc tgaagatgag ataaaccgtc 6720
tgcgctataa ctggcaggag agccgcccct ccgcgcttac ggcgttgctg ggcgagagg 6780
cagaaaacat cgatctcttc gaccgtatgc aactggaaca cgttatcgct atctgccgcc 6840
aggcaaagtc gctttccgca gccggacgcc agctttttga cgtttcgcgc cagggcaaag 6900
ccagcgtcaa tgacgcggat cggctacgca aatacctggc gcgttttggc ctgacgtggg 6960
aagccgtgca ggatcagcac agctccagtt gaatatggtg gtccgtcagc acctgcatca 7020
cgcttactgg cggcggggcg tcggtgtaga cggcatctac catgctgatg ctgcccatac 7080
tgaccatcgc gttacggcca aatttcgagt gatcgacaac cagcataacg tggcgcgagt 7140
tctcaataat ggcgcgtttg gtgcgaactt cgtggtaatc gaactccagc agcgagccgt 7200
cgctatcgat gccgcttacc ccagaaatgc cgaaatcaag gcggaactgg gagataaaat 7260
cgagcgtcgc ttcgccaatg atcccgccat cgcggctgcg taattcgcca ccggcgagaa 7320
tgatgcgaaa atcttctttt accatcaacg tgtagcaac gttgagattg ttggtgacaa 7380
tgcgcaaatt gctgtgattg agcagtgcgt gcgctaccgc ttccggcgtg gtgccgatat 7440
cgataaacag cgtcgagcca ttggggattht gctccgccac tttgcgggag atgcgctctt 7500
tttcttcggc ctgggtggcc ttgcgatcgt gccacggcgt gtttaaccgaa ctggaaggca 7560
gcgccgcacc gccatgatgg cgcaggatca ggttttgctc cgccagctca ttgaggctgc 7620
ggcgaatagt ctgcgggctg acggagaaat gctctaccag ctcttcggta ctgacataac 7680
cctgctgttt aaccagttcg ataataccgt tgtgacgttg tgtttgtttc atttataaat 7740
ccctggaatt attttcgttt tcgcgcattg agcgaatcaa caaaagccat cgctaaacc 7800
acggctaacc cggcgatgtg tgctccgttc gccatcgaca tcccaaaca atcaaaccat 7860
ccggcgacaa tccagatcag cgcaaagata attaaaccac gttgcaggta aatgccactt 7920
tgcggatcgc gttcgccacg tagccagacg tagcccatca gcgcatacac cagccagaa 7980
agcccgccaa accacggccc gctgaatttt tgctgcacat agccgcttaa cagggcgctg 8040
ataagcgtaa tgacaattag cttaccgcta ccgaggcgtt tttccaccgc accgccgaga 8100
taccaccacc agagcagggt aaagaggata tgcatcagcg agaagtgcac taacgcgtgg 8160
gtgaagtaac gccagaactc aaatttcagt gttggatcga atggccaggc cagccataac 8220
atcacttcct gatcgccgag aatttgcatg gcaataaaca ccaccacgca ggcgatcatc 8280

atcacc	cagg	ttacc	ggacc	tgcg	cgttca	cgca	aggcgg	caa	agaaagg	ata	acggcga	8340				
taatgc	caggc	cactg	ccggt	atgg	cctgcc	tgcc	agctcg	ccg	ccagata	acgc	ggatct	8400				
gccggg	gtttt	cgaga	aaaacg	cgcc	acgtcc	gccc	gtacgc	gctc	ggcctg	ggact	catcc	8460				
gccagc	ccaga	catcg	ctttg	gttat	gtttgt	tgaat	cgtaga	ggata	acacc	ctgc	gtcgcc	8520				
atgta	atcaa	caaac	gcctg	cgcc	acgcgg	gggt	tagcaa	aagag	gtaat	catc	aacatc	8580				
gttgct	gtcg	cttatt	ccac	acaaa	agggg	acagt	tataaa	gcgt	tacgcg	ccgt	acgcca	8640				
cctctg	cggg	aaact	gacgt	tgcc	aggctt	caaag	ccgcc	gtca	atgcta	tagac	cat	8700				
cgtag	ccctg	ttgc	cagcaga	tactg	cgccg	cgctt	ttgct	gctatt	gccg	tgata	acaca	8760				
tcacc	atcac	cggag	tgtca	aagtc	gttat	cacgc	ataaa	agcgc	cccagc	gtgtc	gttgg	8820				
ttaa	atggaa	agcct	gcacc	gcat	gtccca	ttgc	gaaact	ctgt	ggatcg	cgaat	atcga	8880				
ccagc	accgc	ctctt	tttcc	tgca	acttct	ggtg	cgcgtc	ggca	acgtta	atac	attcga	8940				
actga	tccat	gcgt	ctctct	ttctt	tacaa	acaag	tgggc	aaatt	taccg	cacag	tttac	9000				
gtcga	agcgg	cagata	aaacg	ccata	atgtt	atac	atatca	ctcta	aaaatg	ttttt	tcaat	9060				
gttac	ctaaa	gcgc	gattct	ttgct	aatat	gttc	gataac	gaac	atttat	gagct	tttaac	9120				
gaaagtgaat gagggcagc atg gaa acc aaa gat ctg att gtg ata ggg ggc												9172				
Met Glu Thr Lys Asp Leu Ile Val Ile Gly Gly																
1 5 10																
ggc	atc	aat	ggt	gct	ggt	atc	gcg	gca	gac	gcc	gct	gga	cgc	ggt	tta	9220
Gly	Ile	Asn	Gly	Ala	Gly	Ile	Ala	Ala	Asp	Ala	Ala	Gly	Arg	Gly	Leu	
15 20 25																
tcc	gtg	ctg	atg	ctg	gag	gcg	cag	gat	ctc	gct	tgc	gcg	acc	tct	tcc	9268
Ser	Val	Leu	Met	Leu	Glu	Ala	Gln	Asp	Leu	Ala	Cys	Ala	Thr	Ser	Ser	
30 35 40																
gcc	agt	tca	aaa	ctc	att	cac	ggt	ggc	ctg	cgc	tac	ctt	gag	cac	tat	9316
Ala	Ser	Ser	Lys	Leu	Ile	His	Gly	Gly	Leu	Arg	Tyr	Leu	Glu	His	Tyr	
45 50 55																
gaa	ttc	cgc	ctg	gtc	agc	gag	gcg	ctg	gct	gaa	cgt	gaa	gtg	ctg	ctg	9364
Glu	Phe	Arg	Leu	Val	Ser	Glu	Ala	Leu	Ala	Glu	Arg	Glu	Val	Leu	Leu	
60 65 70 75																
aaa	atg	gcc	ccg	cat	atc	gcc	ttc	ccg	atg	cgt	ttt	cgc	ctg	cca	cat	9412
Lys	Met	Ala	Pro	His	Ile	Ala	Phe	Pro	Met	Arg	Phe	Arg	Leu	Pro	His	
80 85 90																
cgt	ccg	cat	ctg	cgc	ccg	gcg	tgg	atg	att	cgc	att	ggt	ctg	ttt	atg	9460
Arg	Pro	His	Leu	Arg	Pro	Ala	Trp	Met	Ile	Arg	Ile	Gly	Leu	Phe	Met	
95 100 105																
tac	gat	cat	ctg	ggt	aaa	cgc	acc	agc	ttg	ccg	gga	tca	act	ggt	ttg	9508
Tyr	Asp	His	Leu	Gly	Lys	Arg	Thr	Ser	Leu	Pro	Gly	Ser	Thr	Gly	Leu	
110 115 120																

- 93 -

cgt ttt ggc gca aat tca gtg tta aaa ccg gaa att aag cgc gga ttc	9556
Arg Phe Gly Ala Asn Ser Val Leu Lys Pro Glu Ile Lys Arg Gly Phe	
125 130 135	
gaa tat tct gac tgt tgg gta gac gac gcc cgt ctg gta ctc gcc aac	9604
Glu Tyr Ser Asp Cys Trp Val Asp Asp Ala Arg Leu Val Leu Ala Asn	
140 145 150 155	
gcc cag atg gtg gtg cgt aaa ggc ggc gaa gtg ctt act cgg act cgc	9652
Ala Gln Met Val Val Arg Lys Gly Gly Glu Val Leu Thr Arg Thr Arg	
160 165 170	
gcc acc tct gct cgc cgc gaa aac ggc ctg tgg att gtg gaa gcg gaa	9700
Ala Thr Ser Ala Arg Arg Glu Asn Gly Leu Trp Ile Val Glu Ala Glu	
175 180 185	
gat atc gat acc ggc aaa aaa tat agc tgg caa gcg cgc ggc ttg gtt	9748
Asp Ile Asp Thr Gly Lys Lys Tyr Ser Trp Gln Ala Arg Gly Leu Val	
190 195 200	
aac gcc acc ggc ccg tgg gtg aaa cag ttc ttc gac gac ggg atg cat	9796
Asn Ala Thr Gly Pro Trp Val Lys Gln Phe Phe Asp Asp Gly Met His	
205 210 215	
ctg cct tcg cct tat ggc att cgc ctg atc aaa ggc agc cat att gtg	9844
Leu Pro Ser Pro Tyr Gly Ile Arg Leu Ile Lys Gly Ser His Ile Val	
220 225 230 235	
gtg ccg cgc gtg cat acc cag aag caa gcc tac att ctg caa aac gaa	9892
Val Pro Arg Val His Thr Gln Lys Gln Ala Tyr Ile Leu Gln Asn Glu	
240 245 250	
gat aaa cgt att gtg ttc gtg atc ccg tgg atg gac gag ttt tcc atc	9940
Asp Lys Arg Ile Val Phe Val Ile Pro Trp Met Asp Glu Phe Ser Ile	
255 260 265	
atc ggc act acc gat gtc gag tac aaa ggc gat ccg aaa gcg gtg aag	9988
Ile Gly Thr Thr Asp Val Glu Tyr Lys Gly Asp Pro Lys Ala Val Lys	
270 275 280	
att gaa gag agt gaa atc aat tac ctg ctg aat gtg tat aac acg cac	10036
Ile Glu Glu Ser Glu Ile Asn Tyr Leu Leu Asn Val Tyr Asn Thr His	
285 290 295	
ttt aaa aag cag tta agc cgt gac gat atc gtc tgg acc tac tcc ggt	10084
Phe Lys Lys Gln Leu Ser Arg Asp Asp Ile Val Trp Thr Tyr Ser Gly	
300 305 310 315	
gtg cgt ccg ctg tgt gat gat gag tcc gac tcg ccg cag gct att acc	10132
Val Arg Pro Leu Cys Asp Asp Glu Ser Asp Ser Pro Gln Ala Ile Thr	
320 325 330	
cgt gat tac acc ctt gat att cat gat gaa aat ggc aaa gca ccg ctg	10180
Arg Asp Tyr Thr Leu Asp Ile His Asp Glu Asn Gly Lys Ala Pro Leu	
335 340 345	
ctg tcg gta ttc ggc ggt aag ctg acc acc tac cga aaa ctg gcg gaa	10228
Leu Ser Val Phe Gly Gly Lys Leu Thr Thr Tyr Arg Lys Leu Ala Glu	

- 94 -

350	355	360	
cat gcg ctg gaa aaa cta acg ccg tat tat cag ggt att ggc ccg gca			10276
His Ala Leu Glu Lys Leu Thr Pro Tyr Tyr Gln Gly Ile Gly Pro Ala			
365	370	375	
tgg acg aaa gag agt gtg cta ccg ggt ggc gcc att gaa ggc gac cgc			10324
Trp Thr Lys Glu Ser Val Leu Pro Gly Gly Ala Ile Glu Gly Asp Arg			
380	385	390	395
gac gat tat gcc gct cgc ctg cgc cgc cgc tat ccg ttc ctg act gaa			10372
Asp Asp Tyr Ala Ala Arg Leu Arg Arg Arg Tyr Pro Phe Leu Thr Glu			
	400	405	410
tcg ctg gcg cgt cat tac gct cgc act tac ggc agc aac agc gag ctg			10420
Ser Leu Ala Arg His Tyr Ala Arg Thr Tyr Gly Ser Asn Ser Glu Leu			
	415	420	425
ctg ctc ggc aat gcg gga acg gta agc gat ctc ggg gaa gat ttc ggt			10468
Leu Leu Gly Asn Ala Gly Thr Val Ser Asp Leu Gly Glu Asp Phe Gly			
	430	435	440
cat gag ttc tac gaa gcg gag ctg aaa tac ctg gtg gat cac gaa tgg			10516
His Glu Phe Tyr Glu Ala Glu Leu Lys Tyr Leu Val Asp His Glu Trp			
	445	450	455
gtc cgc cgc gcc gac gac gcc ctg tgg cgt cgc aca aaa caa ggc atg			10564
Val Arg Arg Ala Asp Asp Ala Leu Trp Arg Arg Thr Lys Gln Gly Met			
	460	465	470
tgg cta aat gcg gat caa caa tct cgt gtg agt cag tgg ctg gtg gag			10612
Trp Leu Asn Ala Asp Gln Gln Ser Arg Val Ser Gln Trp Leu Val Glu			
	480	485	490
tat acg cag cag agg tta tcg ctg gcg tcg taa attaacgtaa ggtgatcagg			10665
Tyr Thr Gln Gln Arg Leu Ser Leu Ala Ser			
	495	500	
tcagattttca atctggcctg agactgatga caaacacaaa actgcctgat gcgctacgct			10725
tatcaggcct acgtggttta tgcaatatat tgaatttgca tgggtcttgta g			10776

<210> 23

<211> 501

<212> PRT

<213> Escherichia coli

<400> 23

Met	Glu	Thr	Lys	Asp	Leu	Ile	Val	Ile	Gly	Gly	Gly	Ile	Asn	Gly	Ala
1				5					10					15	

Gly	Ile	Ala	Ala	Asp	Ala	Ala	Gly	Arg	Gly	Leu	Ser	Val	Leu	Met	Leu
			20					25					30		

Glu	Ala	Gln	Asp	Leu	Ala	Cys	Ala	Thr	Ser	Ser	Ala	Ser	Ser	Lys	Leu
		35					40					45			

Ile	His	Gly	Gly	Leu	Arg	Tyr	Leu	Glu	His	Tyr	Glu	Phe	Arg	Leu	Val

WO 01/70776

PCT/US01/07478

- 95 -

50		55		60												
Ser 65	Glu	Ala	Leu	Ala	Glu	Arg	Glu	Val	Leu	Leu	Lys	Met	Ala	Pro	His	80
Ile	Ala	Phe	Pro	Met	Arg	Phe	Arg	Leu	Pro	His	Arg	Pro	His	Leu	Arg	95
Pro	Ala	Trp	Met	Ile	Arg	Ile	Gly	Leu	Phe	Met	Tyr	Asp	His	Leu	Gly	110
Lys	Arg	Thr	Ser	Leu	Pro	Gly	Ser	Thr	Gly	Leu	Arg	Phe	Gly	Ala	Asn	125
Ser	Val	Leu	Lys	Pro	Glu	Ile	Lys	Arg	Gly	Phe	Glu	Tyr	Ser	Asp	Cys	140
Trp	Val	Asp	Asp	Ala	Arg	Leu	Val	Leu	Ala	Asn	Ala	Gln	Met	Val	Val	160
Arg	Lys	Gly	Gly	Glu	Val	Leu	Thr	Arg	Thr	Arg	Ala	Thr	Ser	Ala	Arg	175
Arg	Glu	Asn	Gly	Leu	Trp	Ile	Val	Glu	Ala	Glu	Asp	Ile	Asp	Thr	Gly	190
Lys	Lys	Tyr	Ser	Trp	Gln	Ala	Arg	Gly	Leu	Val	Asn	Ala	Thr	Gly	Pro	205
Trp	Val	Lys	Gln	Phe	Phe	Asp	Asp	Gly	Met	His	Leu	Pro	Ser	Pro	Tyr	220
Gly	Ile	Arg	Leu	Ile	Lys	Gly	Ser	His	Ile	Val	Val	Pro	Arg	Val	His	240
Thr	Gln	Lys	Gln	Ala	Tyr	Ile	Leu	Gln	Asn	Glu	Asp	Lys	Arg	Ile	Val	255
Phe	Val	Ile	Pro	Trp	Met	Asp	Glu	Phe	Ser	Ile	Ile	Gly	Thr	Thr	Asp	270
Val	Glu	Tyr	Lys	Gly	Asp	Pro	Lys	Ala	Val	Lys	Ile	Glu	Glu	Ser	Glu	285
Ile	Asn	Tyr	Leu	Leu	Asn	Val	Tyr	Asn	Thr	His	Phe	Lys	Lys	Gln	Leu	300
Ser	Arg	Asp	Asp	Ile	Val	Trp	Thr	Tyr	Ser	Gly	Val	Arg	Pro	Leu	Cys	320
Asp	Asp	Glu	Ser	Asp	Ser	Pro	Gln	Ala	Ile	Thr	Arg	Asp	Tyr	Thr	Leu	335
Asp	Ile	His	Asp	Glu	Asn	Gly	Lys	Ala	Pro	Leu	Leu	Ser	Val	Phe	Gly	350
Gly	Lys	Leu	Thr	Thr	Tyr	Arg	Lys	Leu	Ala	Glu	His	Ala	Leu	Glu	Lys	365

- 96 -

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

<210> 24
 <211> 15067
 <212> DNA
 <213> Escherichia coli

<400> 24
 ataacagcaa tcaaggaaaa ggggaaaatc agcaattttc tgaaagagat gccctttccg 60
 gcggcaaagg gcatcatggg aaatcagtta tagttaatct taaaaagcac cactgtatcg 120
 aaggggccag cctcgatggg ggtgcctgga atggacgaaa gttcagccgt atagggttttt 180
 tcgtagacca acatgtcact ggtgaaatcg gcgtattcag catatttatt gaagggtataa 240
 ggcgaggcgt tttcatccag taaacgtagc ttcagaccat taccaatcaa caatgcttta 300
 tcttcttcaa ccaacgtttc ttcggtataa aaggaggaac tgactttaaa tccgtcgggtg 360
 cattgatcgt tttggctctt ggtcgtttta atactgaacg ttttcgacat tgtgtggcga 420
 cgtatatcca gcaaattaaa tttaccaaaa tcaatcacct gtgtttccgg tgaaatggaa 480
 aaattgacgc tgcaatccag cacgcgaatg ttttccagac cagtaatatg atatttcagg 540
 ttttgggcag tagggtcttc gttaacgcta ccggcaccgt caaattgcac gacgatatag 600
 tcgctaagcg tactttgata atcatgcggc ggcaatttctc gtattttgac gtataaccgc 660
 atgcgagcga gaaaggtagc cgacatatga atattagcgg gatcgcccga gcaaattttg 720
 ttctgccatc ccatcgcgat gatctgctca ggggtataaa tatcaatggt cttactgtcg 780

- 97 -

atacactggg tagtgtcgat gcgacttttt cccagactgg catcgtaatc gacaccggtg 840
tatgtcacgc cgagttgata ataacgatcc tgcacccccg gataggggtt taccagggcg 900
tagacatgct cgctttcaaa gttgccggtg gtattgttgt cacaatagac ggggaatttta 960
atatcgtcag attcccagat cttatcacgc ggtttggcat tgcctggtag agcaaacggt 1020
tgaatcgctt ctgatttttc tactgcgcca ccagatgtac caaaatagca gttcagcgcc 1080
attgcagact tcaacggcag aaataacagg cagatgacga gcaacaatcc cttgccagca 1140
ctcatttcgc ctccactgct tgtaaaccgc attgtccgct gcagttgagc gtcgtcatgc 1200
gtaatgcgcc gtaatcatcc atgtaaccga gataaaactg gctgccgcta taacttccgg 1260
tgtttgtgtt gaccgtacta aaagggtgcaa ccatcacggt tttgaagcct gggagaacgc 1320
ctttttcatt tctcccaga taggcaatgg tcagataata cgcggtagga tttttcagtg 1380
tcagttgggt gccctgctgg ctgacctgta actgcagttc gaccttttct cctgctttct 1440
tgcgtaatgc agccggtcgc cagaacaatt taatacgact ttgaatagcg acctgaagta 1500
tcgcatggtc gctgctttta tccggagcgg gaggaatttc gcgcatattg tagaaaaata 1560
gcgtttcgcg atcgccaggt aattgtgtgg ttgaggcttg tttcacaacc cgcacctgcg 1620
atgtcgcttt cggttccaga cgttgaatag ggggcaaagc gaccagaaga gcatcgcttt 1680
tctccccttt ttcattttct atccaggaat aagcaagata aggcagtttg tcactttgggt 1740
tttcgatacg caggctgggtg gctttgtcat tggcggttaa cacaatgcgt gtgcgatcgg 1800
gttgcaccgc agcattgcaa cccaagctga tgggttaacaa cattaagggt aatcctttta 1860
taaattgtcat tctcttttcc ttaattaatg acaaggtaaa atcaggcggt tggtcgtgtc 1920
ttccagacgt tctggcagat gaataatgca gctttgctca cccagacca cggtaaacag 1980
ttgatattca gcgacaccgc ttaaccaggc atgtccttcc tcgccaacca taccacgct 2040
aatgccgctg tcatcctggc ggatatccgc accgagggga gggaaactgac cgctggcggt 2100
gcgtatgatg acgttaacgt ctttaccgga acgggaagcc agtgatttgt aacctatgc 2160
gccttcaatc catgtttcct tgatgacggt ttctgcaact gttacgccgt cgggtaagtc 2220
attcatgttc accgccacgg tggaaggctg gtaactggaa atcaacggca ccaccgcaat 2280
gccaaaatgg ttggtgtagt cgagattgcc ctgaaccgga atatctgcca cgccatcggt 2340
gctgaccatc aggcgtgggt cattgggtgga gctgcggcga tgaaatgctg caccatattg 2400
ggttgcggtg aaagaaccgc tccagctgct gctgacggaa ctgtaatcac tggcggcata 2460
ggtaccagaa atatccact caccgctga actcagggtg tgatagttac cgctcacctg 2520

- 98 -

ggctccattg tccggacgat cggattgcag tccggcagac atgccccagc tattacgttc 2580
atccagcgta tcgttccacg acatgcggtg tatggtgctg tggctactgt tttgcatgtc 2640
atagccaacc cgaccaccgt taccgaaggg gagcgaaatc gacagataaa tctgggtgtc 2700
gcgatcttta tcttcgtagt gggtcgtatt gaatgacgtc gaaatcgaga tatctctcca 2760
gtcaccaata tcaacattaa aacctgctgt gatgttggcg gtcgtggagg cgtctgcatt 2820
ccaccagggt tgatgtagca ggttggcgta aagattgagg tttagtgggg taatcgggtg 2880
gcccacggat aagctgatcg tctgtttttc gtcctgcgca tcgctgtcgt tgtatttgtg 2940
atccagatag ttggcgtagc tgtgaaaatg acgatctgag aagcgataag cggcgagcga 3000
aatcgtgctg ttggtggcat ccacttgttt gctgtaatta aaacggtagc ttaagccccg 3060
ctcgtcctgc tgagtatcaa aatggctact ggcccagggtg acatcaaagg acagtgcgcc 3120
aagccacagc atattttgcc cgatacccat tgcggcagaa tggtagtcat catcagaaat 3180
cagcaggccg ccgtacagcg aggtgtttga gagcatcccc caggaaactt cattgctaaa 3240
aaaggtttca ttttcagttt gatgtgacat ggaaggccgt ggctgaccgg cggccaattt 3300
atagcgaacc tgtccctggc gagtcaggaa gggcgtcgat gctgccgaaa cctggaaatt 3360
gttcaccoga ccatcttctt ccgtcacttt gacatccagc gtgccctgaa cagactgatt 3420
gaggtcatca ataataaacg ggcctggcgg gacttttttc tggtaaataga cgcggcctga 3480
ttgactgate gtcaccgtgg cattggtctg tgcaataccg ctaatttgtg gggcgtagcc 3540
acgcagctcc cacggtaaca ttcgatcgtc actcgccagt gccgcgccgg tataagaaaa 3600
accatcgaaa atattggaac tgaaatcggg ttcgccgagg gttaacttag agcctaattg 3660
cggtaatgga cgaaaaagat aggtgcgcga tattccgcct gactgggtcat ggttatcttc 3720
gctatcggtc ttattaagct ggtaatcact gcgtaagcgc catgccccgg cgtaatttcc 3780
ggcggtagcg taggcattca ggtagtgct gctgctgccg tcctgtgggc ggtagctgct 3840
ggcaaacaag ttgtaatcca tcaggacacc ggcaacacct tctttccatg tagagggggg 3900
agcccagttt tctgagtgcc acgccagcca ggcttgcgga atactaatat tcagttgctg 3960
attggcttga tcgaaattga agagcatttc aggtcgggaa ctgaaatcaa tacaccgatc 4020
tatctgtggc aaggactgac ggatatctgg ttttaaacca aatttatcga ccagtgaatc 4080
attgatgcat ggaatggttt tgtcaccctt tttttgccag ttaatttttt gccattact 4140
gattttgttg ttattcaccg caacgctaac aaaatattca ccgggagcaa tgactccttt 4200
ttcctttaat aatgaaatat caatgcggtc gcgcattgat ttatcgagaa cattcagatt 4260
gaattcgacc gccattgcgc aaggcatcgc catgaccagg caggatacaa aagagagtcg 4320

- 99 -

ataaatattc acggtgtcca tacctgataa atatTTTTatg aaaggcggcg atgatgccgc 4380
caaaataata cttatTTtata atccagcacg taggttgCGt tagcggttac ttcacctgcc 4440
gtgacatcga ctgcattatc aatttgTtcc atccaggcga aaaagtTcag cgtctgtTct 4500
gatgagcttg catccaggTc aagatctggc gcggctgaac ctaatacgat gttaccgtca 4560
TTTTtgTcca tcagtcgtac accgacccca gttgctTcgc ctgcactggT gttgctcaac 4620
aaaggcgtag caccagTtgt cttagccgtg ctatcgaagg ttacgccaac tttggatacc 4680
ggcattccgc taccgtTgtc agaagcaggc agatcacagt tgatcaagcg aatgtcgacg 4740
gccactTTat tgctatgatg ctcccggttt atatgggttg tcgtgacttg tccaagatct 4800
atgTTTTtat caatatcttc tggatgaatt tcacaaggTg cttcaataac ctccccctta 4860
aagtgaattt cgccagaacc ttcatcagca gcataaacag gtgcagtga cagcagagat 4920
acggccagtg cggccaatgt tttttgtcct ttaaacataa cagagtcctt taaggatata 4980
gaataggggt atagctacgc cagaatatcg tatttgatta ttgctagttt ttagttttgc 5040
ttaaaaatat tgttagtttt attaaatgca aaactaaatt attggtatca tgaatttgTt 5100
gtatgatgaa taaaatatag gggggtatag atagacgtca ttttcatagg gttataaatg 5160
cgactaccat gaagTTTTta attgaaagta ttgggttgct gataatttga gctgttctat 5220
tctTTTTaaa tatctatata ggtctgttaa tggattttat ttttacaatt ttttgTgttt 5280
aggcatataa aaatcaacc gccatatgaa cggcgggtta aaatatTtac aacttagcaa 5340
tcaaccatta acgcttgata tcgctTTTTa agtcgcgttt ttcatatcct gtatacagct 5400
gacgcggacg ggcaatcttc ataccgtcac tgtgcatttc gctccagtgg gcgatccagc 5460
caacggTacg tgccattgcg aaaatgacgg tgaacatgga agacggaata cccatcgctt 5520
tcaggatgat accagagtag aaatcgacgt tcgggtacag tttcttctcg ataaagtacg 5580
ggtcgtTcag cgcgatgttt tccagctcca tagccacttc cagcaggTca tccttcgtgc 5640
ccagctcttt cagcacttca tggcaggTtt cacgcattac ggtggcgcgc gggtcgtaat 5700
TTTTgtacac gcggtgaccg aagcccatca ggcggaaaga atcatttttg tctttcgcac 5760
gacgaacaaa ttccggaatg tgTTtaacgg agctgatttc ttccagcatt ttcagcgccg 5820
cttcgttage accgccgtgc gcaggTcccc acagtgaagc aatacctgct gcgatacagg 5880
caaacgggtt cgcacccgaa gagccagcgg tacgcacggT ggaggtagag gcgttctgtt 5940
catggTcagc gtgcaggatc agaatacggT ccatagcacg ttccagaatc ggattaactt 6000
catacggTtc gcacggcgtg gagaacatca tattcaggaa gttaccggcg taggagagat 6060

- 100 -

cgttgcgcg gtaaacaat ggctgaccaa tggaatactt gtaacacatc gcggccatgg 6120
tcggcathtt cgacagcagg cggaacgcgg caatttcacg gtgacgagga ttgttaacat 6180
ccagcgagtc gtgatagaac gccgccagcg cgccggtaat accacacatg actgccattg 6240
gatgcgagtc gcgacggaaa gcatggaaca gacgggtaat ctgctcgtgg atcatgggat 6300
gacgggtcac cgtagtthta aattcgtcat actgttcctg agtcgggttt tcaccattca 6360
gcaggatgta acaaacttcc aggtagttag aatcggtcgc cagctgatcg atcgggaaac 6420
cgcggtgcag caaaatacct tcatcaccat caataaaagt aatttttagat tcgcaggatg 6480
cggttgaagt gaagcctggg tcaaagggtga acacaccttt tgaaccgaga gtacggatat 6540
caataacatc ttgacccagc gtgcctttca gcacatccag ttcaacagct gtatccccgt 6600
tgaggggtgag ttttgctttt gtatcagcca ttttaagggtct ccttagcgcc ttattgcgta 6660
agactgccgg aacttaaatt tgccttcgca catcaacctg gctttaccgg ttttttattt 6720
ggctcgccgc tctgtgaaag aggggaaaac ctgggtacag agctctgggc gcttgcagg 6780
aaaggatcca ttgatgacga ataaatggcg aatcaagtac ttagcaatcc gaattattaa 6840
acttgtctac cactaataac tgtcccgaaat gaattgggtca atacttccac actgtttacat 6900
aagttaatct taggtgaaat accgacttca taacttttac gcattatatg cttttcctgg 6960
taatgtttgt aacaactttg ttgaatgatt gtcaaattag atgattaaaa attaaataaa 7020
tgttgttatc gtgacctgga tcaactgttca ggataaaacc cgacaaacta tatgtaggtt 7080
aattgtaatg attttgtgaa cagcctatac tgccgccagg totccggaac accctgcaat 7140
cccgagccac ccagcgttgt aacgtgtcgt tttcgcatct ggaagcagtg ttttgcata 7200
cgcgagttta tagaaaggac gctgtctgac ccgcaagcag accggaggaa ggaaatccc 7260
acgtctccag gtaacagaaa gttaacctct gtgcccgtag tccccaggga ataataagaa 7320
cagcatgtgg gcgttattca tgataagaaa tgtgaaaaaa caaagacctg ttaatctgga 7380
cctacagacc atccggttcc ccatcacggc gatagcgtcc attctccatc gcgtttccgg 7440
tgtgateacc tttgttgagc tgggcatcct gctgtggctt ctgggtacca gcctctcttc 7500
ccctgaagg 6120
ttcgagcaag cttccgcgat tatgggcagc ttcttcgtca aatttatcat 7560
gtggggcatc cttaccgctc tggcgatatc cgtcgtcgta ggtattcgcc acatgatgat 7620
ggatthttggc tatctggaag aaacattcga agcgggtaaa cgctccgcca aaatctcctt 7680
tgttattact gtcgtgcttt cacttctcgc aggagtcctc gtatggtaag caacgcctcc 7740
gcattaggac gcaatggcgt acatgatttc atcctcgttc gcgctaccgc tatcgtcctg 7800
acgtctaca tcatttatat ggtcgggttt ttcgctacca gtggcgagct gacatatgaa 7860

gtctggatcg gtttcttcgc ctctgcgttc accaaagtgt tcaccctgct ggcgctgttt 7920
tctatcttga tccatgcctg gatcggcatg tggcaggtgt tgaccgacta cgttaaaccg 7980
ctggctttgc gcctgatgct gcaactgggtg attgtcgttg cactgggtggg ttacgtgatt 8040
tatggattcg ttgtgggtgtg ggggtgtgtga tgaaattgcc agtcagagaa tttgatgcag 8100
ttgtgattgg tgccgggtggc gcaggtatgc gcgcggcgct gcaaatttcc cagagcggcc 8160
agacctgtgc gctgctctct aaagtcttcc cgaccctgtc ccataccgtt tctgcgcaag 8220
gcggcattac cgttgcgctg ggtaataccc atgaagataa ctgggaatgg catatgtacg 8280
acaccgtgaa agggtcggac tatatcgggtg accaggacgc gattgaatat atgtgtaaaa 8340
ccggggccgga agcgattctg gaactcgaac acatgggcct gccgttctcg cgtctcgatg 8400
atggtcgtat ctatcaacgt ccgtttggcg gtcagtcgaa aaacttcggc ggcgagcagg 8460
cggcacgcac tgcggcagca gctgaccgta ccggtcacgc actgttgcaac acgctttatc 8520
agcagaacct gaaaaaccac accaccattt tctccgagtg gtatgcgctg gatctgggtga 8580
aaaaccagga tggcgcgggtg gtgggttgta ccgcactgtg catcgaaacc ggtgaagtgg 8640
tttatttcaa agcccgcgt accgtgctgg cgactggcgg agcagggcgt atttatcagt 8700
ccaccaccaa cgccacatt aacaccggcg acggtgtcgg catggctatc cgtgccggcg 8760
taccggtgca ggatatggaa atgtggcagt tccacccgac cggcattgcc ggtgcggggcg 8820
tactggtcac cgaaggttgc cgtggtgaag gcggttatct gctgaacaaa catggcgaac 8880
gttttatgga gcgttatgcg ccgaacgcca aagacctggc gggccgtgac gtggttgcg 8940
gttccatcat gatcgaaatc cgtgaaggtc gcggctgtga tggcccggtg gggccacacg 9000
cgaaactgaa actcgatcac ctgggtaaag aagttctcga atcccgtctg ccgggtatcc 9060
tggagctttc ccgtaccttc gctcacgtcg atccggtgaa agagccgatt ccggttatcc 9120
caacctgtca ctacatgatg ggcggtatcc cgaccaaagt taccggtcag gcactgactg 9180
tgaatgagaa aggcgaagat gtggttggtc cgggactggt tgccgttggt gaaatcgctt 9240
gtgtatcggg acacggcgct aaccgtctgg gcggcaactc gctgctggac ctggtggtct 9300
ttggtcgcgc ggcaggtctg catctgcaag agtctatcgc cgagcagggc gcactgcgcg 9360
atgccagcga gtctgatgtt gaagcgtctc tggatcgct gaaccgctgg aacaataatc 9420
gtaacgggtga agatccgggtg gcgatccgta aagcgtgca agaattgtatg cagcataact 9480
tctcggctct ccgtgaagggt gatgcgatgg cgaaagggct tgagcagttg aaagtgatcc 9540
gcgagcgtct gaaaaatgcc cgtctggatg acacttccag cgagttcaac acccagcgcg 9600

- 102 -

ttgagtgcct ggaactggat aacctgatgg aaacggcgta tgcaacggct gtttctgcca 9660
acttccgtac cgaaagccgt ggcgcgcata gccgcttcga cttcccggat cgtgatgatg 9720
aaaactggct gtgccactcc ctgtatctgc cagagtcgga atccatgacg cgccgaagcg 9780
tcaacatgga accgaaactg cgcccggcat tcccgccgaa gattcgtact tactaatgcg 9840
gagacaggaa aatgagactc gagttttcaa tttatcgcta taaccggat gttgatgatg 9900
ctccgcgtat gcaggattac accctggaag cggatgaagg tcgcgacatg atgctgctgg 9960
atgcgcttat ccagctaaaa gagaaagatc ccagcctgtc gttccgccgc tcctgccgtg 10020
aagggtgtgtg cggttccgac ggtctgaaca tgaacggcaa gaatggtctg gcctgtatta 10080
ccccgatttc ggcactcaac cagccgggca agaagattgt gattcgcccgc ctgccagggt 10140
taccggtgat ccgcgatttg gtggtagaca tgggacaatt ctatgcgcaa tatgagaaaa 10200
ttaagcctta cctggtgaat aatggacaaa atccgccagc tcgcgagcat ttacagatgc 10260
cagagcagcg cgaaaaactc gacgggctgt atgaatgtat tctctgcgca tgttgttcaa 10320
cctcttgtcc gtctttctgg tggaatcccgc ataagtttat cggcccggca ggcttgtag 10380
cggcatatcg tttcctgatt gatagccgtg ataccgagac tgacagccgc ctgcacgggt 10440
tgagtgatgc attcagcgta ttccgctgtc acagcatcat gaactgcgtc agtgtatgtc 10500
cgaaggggct gaacccgacg cgcgccatcg gccatatcaa gtcgatgttg ttgcaacgta 10560
atgcgtaaac cgtaggcctg ataagacgcg caagcgtcgc atcaggcaac cagtgccgga 10620
tgccggcgtga acgccttata' cggcctacaa gtcattaccg gtaggcctga taagcgcagc 10680
gcatcaggcg taacaaagaa atgcaggaaa tctttaaaaa ctgcccctga cactaagaca 10740
gttttttaaag gttccttcgc gagccactac gtagacaaga gctcgcaagt gaaccccggc 10800
acgcacatca ctgtgcgtgg tagtatccac ggcgaagtaa gcataaaaaa gatgcttaag 10860
ggatcacgat gcagaacagc gctttgaaag cctgggttga ctcttcttac ctctctggcg 10920
caaaccagag ctggatagaa cagctctatg aagacttctt aaccgatcct gactcggttg 10980
acgctaactg gcgttcgacg ttccagcagt tacctggtac gggagtcaaa ccggatcaat 11040
tccactctca aacgcgtgaa tatttccgcc gcctggcgaa agacgcttca cgttactctt 11100
caacgatctc cgaccctgac accaatgtga agcagggtta agtcctgcag ctcatthaacg 11160
cataccgctt ccgtggtcac cagcatgcga atctcgatcc gctgggactg tggcagcaag 11220
ataaagtggc cgatctggat ccgtctttcc acgatctgac cgaagcagac ttccaggaga 11280
ccttcaacgt cggttcattt gccagcggca aagaaacat gaaactcggc gagctgctgg 11340
aagccctcaa gcaaacctac tgccggcccga ttggtgccga gtatatgcac attaccagca 11400

ccgaagaaaa acgctggatc caacagcgta tcgagtctgg tcgcgcgact ttcaatagcg 11460
aagagaaaaa acgcttctta agcgaactga ccgccgctga aggtcttgaa cgttacctcg 11520
gcgcaaaatt ccctggcgca aaacgcttct cgctggaagg cggtgacgcg ttaatcccga 11580
tgcttaaaga gatgatccgc cacgctggca acagcggcac ccgcgaagtg gttctcggga 11640
tggcgcaccg tggtcgtctg aacgtgctgg tgaacgtgct gggtaaaaaa ccgcaagact 11700
tgttcgacga gttcgccggt aaacataaag aacacctcgg cacgggtgac gtgaaatacc 11760
acatgggctt ctcgctctgac ttccagaccg atggcggcct ggtgcacctg gcgctggcgt 11820
ttaaccgcgc tcaccttgag attgtaagcc cggtagttat cggttctgtt cgtgccgcgc 11880
tggacagact tgatgagccg agcagcaaca aagtgtgcc aatcaccatc cacggtgacg 11940
ccgcagtgac cgggcagggc gtggttcagg aaaccctgaa catgtcgaaa gcgcgtgggt 12000
atgaagttgg cggtagcgta cgtatcgta tcaacaacca ggttggtttc accacctcta 12060
atccgctgga tgcccgttct acgccgtact gtactgatat cggtaagatg gttcaggccc 12120
cgattttcca cgttaacgcg gacgatccgg aagccgttgc ctttgtgacc cgtctggcgc 12180
tcgatttccg taacaccttt aaacgtgatg tcttcatcga cctgggtgtg taccgccgcgc 12240
acggccacaa cgaagccgac gagccgagcg caaccagcc gctgatgtat cagaaaatca 12300
aaaaacatcc gacaccgcgc aaaatctacg ctgacaagct ggagcaggaa aaagtggcga 12360
cgctggaaga tgccaccgag atggttaacc tgtaccgcga tgcgctggat gctggcgatt 12420
gcgtagtggc agagtggcgt ccgatgaaca tgcactcttt cacctggtcg ccgtacctca 12480
accacgaatg ggacgaagag taccgaaca aagttgagat gaagcgctg caggagctgg 12540
cgaaacgcat cagcacggtg ccggaagcag ttgaaatgca gtctcgcgtt gccaaagattt 12600
atggcgatcg ccaggcgatg gctgccggtg agaaactgtt cgactggggc ggtgcggaaa 12660
acctcgctta cgccacgctg gttgatgaag gcattccggt tcgcctgtcg ggtgaagact 12720
ccggtcgcgg taccttcttc caccgccacg cggatgatca caaccagtct aacggttcca 12780
cttacacgcc gctgcaacat atccataacg ggcagggcgc gttccgtgtc tgggactccg 12840
tactgtctga agaagcagtg ctggcgtttg aatatggtta tgccaccgca gaaccacgca 12900
ctctgaccat ctgggaagcg cagttcgggtg acttcgcaa cggtgccgag gtggttatcg 12960
accagttcat ctctctggc gaacagaaat ggggccggat gtgtggtctg gtgatgttgc 13020
tgccgcacgg ttacgaaggg caggggccgg agcactctc cgcgcgctct gaacgttatc 13080
tgcaactttg tgctgagcaa aacatgcagg tttgcgtacc gtctaccccgc gcacaggttt 13140

- 104 -

accacatgct gcgctgctcag gcgctgcgcg ggatgcgtcg tccgctggtc gtgatgtcgc 13200
cgaaatccct gctgcgtcat ccgctggcgg ttccagcct cgaagaactg gcgaacggca 13260
ccttcctgcc agccatcggg gaaatcgacg agcttgatcc gaagggcgtg aagcgcgtag 13320
tgatgtgttc tggtaagggt tattacgacc tgctggaaca gcgtcgtaag aacaatcaac 13380
acgatgtcgc cattgtgcgt atcgagcaac tctaccggtt cccgcataaa gcgatgcagg 13440
aagtgttgca gcagtttgct cacgtcaagg attttgtctg gtgccaggaa gagccgctca 13500
accagggcgc atggtactgc agccagcatc atttccgtga agtgattccg tttggggcctt 13560
ctctgcgtta tgcaggccgc ccggcctccg cctctccggc ggtaggggat atgtccgttc 13620
accagaaaca gcaacaagat ctggttaatg acgcgctgaa cgtcgaataa ataaaggata 13680
cacaatgagt agcgtagata ttctgggtccc tgacctgcct gaatccgtag ccgatgccac 13740
cgtcgcaacc tggcataaaa aaccggcga cgcagtcgta cgtgatgaag tgctggtaga 13800
aatcgaaact gacaaagtgg tactggaagt accggcatca gcagacggca ttctggatgc 13860
ggttctggaa gatgaaggta caacggtaac gtctcgtcag atccttggtc gcctgcgtga 13920
aggcaacagc gccggtaaag aaaccagcgc caaatctgaa gagaaagcgt ccactccggc 13980
gcaacgccag caggcgtctc tggaaagagca aaacaacgat gcgttaagcc cggcgatccg 14040
tcgcctgctg gctgaacaca atctcgacgc cagcgccatt aaaggcaccg gtgtgggtgg 14100
tcgtctgact cgtgaagatg tggaaaaaca tctggcgaaa gccccggcga aagagtctgc 14160
tccggcagcg gctgctccgg cggcgcaacc ggctctggct gcacgtagtg aaaaacgtgt 14220
cccgatgact cgctgcgtga agcgtgtggc agagcgtctg ctggaagcga aaaactccac 14280
cgccatgctg accacgttca acgaagtcaa catgaagccg attatggatc tgcgtaagca 14340
gtacggtgaa gcgtttgaaa aacgccacgg catccgtctg ggctttatgt ccttctacgt 14400
gaaagcgggtg gttgaagccc tgaaacgtta ccggaagtg aacgcttcta tcgacggcga 14460
tgacgtgggt taccacaact atttcgacgt cagcatggcg gtttctacgc cgcgcggcct 14520
ggtgacgccg gttctgcgtg atgtcgatac cctcggcatg gcagacatcg agaagaaaat 14580
caaagagctg gcagtcaaag gccgtgacgg caagctgacc gttgaagatc tgaccggtgg 14640
taacttcacc atcaccaacg gtggtgtgtt cggttccctg atgtctacgc cgatcatcaa 14700
cccgccgcag agcgcaattc tgggtatgca cgctatcaaa gatcgtccga tggcggtgaa 14760
tggtcagggt gagatcctgc cgatgatgta cctggcgctg tcctacgatc accgtctgat 14820
cgatggtcgc gaatccgtgg gcttcctggg aacgatcaaa gagttgctgg aagatccgac 14880
gcgtctgctg ctggacgtgt agtagtttaa gtttcacctg cactgtagac cggataaggc 14940

- 105 -

attatcgctt tctccggcaa ttgaagcctg atgcgacgct gacgcgtctt atcaggccta 15000
 cgggaccacc aatgtaggtc ggataaggcg caagcgccgc atccgacaag cgatgcctga 15060
 tgtgacg 15067

<210> 25

<211> 427

<212> PRT

<213> Escherichia coli

<220>

<223> complement of position 5348-6631 of seq id 24

<400> 25

Met Ala Asp Thr Lys Ala Lys Leu Thr Leu Asn Gly Asp Thr Ala Val
 1 5 10 15

Glu Leu Asp Val Leu Lys Gly Thr Leu Gly Gln Asp Val Ile Asp Ile
 20 25 30

Arg Thr Leu Gly Ser Lys Gly Val Phe Thr Phe Asp Pro Gly Phe Thr
 35 40 45

Ser Thr Ala Ser Cys Glu Ser Lys Ile Thr Phe Ile Asp Gly Asp Glu
 50 55 60

Gly Ile Leu Leu His Arg Gly Phe Pro Ile Asp Gln Leu Ala Thr Asp
 65 70 75 80

Ser Asn Tyr Leu Glu Val Cys Tyr Ile Leu Leu Asn Gly Glu Lys Pro
 85 90 95

Thr Gln Glu Gln Tyr Asp Glu Phe Lys Thr Thr Val Thr Arg His Thr
 100 105 110

Met Ile His Glu Gln Ile Thr Arg Leu Phe His Ala Phe Arg Arg Asp
 115 120 125

Ser His Pro Met Ala Val Met Cys Gly Ile Thr Gly Ala Leu Ala Ala
 130 135 140

Phe Tyr His Asp Ser Leu Asp Val Asn Asn Pro Arg His Arg Glu Ile
 145 150 155 160

Ala Ala Phe Arg Leu Leu Ser Lys Met Pro Thr Met Ala Ala Met Cys
 165 170 175

Tyr Lys Tyr Ser Ile Gly Gln Pro Phe Val Tyr Pro Arg Asn Asp Leu
 180 185 190

Ser Tyr Ala Gly Asn Phe Leu Asn Met Met Phe Ser Thr Pro Cys Glu
 195 200 205

Pro Tyr Glu Val Asn Pro Ile Leu Glu Arg Ala Met Asp Arg Ile Leu
 210 215 220

- 106 -

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

<210> 26

<211> 12354

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (10196)..(11146)

<400> 26

gacagcgcgt tttgggctac gccggaaaat ttgccaacaa tttaccgcaa gccgcgcgtc 60
 atgtacatgg aacatccttt tgccgcttca gaaatctctg gatcatgctc gcatgttgcg 120
 caatctactc gcccgctccgc tgcgcttttc cttatactga gactgagcgt cgattcacct 180
 gcaaacggcg cattttttaga ataatcctga ccttgtgctg aagagaaaac atgaaaattc 240
 gcgccttatt ggtagcaatg agcgtggcaa cggtactgac tggttgccag aatatggact 300

- 107 -

ccaacggact gctctcatca ggagcggaag cttttcaggc ttacagtttg agtgatgcgc 360
aggtgaaaac cctgagcgat caggcatgtc aggagatgga cagcaaggcg acgattgcgc 420
cagccaatag cgaatacgct aaacgtctga caactattgc caatgcgcta ggcaacaata 480
tcaacgggtca gccggtaaata taaaaagtgt atatggcgaa ggatgtgaac gcctttgcaa 540
tggctaacgg ctgtatccgc gtctatagcg ggctgatgga tatgatgacg gataacgaag 600
tcgaagcggg gatcgggtcac gaaatggggc acgtggcggt aggccatgtg aaaaaaggaa 660
tgcagggtggc acttggtaca aatgccgtgc gagtagctgc ggctctgcg ggcgggattg 720
tcggaagttt atctcaatca caacttggtg atctgggcga gaaattagtc aattcgcaat 780
tctcccagcg ccaggaagca gaagccgatg attattctta cgatcttctg cgccaacgcg 840
gcatcagccc ggcagggtctt gccaccagct ttgaaaaact ggcaaaactg gaagaaggtc 900
gccaaagctc aatgtttgac gaccatcctg catccgccga acgcgccag catattcgcg 960
atcgcatgag cgcggatggg attaagtaaa gcctgggtgg gtggaacgta ggtcagataa 1020
ggcgttcacg ccgcatctga catgaatggc acatttggtt ccttggtgcg attgccggat 1080
gogatgctgg cgcattctat ccggcctacg ggtgccgaac gtaggtcgga taaggcggtc 1140
acgcgcgcatc cgacattaat ggcacgtttt acccgtgcgc atcgcatctg gtgcttactc 1200
gccctttttc gccgcctgaa tatacagcat ttccagcgcc agcgttgcg ctgccagagc 1260
agtgatttcc gactgatcgt atgccggagc cacttccact acgtccatcc caacaatgtt 1320
gagatctttc aggcgcgta ccagtttaat agcgcgatcg gaggtcaggc cgccaatcac 1380
tggcgtaccg gtgcctgggtg caaaagcagg atccaggcag tcgatatcaa aagtcaggta 1440
aaccggcata tcaccacaaa tctgtttcac ttgggcgata acgtcatcca cgctgcgac 1500
gttcacctgg caggcgtcca gcacggtaaa gccgttgtct ttatcaaact cggtacgaat 1560
accaatctgc acggaatgat tcgggtcgat cagaccttct ttccggcgcg tatagaacat 1620
agtgccgtgg tcaaattcac aaccgttcgc ataggtatcg gtgtgggcgt caaagtgtac 1680
cagcgccatt ttgccgaaat gcttcgcatg agcacgcagc agcggcagcg taacaaagtg 1740
gtcaccaccg aaagagagca tacgcttacc ggcagccagc agcttctcgg cgtgcgcctg 1800
cagcttttcg ctcatctcac gggcatcgcc aaaggcatat accagatcgc cgcagtccac 1860
gacgttcaga cgctcacgca tgtcgaaatt ccacgggaag cggttggtgtt ccagggccag 1920
attcgtcgaa acctgacgga tcgctgccgg accgtggcga ccaccgcac gaccagaagt 1980
ggccatatcg aacggcacgc cagtaatcac ccagtctgca tcgctgtcat acggctggaa 2040

- 108 -

gttcatcggc aggcgtaaaa aaccaaaggc attggaaacc agtgagttat cgtattgatg 2100
acctaagggtg ctcatggatg gactcctatt acaaagtcga tataaaaaaa cccttccgcg 2160
tcgttaggcc cgacgaggaa gggttggatt tgtcacaata aattgtggcg gattatcacc 2220
gctaaattaa gcggattcaa gtaacacagg acttactcat cttcaagata agtataaccg 2280
tacaacctg cctcgaactc ttcaaggaac tgttggtgca gttcagcatc aagatcgggtt 2340
ttcttcactt gatcgcggaa ctgggttaac agcgttttcg gatcagagctg tacatattgc 2400
agcatgtccg ccacgggtatc gccttcgtca gacagttcta cttctacgct accgtcaggg 2460
aagacgaaca cgtcaaccgc ttcggtatca ccgaacagggt tgtgcatgtt gccgaggatc 2520
tcctgatatg cgccgaccat aaagaaaccg agcatcggcg gattctctgg atcgtactec 2580
ggcattggca ttgtcgtggc aataccgtca ccatcaatat agtggtcgat agcaccgta 2640
gagtcacagg taatatccag cagcacagcg cgacgttccg gcacttgatc cagcccttc 2700
agcggcagaa ccgggaacaa ctggtcgatc ccccatgcgt ccggcatcga ctggaacagc 2760
gagaagttga cgtacatfff gtccgccata cgttcctgca gtcgctcgat aatcggacga 2820
tgagcacgggt tttgcggatc cagctgcttt tgcacttcat ggcacatgct caaataaagc 2880
tgctcagccc atgcacgttc ttgcaggcta aagatgccgg aagagtagcc gatatgaatg 2940
tcgtgcagat ccatctgact gtcgtgtaac cattcacgca gagaacggcg agttcccgggt 3000
tcgtgcatct cctgccagggt ttcccacatg ctttgcagcg cgcgcggcgc atcttctgca 3060
ggcgcggtcg gcaccgtgta ttcgttacgt tccacgccga tgatattaga caccagcacg 3120
gtgtgatgcg cagtcaccgc acgaccgat tcggtgatta ccgtcggatg cggcagaccg 3180
ttttcttcac acgcatcgcc aatcgcccag ataatgttgt tggcgtattc attgaggccg 3240
tagttcaccg aacagtcgga ctgcgaacga gtaccttcat aatccacgcc cagaccgccg 3300
ccgacgtcga agcactgaat attgacgcc agcttgtgca gttccacata gaaacgcgcg 3360
gattcacgaa cgctgtcgc gatatcgca atattcgcca tctgcgaacc gaggtggaag 3420
tgcagtagtt gcaggctgtc gagacgcccg gcttcacgca gggtttcaac cagttgcagt 3480
acctgagtcg cagccaggcc gaacttcgat ttttccccgc cggaggactg ccatttacc 3540
gaaccctgcg aagccagacg tgcacgcacg cccagacgag gaacgacatt cagacgttct 3600
gcttcatcca gcacaatggc gatttctgac atcttctcaa tgaccagata gacctgtgc 3660
cccatcttct cgccaattaa tgccaggcgg atatattcgc ggtctttata accgttgcag 3720
acgatgacgc tacgggtcat gccagcatgt gccagtactg ccatcaactc ggctttggaa 3780
ccggcttcca gaccagcgg ttcccccga tgaatcaggg actcaatcac gcggcgggtgc 3840

- 109 -

tggttaactt tgatcggata aacaaggaag taatcgccgt tatagccgta ggattccctc 3900
gcacgtttga acgcggcggtt aatggaacgc aaacgggtgct gcaggatctg tgggaaacag 3960
aacagtgcag gcagacgctg gccctgtgct tcacgagttt tcactaactg cgcgagatcg 4020
acgcgagctt ccgggacgctc cgggtccggg cacacgctaa tgtggcccag ctcgttaacg 4080
tcatagtagt tattgccccca ccaggcaata ttgtaagtac gcagcatctt gctggcttcc 4140
tgggagctca ttgcaacctc ctgcatggag cgtagtacac cgtgttcgcc cgctgacgaa 4200
ggcaaaccce tagacatgtc gtcagacata gcgaacctca aattatttta ttaagtgtaa 4260
aacagttaac gactatcgca gcctggaact gcgataacaa cccataagca catggatttt 4320
ccagcagtga atgctgacgc tccaactgcg caaccggttt ctttttcata acattattaa 4380
gcacataacc gaacgtaagt gtgaaagttc ggcgaaacca cgagaaaact cttgttttta 4440
caagagcgcc cttgttcagt cctcagtaac tgtaaccagc tcttgaatcc tgagaagcgc 4500
cgagatgggt ataacatcgg caggtatgca aagcagagat gcagagtgcg gggaacgaat 4560
cttcaccaga acggtgagac aggttaagca gcagacaacg gttcattatt tcgtatcacc 4620
tccacggccg cctgttaaga cgaaccacaca agccaaaact ctgatttcaa cccggctgga 4680
agtggcaaca cgaaagaaac gtcgtgtgct ttttatattaa gccgcgcgcc gcgttttata 4740
ccccacaatg gcagaaaatt gcaaaagata aatacgcaga atgccggcat tgtcaggaaa 4800
aatttccagc cacgttttta acagaatgag acacgattca aaaaaaagtg gaaatagggt 4860
gaagaattga cctaaaatag ccatccagat gttaatccat ccataccgat taacactcag 4920
actgccagtg tttttaacct gcagagtcgt ggtaggatcc gctaccacag aaaatccaca 4980
caacagtttg agctaaccac attctcttta ggtgatatta aatatggcaa aacacctttt 5040
tacgtccgag tccgtctctg aagggcatcc tgacaaaatt gctgaccaa tttctgatgc 5100
cgttttagac gcgacccctg aacaggatcc gaaagcacgc gttgcttgcg aaacctacgt 5160
aaaaaccggc atggtttttag ttggcggcga aatcaccacc agcgcctggg tagacatcga 5220
agagatcacc cgtaacaccg ttcgcgaaat tggctatgtg cattccgaca tgggctttga 5280
cgctaactcc tgtgcggttc tgagcgctat cggcaaacag tctcctgaca tcaaccaggg 5340
cgttgaccgt gccgatccgc tggaacaggg cgcgggtgac cagggtctga tgtttggcta 5400
cgcaactaat gaaaccgacg tgctgatgcc agcacctatc acctatgcac accgtctggt 5460
acagcgtcag gctgaagtgc gtaaaaacgg cactctgccg tggctgcgcc cggacgcgaa 5520
aagccagggt acttttcagt atgacgacgg caaatcggtt ggtatcgatg ctgtcgtgct 5580

- 110 -

ttccactcag cactctgaag agatcgacca gaaatcgctg caagaagcgg taatggaaga 5640
gatcatcaag ccaattctgc ccgctgaatg gctgacttct gccaccaa at tcttcatcaa 5700
cccgaccggt cgtttcgtta tcggtggccc aatgggtgac tgcggtctga ctggtcgtaa 5760
aattatcggt gatacctacg gcggcatggc gcgtcacggt ggcggtgcat tctctggtaa 5820
agatccatca aaagtggacc gttccgcagc ctacgcagca cgttatgtcg cgaaaaacat 5880
cgttgctgct ggccctggccg atcgttgtga aattcagggt tcctacgcaa tcggcgtggc 5940
tgaaccgacc tccatcatgg tagaaacttt cgg tactgag aaagtgcctt ctgaacaact 6000
gaccctgctg gtacgtgagt tcttcgacct gcgcccatac ggtctgattc agatgctgga 6060
tctgctgcac ccgatctaca aagaaaccgc agcatacggc cactttggtc gtgaacat 6120
cccgtgggaa aaaaccgaca aagcgcagct gctgcgcgat gctgccggtc tgaagtaatc 6180
tttcttcacc tgcgttcaaa ggccagcctc gcgctggcct ttttcttttg gataggcggt 6240
cacgccgat ccggcaaaaa aaccgcccgc acaataacat cattcttcct gatcacgttt 6300
caccgcagat tatcatcaca actgaaaccg attacaccaa ccacaacaga caaagatttg 6360
taatattttc atattattat tcggttttca cagttgttac atttcttttc agtaaagtct 6420
taattgcaga taacagcggt taatctatga tgatataact caattatttt catgcactta 6480
aatcataact aagataaatg ttagtgtaag cgattacact gatgtgattt gcttcacatc 6540
tttttacgtc gtactcacct atcttaattc acaataaaaa ataaccatat tggagggcat 6600
catgcctgac gctaaaaaac aggggcggtc aaacaaggca atgacgtttt tcgtctgctt 6660
ccttgccgct ctggcgggat tactcttttg cctggatatc ggtgtaattg ctggcgact 6720
gccgtttatt gcagatgaat tccagattac ttgcacacg caagaatggg tcgtaagctc 6780
catgatgttc ggtgcggcag tcggtgcggt gggcagcggc tggtctctct ttaaaactcg 6840
gcgcaaaaag agcctgatga tcggcgcaat tttgtttgtt gccggttcgc tgttctctgc 6900
ggctgcgcca aacgttgaag tactgattct ttcccgcgtt ctactggggc tggcggtg 6960
tgtggcctct tataccgcac cgctgtacct ctctgaaatt gcgccgaaa aaattcgtgg 7020
cagtatgatc tcgatgtatc agttgatgat cactatcggg atcctcggtg cttatctttc 7080
tgataccgcc ttcagctaca ccggtgcatg gcgctggatg ctgggtgtga ttatcatccc 7140
ggcaattttg ctgctgattg gtgtcttctt cctgccagac agcccacgtt ggtttgccgc 7200
caaacgccgt tttgttgatg ccgaacgcgt gctgctacgc ctgcgtgaca ccagcgcgga 7260
agcgaaacgc gaactggatg aaatccgtga aagtttgcag gttaaacaga gtggctgggc 7320
gctgtttaaa gagaacagca acttccgccg cgcggtgttc cttggcgtag tgttgagg 7380

- 111 -

aatgcagcaa ttcaccggga tgaacgtcat catgtattac gcgccgaaaa tcttcgaact 7440
ggcgggttat accaacta cccagcaaat gtgggggacc gtgattgtcg gcctgaccaa 7500
cgtacttgcc acctttatcg caatcggcct tggtgaccgc tggggacgta aaccaacgct 7560
aacgctgggc ttcctggtga tggctgctgg catgggcgta ctcggtacaa tgatgcatat 7620
cggattcac tctccgtcgg cgcagtattt cgccatcgcc atgctgctga tgtttattgt 7680
cggttttgcc atgagtgccg gtccgctgat ttgggtactg tgctccgaaa ttcagccgct 7740
gaaaggccgc gattttggca tcacctgctc cactgccacc aactggattg ccaacatgat 7800
cgttggcgca acgttcctga ccatgctcaa cacgctgggt aacgccaaca ccttctgggt 7860
gtatgcggct ctgaacgtac tgtttatcct gctgacattg tggctggtac cggaaaccaa 7920
acacgtttcg ctggaacata ttgaacgtaa tctgatgaaa ggctcgtaaac tgcgcgaaat 7980
aggcgtcac gattaatctc cccaagcttc ctcccatcgc ggaggaagcc acctcttgca 8040
gtcatctttt ctctgctcta tctctgccc ctatgaaaac atcccgtctc cctatcgcca 8100
tccaacaggc cgttatgcgt cgcctgcggg aaaaactcgc ccaggccaac ctgaagctag 8160
ggcgtaacta cccggagcca aaactctctt acaccagcg cggaacctcc gccggaacgg 8220
cctggctgga aagctatgaa attcgcctca atcccgtttt gctggttgaa aacagtgaag 8280
cttttattga agaagtggta ccgcacgaac tggcacattt gctggtatgg aaacatttcg 8340
gccgcgtagc gccacatggc aaagagtgga agtggtatgat ggaaaacgtg ctgggtgttc 8400
ccgccgctcg tacgcatcag ttcgaactgc aatccgtcgc tcgcaacacc ttcccctacc 8460
gctgcaagtg ccaggagcat cagcttaccg tacgccgcca taatcgcgta gttcgtggcg 8520
aggccgtcta tcgctgtgtt cactgcggtg aacagctggt tgcgaaataa ccatctgaac 8580
tatcaggaac tttcctgatc tggctgattg cataccaaaa cagctttcgc tacgttgctg 8640
gctcgtttta acacggagta agtgatgtac cgttatttgt ctattgctgc ggtggtactg 8700
agcgcagcat tttccggccc ggcgttgccc gaaggatatca atagtttttc tcaggcgaaa 8760
gccgcggcgg taaaagtcca cgctgacgcg ccgggtacgt tttattgcgg atgtaaaatt 8820
aactggcagg gcaaaaaagg cgttgttgat ctgcaatcgt gcggctatca ggtgcgcaaa 8880
aatgaaaacc gcgccagccg cgtagagtgg gaacatgtcg ttcccgctg gcagttcgggt 8940
caccagcgcc agtgctggca ggacggtgga cgtaaaaact gcgctaaaga tccggtctat 9000
cgcaagatgg aaagcgatat gcataacctg cagccgtcag tcggtgaggt gaatggcgat 9060
cgcggaact ttatgtacag ccagtggaaat ggcggtgaag gccagtacgg tcaatgcgcc 9120

- 112 -

atgaaggctcg atttcaaaga aaaagctgcc gaaccaccag cgcggtgcacg cgggtgccatt 9180
 gcgcgcacct acttctatat gcgcgaccaa tacaacctga cactctctcg ccagcaaacg 9240
 cagctgttca acgcatggaa caagatgtat ccggttaccg actgggagtg cgagcgcgat 9300
 gaacgcatcg cgaagggtgca gggcaatcat aaccctgatg tgcaacgcgc ttgccaggcg 9360
 cgaaagagct aacctacact agcgggattc tttttgttaa cccctacccc acgcgtacaa 9420
 ccgcgtgggg agacgacgcg gatttttaac tatgcgtatc ccccgcatth atcatcctga 9480
 accactgacc agccattctc acatcgcgct ttgcgaagat gccgccaacc atatcgggcg 9540
 cgtactgcgc atggggcccg ggcaggcggt gcaattgttt gacggtagca accaggtctt 9600
 tgacgccgaa attaccagcg ccagcaaaaa aagcgtggaa gtgaagggtgc tggaaggcca 9660
 gatcgacgat cgcgaatctc cgctgcatat tcacctcggt caggtgatgt cgcggtggtga 9720
 aaaaatggaa ttactatcc agaaatcgat cgaactcggt gtaagcctca ttacgccact 9780
 tttttctgag cgctgcggcg ttaaaactgga tagtgaacgt ctgaacaaga agcttcagca 9840
 gtggcagaag attgcaattg ctgcctgtga gcagtgtggt cgtaaccggg tgccggaaat 9900
 ccgtccagcg atggatctgg aagcctgggtg tgcagagcag gatgaaggac tgaaactgaa 9960
 tcttcacccg cgcgccagta acagcatcaa tacgttgccg ttaccgggtg aacgcgtccg 10020
 cctgctgatt ggcccgaag gcggtttatc ggcagatgaa attgccatga ctgcccgtca 10080
 tcaatttact gatatcctgt tgggacctcg cgttttgctg acagagacaa ctgcgctcac 10140
 cgccattacc gcgctacaag tacgatttgg cgatttgggc taacggagaa gaata atg 10198
 Met
 1

atc aag ctc ggc atc gtg atg gac ccc atc gca aac atc aac atc aag 10246
 Ile Lys Leu Gly Ile Val Met Asp Pro Ile Ala Asn Ile Asn Ile Lys
 5 10 15

aaa gat tcc agt ttt gct atg ttg ctg gaa gca cag cgt cgt ggt tac 10294
 Lys Asp Ser Ser Phe Ala Met Leu Leu Glu Ala Gln Arg Arg Gly Tyr
 20 25 30

gaa ctt cac tat atg gag atg ggc gat ctg tat ctg atc aat ggt gaa 10342
 Glu Leu His Tyr Met Glu Met Gly Asp Leu Tyr Leu Ile Asn Gly Glu
 35 40 45

gcc cgc gcc cat acc cgc acg ctg aac gtg aag cag aac tac gaa gag 10390
 Ala Arg Ala His Thr Arg Thr Leu Asn Val Lys Gln Asn Tyr Glu Glu
 50 55 60 65

tgg ttt tcg ttc gtc ggt gaa cag gat ctg ccg ctg gcc gat ctc gat 10438
 Trp Phe Ser Phe Val Gly Glu Gln Asp Leu Pro Leu Ala Asp Leu Asp
 70 75 80

gtg atc ctg atg cgt aaa gac ccg ccg ttt gat acc gag ttt atc tac 10486

Val	Ile	Leu	Met	Arg	Lys	Asp	Pro	Pro	Phe	Asp	Thr	Glu	Phe	Ile	Tyr	
			85					90					95			
gcg	acc	tat	att	ctg	gaa	cgt	gcc	gaa	gag	aaa	ggg	acg	ctg	atc	gtt	10534
Ala	Thr	Tyr	Ile	Leu	Glu	Arg	Ala	Glu	Glu	Lys	Gly	Thr	Leu	Ile	Val	
		100					105					110				
aac	aag	ccg	cag	agc	ctg	cgc	gac	tgt	aac	gag	aaa	ctg	ttt	acc	gcc	10582
Asn	Lys	Pro	Gln	Ser	Leu	Arg	Asp	Cys	Asn	Glu	Lys	Leu	Phe	Thr	Ala	
	115					120					125					
tgg	ttc	tct	gac	tta	acg	cca	gaa	acg	ctg	gtt	acg	cgc	aat	aaa	gcg	10630
Trp	Phe	Ser	Asp	Leu	Thr	Pro	Glu	Thr	Leu	Val	Thr	Arg	Asn	Lys	Ala	
130					135					140					145	
cag	cta	aaa	gcg	ttc	tgg	gag	aaa	cac	agc	gac	atc	att	ctt	aag	ccg	10678
Gln	Leu	Lys	Ala	Phe	Trp	Glu	Lys	His	Ser	Asp	Ile	Ile	Leu	Lys	Pro	
				150				155						160		
ctg	gac	ggg	atg	ggc	ggc	gcg	tcg	att	ttc	cgc	gtg	aaa	gaa	ggc	gat	10726
Leu	Asp	Gly	Met	Gly	Gly	Ala	Ser	Ile	Phe	Arg	Val	Lys	Glu	Gly	Asp	
			165					170					175			
cca	aac	ctc	ggc	gtg	att	gcc	gaa	acc	ctg	act	gag	cat	ggc	act	cgc	10774
Pro	Asn	Leu	Gly	Val	Ile	Ala	Glu	Thr	Leu	Thr	Glu	His	Gly	Thr	Arg	
		180					185					190				
tac	tgc	atg	gcg	caa	aat	tac	ctg	cca	gcc	att	aaa	gat	ggc	gac	aaa	10822
Tyr	Cys	Met	Ala	Gln	Asn	Tyr	Leu	Pro	Ala	Ile	Lys	Asp	Gly	Asp	Lys	
	195					200					205					
cgc	gtg	ctg	gtg	gtg	gat	ggc	gag	ccg	gta	ccg	tac	tgc	ctg	gcg	cgt	10870
Arg	Val	Leu	Val	Val	Asp	Gly	Glu	Pro	Val	Pro	Tyr	Cys	Leu	Ala	Arg	
210					215					220					225	
att	ccg	cag	ggg	ggc	gaa	acc	cgt	ggc	aat	ctg	gct	gcc	ggg	ggg	cgc	10918
Ile	Pro	Gln	Gly	Gly	Glu	Thr	Arg	Gly	Asn	Leu	Ala	Ala	Gly	Gly	Arg	
				230					235					240		
ggg	gaa	cct	cgt	ccg	ctg	acg	gaa	agt	gac	tgg	aaa	atc	gcc	cgt	cag	10966
Gly	Glu	Pro	Arg	Pro	Leu	Thr	Glu	Ser	Asp	Trp	Lys	Ile	Ala	Arg	Gln	
			245					250					255			
atc	ggg	ccg	acg	ctg	aaa	gaa	aaa	ggg	ctg	att	ttt	gtt	ggg	ctg	gat	11014
Ile	Gly	Pro	Thr	Leu	Lys	Glu	Lys	Gly	Leu	Ile	Phe	Val	Gly	Leu	Asp	
		260					265					270				
atc	atc	ggc	gac	cgt	ctg	act	gaa	att	aac	gtc	acc	agc	cca	acc	tgt	11062
Ile	Ile	Gly	Asp	Arg	Leu	Thr	Glu	Ile	Asn	Val	Thr	Ser	Pro	Thr	Cys	
	275					280					285					
att	cgt	gag	att	gaa	gca	gag	ttt	ccg	gtg	tcg	atc	acc	gga	atg	tta	11110
Ile	Arg	Glu	Ile	Glu	Ala	G										

- 114 -

cgagaaggat ctctgttgaga ctctgagtga cagcgccctt ctttccacgc atactgggcg 11216
 ctgttgcttt tttgaaccag gaaacagaac ctctgacaat gaatttacag catcactttc 11276
 ttattgccat gcctgctctc caggatccga ttttccgctcgtc ttccgtgggc tacatttgcg 11336
 aacataatac caatgggtgca atgggggatca tcgtcaacaa gccgctggaa aatctcaaaa 11396
 ttgaagggat tctggaaaag ctgaagatca cgccggagcc gcgtgatgaa tcaatccgtc 11456
 tggataaacc gggttatgctc ggcggtccgc tggctgaaga tcgcggggttt attttgcata 11516
 ctccgcccctc caattttgct tccagcattc gcatttcaga caacacggta atgaccactt 11576
 cccgcgatgt gctggaaacg ctcggcaccg ataaacaacc gtctgacgta ttgggtggctc 11636
 tgggttatgc ctctgggag aaaggtcaac tggaacaaga aattctcgat aacgcgtggc 11696
 taacggcccc ggcagatctg aatattctgt tcaaaacgcc gattgccgac cgctggcgcg 11756
 aggcgggcaaa actgattggc gtggatattc tcaccatgcc tgggtgtggca ggacacgcct 11816
 gatgagtggga accttactcg ccttcgactt cggcaccaaa agcattggcg tagcggtcgg 11876
 ccaacgcatt accggcaccg ctcgcccttt gcctgcaatt aaagcacagg acggtacgcc 11936
 ggactggaac attatcgagc gtttactgaa agagtggcag ccggacgaaa tcatcgtcgg 11996
 tttgccgctg aatatggacg gcaccgagca accattgact gccagagcgc gttaaatttgc 12056
 caaccgtatt catggccggt tcggtgttga agtaaagctc catgacgagc gtcttagcac 12116
 tgtggaagcc cgttccggtc tgtttgaaca gggcggtat cgggcgctca acaaaggcaa 12176
 agttgactct gcctctgcgg ttattattct cgaaagctat ttcgagcagg gatattaagg 12236
 cgatttaaac gcctggcgga gtgtaaataa tatcatcgcg tcttattgcc ggatgcggcg 12296
 tgaacacctt atccagcaca catctggcag cggctatagg tctgataaga cgcgttag 12354

<210> 27

<211> 316

<212> PRT

<213> Escherichia coli

<400> 27

Met Ile Lys Leu Gly Ile Val Met Asp Pro Ile Ala Asn Ile Asn Ile
 1 5 10 15

Lys Lys Asp Ser Ser Phe Ala Met Leu Leu Glu Ala Gln Arg Arg Gly
 20 25 30

Tyr Glu Leu His Tyr Met Glu Met Gly Asp Leu Tyr Leu Ile Asn Gly
 35 40 45

Glu Ala Arg Ala His Thr Arg Thr Leu Asn Val Lys Gln Asn Tyr Glu
 50 55 60

WO 01/70776

PCT/US01/07478

- 115 -

Glu Trp Phe Ser Phe Val Gly Glu Gln Asp Leu Pro Leu Ala Asp Leu
 65 70 75 80
 Asp Val Ile Leu Met Arg Lys Asp Pro Pro Phe Asp Thr Glu Phe Ile
 85 90 95
 Tyr Ala Thr Tyr Ile Leu Glu Arg Ala Glu Glu Lys Gly Thr Leu Ile
 100 105 110
 Val Asn Lys Pro Gln Ser Leu Arg Asp Cys Asn Glu Lys Leu Phe Thr
 115 120 125
 Ala Trp Phe Ser Asp Leu Thr Pro Glu Thr Leu Val Thr Arg Asn Lys
 130 135 140
 Ala Gln Leu Lys Ala Phe Trp Glu Lys His Ser Asp Ile Ile Leu Lys
 145 150 155 160
 Pro Leu Asp Gly Met Gly Gly Ala Ser Ile Phe Arg Val Lys Glu Gly
 165 170 175
 Asp Pro Asn Leu Gly Val Ile Ala Glu Thr Leu Thr Glu His Gly Thr
 180 185 190
 Arg Tyr Cys Met Ala Gln Asn Tyr Leu Pro Ala Ile Lys Asp Gly Asp
 195 200 205
 Lys Arg Val Leu Val Val Asp Gly Glu Pro Val Pro Tyr Cys Leu Ala
 210 215 220
 Arg Ile Pro Gln Gly Gly Glu Thr Arg Gly Asn Leu Ala Ala Gly Gly
 225 230 235 240
 Arg Gly Glu Pro Arg Pro Leu Thr Glu Ser Asp Trp Lys Ile Ala Arg
 245 250 255
 Gln Ile Gly Pro Thr Leu Lys Glu Lys Gly Leu Ile Phe Val Gly Leu
 260 265 270
 Asp Ile Ile Gly Asp Arg Leu Thr Glu Ile Asn Val Thr Ser Pro Thr
 275 280 285
 Cys Ile Arg Glu Ile Glu Ala Glu Phe Pro Val Ser Ile Thr Gly Met
 290 295 300
 Leu Met Asp Ala Ile Glu Ala Arg Leu Gln Gln Gln
 305 310 315

<210> 28

<211> 11574

<212> DNA

<213> Escherichia coli

<400> 28

ataaccctga ttaatgaatt attacgttta tcatgttaat tcatcattat tacatcatca 60

ttgtaaataa ttaaattaac ttccataaca ttaaaatatg tatccactga cgcttttttta 120

- 116 -

cataacgaag aattgaccat tttgtcctgt tgtgccttaa tgtaagtacc gtccacagcg 180
tgggacatac ttcaaggaac cttttgtgag tcaggcaacc agtatgcgaa aacgacaccg 240
atttaacagt cgcatgaccc gtatcgtact gctcatcagc tttatcttct tctttggccg 300
ttttatctac tcgtccgtcg gtgcctggca gcaccatcag agcaaaaaag aagctcagca 360
atccacactc tccgtcgaat caccggtaca acgttagcgg ttaccttctc cactttcaca 420
gaacataacg gcacttcgct gtccgatgct tttgctgttt gggattatca aagcggcaga 480
tattctttca tcttaaattt tacgtcttta tcctgactga tgtttatcct gtttggctgc 540
gaaataaata taaaattaat atatatgttg taatgatata tttttataaa ttattccctg 600
cgtgaatttt aataaattta atctatccct ttatacgcaa tacatttact ttctcttttt 660
gatgatctta aatgtcttat ttttcgtaat gtgtataaca aggaatagtg atgaaattta 720
aaaaatgtct tctgcctgtg gcaatgtag cgtcattcac tctggcagga tgccagtcaa 780
atgctgacga tcatgctgcc gatgtttatc aaaccgatca actgaatacc aaacaagaaa 840
ctaaaaccgt taatattatt tccattcttc ccgcaaaagt tgccgtagac aactcccaaa 900
ataaacggaa cgcacaagcc ttccggcgcgc ttattggcgc agtcgctggc ggtgttatcg 960
gccacaacgt cgggtctggc agcaattccg gaacgacggc aggtgcagtt ggcggcggag 1020
ctgtaggcgc ggcagcgggt tctatggtga atgataaaac cttagtggaa ggtgtttctt 1080
taacctataa ggaaggcacc aaagtgtata cctctacca ggtgggtaaa gagtgccagt 1140
ttacgacagg tttagccgtt gttattacca cgacgtataa cgaaacgcgt attcagccaa 1200
ataccaaatg tcctgaaaag agctaataat caggaggagt catgaagaaa gtttttcttt 1260
gcgccatctt agcctcctta agctatccgg ctatcgctc atcattgcag gatcaactct 1320
ctgctgtcgc agaagcggaa cagcaaggta aaaatgaaga gcaaaggcag catgacgaat 1380
gggtcgcgga gcgcaacagg gaaatccagc aagagaagca acgtcgcgca aatgcccgag 1440
ccgccgctaa caaaagagcg gcaacggcag cggcaaataa gaaagctcgt caggataaac 1500
tggacgccga agcctctgcg gacaaaaaac gcgatcaaag ttatgaagat gagctacgca 1560
gcttagagat tcagaaacaa aaactggcgc tggcgaaaga agaagccgc gttaagcgag 1620
aaaacgaatt tatcgatcag gaactgaagc acaaagctgc gcaaaccgat gtggtgcaat 1680
ctgaagctga cgccaacaga aatatgactg aaggcggtcg cgatctgatg aaaagcgtgg 1740
gcaaagcaga agagaacaaa tcggacagct ggtttaatta atcgatgtta gtaacttcaa 1800
tcctataatt cttgaagata aaaaaccctc tgtagtaaca gagggttttg ttcattcata 1860
gtgcagggtc aaatcattcc cactcaatgg tagctggcgg cttgccgctg atgtcataca 1920

- 117 -

ccacgcggga aataccgttc acttcattga taatgcgggtt ggaaacgcga ccgaggaaat 1980
cgtacggcag atgcgcccag tgtgcgggtca taaagtcgat ggtttcgaca gcacgcagag 2040
agacaacca gtcatactta cgaccatcgc ccattacgcc aacggaacgt accggcagga 2100
acacagtga cgcctggctg actttgtcgt acaggtccgc tttacgcagt tcttcaatga 2160
agatggcgctc agcacggcgc agcaggtcac agtactcttt cttcacttca ccagaacac 2220
gaacgccaag gcctgggtccc gggaacgggt gacggtacag catgtcgtac ggcaggccca 2280
gctccagacc aatcttacgc acttcgtctt tgaacagctc tttcagcgggt tcaaccaggc 2340
ccatcttcat ctctttcggc aggccgcca cgttgtggtg agatttgatg acgtgtgctt 2400
taccggttgc agacgcgcga gattcgataa cgtcagggtg gatggtgccc tgcgccagcc 2460
acttcacgtc ttccagtttc agcgcttctt catcgaatac ttcaacgaaa acgcgaccga 2520
tgattttacg ttttgcttcc ggatcgtttt cgccagccag cgctgacagg aagcgatctt 2580
ctgccggtac gtgaacaatg ttaagaccaa agtgatcgcc aaacatatcc agaacctgct 2640
ctgcttcggt gaggcgcagc aggccgttgt cgacgaatac gcaagtcagg tttttaccga 2700
tagcgcggtg cagcagcatt gcggttacgg aggaatccac accaccagag aggccgagga 2760
tgactttatc gtgcctacc tgctcgcgga tgcgagctac agcatcgctc ataatttttcg 2820
ctggcggtcca cagggttca cactggcaga tatcacgcac aaaacgctcc agcatgcgca 2880
taccctggcg ggtatgagtc acttccgggt ggaactgtac gccatagaag cgtttttctt 2940
cgtagccat aatggcaaac gggcagcttt cggtgctggc tacggtgatg aagtcggacg 3000
gaatagcgggt aactttatcg ccgtgggtca tccagacatc gagcagcgggt ttaccgtctg 3060
cggtcagcgc atcttcgata ccgcgaacca gtgcgctgtc gtttacgact tcaacctgcg 3120
cgtagccaaa ttcacgttcg ttagaggctt caacgtgacc gcccaactgc attgccatgg 3180
tctgcatgcc atagcaaacg ccgaataccg gtacgcctgc ttcaaagaca tactgcggcg 3240
cacgcggact gttttcttca gtagtacttt ccgggccgcc ggaaagaata atgccgcttg 3300
gattgaagtc acgaatttgt gcttctgtca catcccacgc ccacagttcg cagtaaacac 3360
ccagctcacg cacgcggcgc gcaaccagtt gagtgtactg agaaccgaag tccagaatga 3420
ggatgcgatg cttatgaatg ttttccgtca ttgacgctta ttccgaggca agtgaaacag 3480
ataatataaa tcgcccgcga tgaagtcggg cgaagagaat caggagccca gacggtagtt 3540
cggggactct ttagtaatgg tcacgtcgtg aacgtggctt tcctgaatgc ccgcaccgct 3600
gatacgtaca aactccgctt tagtacgcag ttcgtcgata gtaccacagc cggtcagacc 3660

- 118 -

catacaggag cgcaggccgc ccatctgctg gtgaatgata tctttcaggc gacctttata 3720
 ggctacgcga ccttcgatac cttccggcac cagtttgtcg gcagcgttat cgctctggaa 3780
 ataacggtca gaggaacctt tggacatcgc gcccagggaa cccataccac ggtaagattt 3840
 gtaagaacgg ccctggtaga gttcgaatttc acccggagat tcttcagtac ccgccagcat 3900
 ggaacctacc atcaccgcgc ttgcgccagc ggcgatagct ttggcgatgt cgccggagaa 3960
 gcgaataaccg ccatcagcga taaccggaat accggtgcct tccagggctt ctactgcgtc 4020
 agcaacagcg gtaatctgcg gaacaccgac gccagtcacg atacgagttg tacagataga 4080
 gccagggcca atgccgactt taaccgcact gcaaccagct tctgccagag cgcgtgcacc 4140
 tgcagctgtt gccacgttgc cgccgataat ttgcagatcc ggatatattag cacgggtttc 4200
 acggatacgt tgcagtacac cttctgagtg accgtgggag gagtcgatca gcagaacgtc 4260
 aacgcctgcg gcaaccagcg cgtcaacacg ctcttcgtta ccgcacctg cgccaaccgc 4320
 tgcaccaaca cgcagacggc cttgctcgtc ttacaggcg ttccggtttac gttccgcttt 4380
 ctggaagtct ttcacggtga tcatgccgat caggtggaat tcgtcatcaa ccaccagcgc 4440
 tttttcaacg cgtttttcgt gcatttttgc cagcaccact tcacgggctt caccttcacg 4500
 cacggtgacc agacgctctt tcggcgatcat gtaaaccgta accggctggg tcaggtcggg 4560
 aacaaaacgc acgtcacgac cggtgataat acccaccagt tcgttttctt cggtaacgac 4620
 cggatagccc gcaaaaccgt tacgctcggg cagttctttc acttcgcgca gcgtcgtggg 4680
 tggcagaaca gtctgcggat cagtcaccac accagattcg tgttttttca cacggcgaac 4740
 ttcttctgcc tggcgttcaa tggacatgtt tttgtggata aagccgatac cgccttcctg 4800
 agccagagca atagccaggc gcgcttccgt tacggatatc attgctgcgg aaagcatagg 4860
 gatattcaga cgaatagttt tcgtcagctg ggtgctgagg tcagcagtat tcggcagaac 4920
 ggtagagtga gcaggaacga ggagaacgtc gtcaaacgtc agagcttctt tagcgatacg 4980
 tagcatgggc aatatctcga ccagagtggg taataaatat tgccgcggca ttatacagag 5040
 cgtaaccgat tgcattctacc cttttttgca aaaaatgctt gctatccccg aagggcgggt 5100
 tactatcgac tgaataacct gctgatttag aatttgatct cgctcacatg ttaccttctc 5160
 aatcccctgc aatttttacc gttagtcgcc tgaatcaaac ggttcgtctg ctgcttgagc 5220
 atgagatggg acaggtttgg atcagcggcg aaatttctaa tttcacgcaa ccagcttccg 5280
 gtcactggta ctttacactc aaagacgaca ccgccaggt acgctgcgcg atgttccgca 5340
 acagcaaccg ccgggtgacc ttccgcccac agcatgggca acaagtttta gttcgcgcca 5400
 atattacgct ctacgagccg cgcggcgact accagataat cgttgagagt atgcagccgg 5460

- 119 -

ccggtgaagg gctgctgcaa cagaagtacg aacagctcaa agcgaagttg caggctgaag 5520
gtttgttcga tcagcaatac aaaaaaccac ttccctcccc tgcgcattgc gttggtgtga 5580
tcacctcaaa aaccggtgct gcgctacatg atattttgca tgtgttaaaa cgtcgcgatc 5640
cttctctgcc ggtgatcatc taccctgccg ccgttcaggg cgatgacgcg ccggggcaaa 5700
ttgttcgcgc cattgaactg gcgaatcagc gcaatgagtg cgacgtattg atcgtcgggc 5760
gcggcggcgg ttcgctggaa gatttatgga gttttaacga cgaacgcgta gcgcgggcga 5820
tttttaccag ccgcattccg gttgtcagcg ccgtcgggca tgagacggat gtgaccattg 5880
ccgattttgt tgccgatctg cgtgcgcaa cgccgtctgc cgccgctgaa gtagtgagcc 5940
gtaatcagca agagttactg cgccagggtgc aatcgaccgc tcaacggctg gagatggcga 6000
tggtattatta tctcgccaac cgcacacgtc gctttacgca aattcatcac cgattacagc 6060
aacagcatcc gcagctccgg ctggcacgcc agcaaaccat gcttgagcgc ctgcaaaagc 6120
gaatgagctt tgcgctggaa aatcaactta agcgtaccgg gcaacagcag cagcggttaa 6180
cacagcggct gaatcagcaa aatccacagc cgaagattca tcgcgcgcaa acgcgcattc 6240
agcaactgga atatcgttta gcagaaacct tgcgcgcaca gcttagcgcc acgcgtgaac 6300
gtttcggtaa tgcagtaacg cacctcgaag ccgtaagccc actgtcaacg ctggcgcgtg 6360
gatacagcgt tactactgct actgacggca atgtactgaa aaaagtgaag caagttaaag 6420
cgggtgaaat gctaaccaca cgtctggaag acggctggat agaaagtga gtaaaaaaca 6480
tccagccagt aaaaaaatcg cgtaaaaagg tgcattaagc cagcacgaat tcaacgcgtt 6540
ttttcgatat caagccatgc ccgtgctgac agaagtaatc tactgcacca caggctttta 6600
acacctgcaa cggttgatgg cagtctggac aaagagcttt catttcaata aattccccgc 6660
aggatcggca acgcgcattg ccgttatcct gatcaagcac atgctgacat tgtggacaat 6720
gcagttccat atgaattcct ccggtagcca ttccctctac attcatagag ggaatggcag 6780
ataaaatact tacggataat tatttatttt tcttgatgtg cttcatcaga cgcttacggt 6840
tacgcactctg ggttggcgtc aggggtgttac gcttatcgc atacgggttt tcccttctt 6900
tgaactgaat acgaatcggc gatccatta cgtccagcga ttgcggaag tagttcatca 6960
agtagcgctt gtaggaatca ggcaggctct tcacctgatt accgtgaatc accacaatcg 7020
gcgggttata accaccggcg tgggcatatt tcagcttcac acgacgaccg cgtaccagcg 7080
gcggttggtg atcttcaaca gccatcgtca tgatgcgcgt cagcatagag gtccccacac 7140
gacgggtgga gctgtcatac gcttcacgta ctgattcaaa caagttacca acaccactgc 7200

- 120 -

cgtgcaaggc agagataaag tgcacacgag caaaatcgat aaagcccaga cggaagtcca 7260
gcgttttcttt cacctgctct ttcacttctt gactcaggcc atcccactta ttcaccacaa 7320
tgacaagtga gcgcccacta ttgagaataa agcccagcag cgagagatcc tgatcggaaa 7380
taccttcgcg cgcatacatc actaacatca ccacgttggc gtcttcaatg gcctgcaacg 7440
ttttgattac ggagaatttc tctacagcat cggtgatttt gccgcgttta cgtacgccag 7500
cgggtgtcaat gagcacatac tcacgtccat cgcgttccat tgggatgtag atgctgtcac 7560
gcgtcgtgcc aggcattgtc taaacaacaa cgcgctcttc accaagaata cggttagtga 7620
gtgtagactt acctacgttc ggacgacca caatcgccag tttgatcggc agactttgcg 7680
ggtcgaagtc gtcttcctct tcttcttcgc cgttctcttc cgcttcaaata tgcgcccagt 7740
attcagcgtc ttcgtcgact tctcttgcg gtgcgagatc tccatccac ggcagcagca 7800
catgctccag cagacttaata acgccacgac cgtgagacgc ggcgatcggg tagatttcac 7860
ctaaaccaag cgagtagaaa tcaaccactg cctgatcggg atccagaccg tcagttttgt 7920
ttgccaccag gaaggtcggg ttttcacggg agcgcagatg tttggcaatc gcttcatctg 7980
ccggcatcag gcccgcgcg gcattccacca taaacagtac gacgtccgct tcttcaatcg 8040
ccagcagcga ctgttccgcc atgcggggtt ctacaccgtc ttctgtgcca tcaatcccgc 8100
cggatatcaat acagataaac tcacggcctt caatttccgc acgaccgtac ttacgggtcac 8160
gagtcagacc cgggaaatcc gcaaccagcg catctcgggt gcgagttaga cggttaaata 8220
acgtggattt tctacgtta gggcgcccga caagcgcgac cacagggtacc atgttttaaag 8280
cctcattttt ataatcatc agacaacgca cgctatattc gcgtcgttgt taaaaacagg 8340
aaaacggccc ctgtccagga gccgttttca aagtgaacga cagagacgat taacgtgtaa 8400
tagagtacac ggttccgtct tttgcctgga tcagcagttt gccgtcagcg gcaaccgggt 8460
cagtctggaa accggaacta tcaacttttt gctgggcaac gaaacgacca tcttcgacgt 8520
taatccagtg cagataacct tcaactgtcac cgaccaccag gttgccatta tacagcaccg 8580
gagaagtcag caggcgatgc agcagatcgc tttgtgtcca cagcgtaacg ccgccatcaa 8640
tggtcaacgc catcaccggg tcattttgat cgaccagata gatgcgattg ccgtcgacga 8700
tgaaatcatt caccgaaccc agttcgcgtt tccacataat ctgaccactg cgcagatcaa 8760
gcgccgtcag gttaccatta taggccagcg cgaaaacaac gccgttaaca acgacgggag 8820
tcgtgtcaac atcgctcaga cgggtcaattt cggtagaacc ggtcgcctgg gaaatacgt 8880
gctgccaaat catctggccc tgttccatca gcaactgcgt gacgcgacca ttatcgcccc 8940
ccacgacggc cgcacaaaaa gccgttgtcg gcgcagactc gccacgcaa gagagcgaag 9000

- 121 -

gcataatcgag gtttaactgtc catttgacag cgccgtcagc ttcgttcagc gcttgtaact 9060
gaccgttact ggtgtggatt aacaccagac cgtcgctgac caccgggagc gaaagtgtt 9120
caccgcgac tttagtttgc catgccacag taccatcgct ggtattcagc gcgtaaacct 9180
gcgccttttc gctgccaatg tagacatgcc caccagacac ggtcacaccg ccagaaagta 9240
atgcaggctc tttagagaac cagccatctt tctcggccag gctgacagac cagatttctt 9300
tgccatcatc cgcattcagc gcttttacta aaccagcgcg gtccgctgca tagacaacgt 9360
tgtccgccag tgccggatga agattggaat agaagttgcc aatgccgcta ccaacggaag 9420
tgctccacgc cgtgggtcggc gtaaaactggt tttcaacggt tggcaatggg gacatcttta 9480
ccacatcttc ttcgctgtta aacagcgaac agccgcttaa aagggtaacg gaaagcagtc 9540
ctggcagcag taatttacgc aattgcatcg ggtccctctc agatggacaa attattaatt 9600
ttcatctgca tcatttcgct cagtgccgga gtaacatcgc ttttcacgcc tgcttcccat 9660
gcactacgcg caccttgctt atcacctttg ctacgcaatg cttcaccacg caggtcggca 9720
acaatggcag cccacccttc acctttgatg gtatcaaggg ttttcagcgc ggcatcagcc 9780
tgcttgagct gtacctgaac gcgagcaaga cgcagattta tcacggcttt gagattttca 9840
tcgctcgtgt ctgccagccc ctggttgtaac tgggcggcag ctttctctag ttcatttttg 9900
tcaacaaatt gctgcgcaag ttccaaagaa gccagcgcac cataagtatt tttattttca 9960
gcagcaaatt tttccgcccgc cgggatgcta tccggtttgc cttcgctcac tgcggtaacc 10020
gcattttgat aggcaagaga agcggagcgt gcagaatcaa cctgatggct gttccagtag 10080
cgccagccaa tcagtgcgcc aacgccc aaa atcaccccaa cagccagtgc tttgccattt 10140
tcagcaaaaa agcgtttaac cgcttctacc tggtcgtttt cgttctcgta aatttccacg 10200
ctgtccttct ccttccttaa cccagtaacg tgccgcaaatg cgcggctacg ctatcctgcg 10260
caactgccgt ttgctcacca gagcgcgaat ccttcactac tgctgtgccg ttagccactt 10320
cagactcacc cagcaccaca gcaacgcggg caccctattt atcagcacgg gcaaactgtt 10380
tcttaaagtt gccgccgccg tggttggtca tcaatttcac gccgggtaat tcatcacgca 10440
gacgctcagc taatgccata gccgcagatt gtgtatcagc acctgaagcc accaggtata 10500
tatcgacaac aggatcggct ttaaattccg gattaacggc ctgtactaac aatacaagac 10560
gttcgaggcc catagcaaaa ccgacagccg gtgttgacg accgcccagt tgttccacaa 10620
gaccgtcata acgaccgcct gcacacacgg tgccctggga gccgagactg ttagtcaccc 10680
actcgaaaac ggtacgggtg tagtaatcca gaccacgcac cagacgctgg tttacgggtg 10740

- 122 -

aagcgatccc cgcgctctcc agcagtttgc acagaccggc aaaatgctca cgagattcct 10800
cgtccagata gtcacctaata gccggagcgt cgttgagaag cgcctgcact tccggatttt 10860
ttgaatccag cagcgcgagc gggtttagtgt acatgcggcg tttgcagtct tcgtccagct 10920
tttctttatg ctgctcaagg aatgccacca gcgcacgcgc gtaattggcg cgtgcttcca 10980
gcgaaccgat agagttcagc tcaagagtta cgtgctcgga aatacccagc gcgcgccacc 11040
agcgggagcgt gagcataatc agttcagcgt cgatatccgg accttgcaga ccgaaaactt 11100
cgcagcccaa ctgatggaac tgacgataac gccctttctg cggacgctcg tgacggaaca 11160
tcggcccgat ataccacaga cgctgttcct gattgtacag aagaccatgc tcgatgccgg 11220
cgcgtacaca gcccgccgtc ccttcagggc gcagagtcag gctgtcgcca ttgcgatact 11280
caaaggtgta catctctttt tcaaccacgt cggtgacttc accaatcgca cgtttgaata 11340
gcgggggtctg ctctacaatc ggcaagcggg tttcactgta accgtagctg ccgagcacgt 11400
ttttcagtggt gccttcaatg cgctgccaga tggccgtttc gccaggcagg taatcgttca 11460
tgccgcgaat ggcttgaatg ttttttgcca cgtttattct ctttctgaat ataaaaatga 11520
accctcaacg cttccctcaa tgtttcggga gccatgcggg ttcaatcata cacg 11574

<210> 29

<211> 488

<212> PRT

<213> Escherichia coli

<220>

<223> complement of position 3520-4986 of seq id 28

<400> 29

Met	Leu	Arg	Ile	Ala	Lys	Glu	Ala	Leu	Thr	Phe	Asp	Asp	Val	Leu	Leu
1				5				10					15		

Val	Pro	Ala	His	Ser	Thr	Val	Leu	Pro	Asn	Thr	Ala	Asp	Leu	Ser	Thr
			20					25					30		

Gln	Leu	Thr	Lys	Thr	Ile	Arg	Leu	Asn	Ile	Pro	Met	Leu	Ser	Ala	Ala
		35					40					45			

Met	Asp	Thr	Val	Thr	Glu	Ala	Arg	Leu	Ala	Ile	Ala	Leu	Ala	Gln	Glu
	50					55					60				

Gly	Gly	Ile	Gly	Phe	Ile	His	Lys	Asn	Met	Ser	Ile	Glu	Arg	Gln	Ala
65					70				75					80	

Glu	Glu	Val	Arg	Arg	Val	Lys	Lys	His	Glu	Ser	Gly	Val	Val	Thr	Asp
				85					90					95	

Pro	Gln	Thr	Val	Leu	Pro	Thr	Thr	Thr	Leu	Arg	Glu	Val	Lys	Glu	Leu
			100					105					110		

- 123 -

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

- 124 -

Gln Gln Met Gly Gly Leu Arg Ser Cys Met Gly Leu Thr Gly Cys Gly
 435 440 445

Thr Ile Asp Glu Leu Arg Thr Lys Ala Glu Phe Val Arg Ile Ser Gly
 450 455 460

Ala Gly Ile Gln Glu Ser His Val His Asp Val Thr Ile Thr Lys Glu
 465 470 475 480

Ser Pro Asn Tyr Arg Leu Gly Ser
 485

<210> 30
 <211> 11204
 <212> DNA
 <213> Escherichia coli

<400> 30
 ttgaggcggc gctggctcat tcccctcagc ccgaagggtg gaggggaatg cccgaattaa 60
 ccgtcagttc gggattatct acgcgttagt ttttcagat cggttcaat ctgcctgatc 120
 ttattggtta cgactgactc caggtgacgt aagtcgtcga ggatcttacg ctttaagatca 180
 acttcgctgc ggtcacgctg gcagatttga tcaagctcat caatgatata ccgtagattc 240
 gggctgattt cctggacttc tttataacct tgaccacac catcagcgc gaccgtttta 300
 cgctgacgcg gatattttaa ctttaacgctc ttggcgaaaa actcgccctt gtctttcttg 360
 aaatagattt tcagaatata gttgttggct tcctgccgga ggctgtaacg atcaatttct 420
 tcaggattgg taatacccag acttttcaga ttgtcgtaca tagcgttacc tcaaaatgag 480
 tcagtaaatt gtgcttattt tagcatttgg cctggccgcg ccggctcgcac tgtgatcgca 540
 gaaagccttt ttcttcaga cttctatcat ggcgtaaaaa tcaaaaaatt acctgcttta 600
 ttctggatgat aaaattcacg atctacacca gttccagcat cgttacttcc gggcggcagt 660
 tcagacgcaa accatacaaa ctgcccacgc cacgggttgt gtagatgtgt ctttcgcaa 720
 aggcatttaa tccggcgacg taacgtttat cttcgacagg ggcaaaagggt tcaccaacca 780
 acggtacgcg cagttgcccg ccgtgagtat ggccgcacag catcagatcc cagggttcat 840
 cacgcatgac ttctttgctg tcgggattat gcgccagcac cagtcttggc agattggctt 900
 cgctggcggg aggcgggttg cattgtccgg ccataagtc accagtgcct accagttcga 960
 attgcctgtt cggcgtggcg atcaccgtag cctgggttaa caacaccgtg atgcccgctg 1020
 atttcaacgt ctgcctaatt aagtgatttt tttctgtacc aacagggcga tcgtgggttg 1080
 cgaagcaggc aaacgtcggc gcacattcgg caagggggga gactacgtca ctaaagccg 1140
 aaaaattcag cgacatatca aataatacgt aatcgccgcc cagtaatatc aaatcgggct 1200

- 125 -

tttgttctat gccaaagagca atcgcgtcag aaatcaggct taaaggaaca aaacgagagt 1260
aatggagatc ggccagaaaa agaattttga atgggtgctgc gttgtcttta aaaaaggcga 1320
ggcggtgacg gattaattca aaccagcctg gctcacagta atgcatataa ccgaaacctg 1380
agctcgtggc tategtcgcg gcagtagcct gcaaaaatcg gcggcgtgaa atcatcgctc 1440
atccctgcaa aaaaatcggg cagcgtcgtg ctgcccgtgt gcatactttt agtcgatggg 1500
acgcagcagt tcgttaatgc cgactttgcc gcgagttttc gcgtcaactt tcttaacgat 1560
aaccgcacag tagaggctgt atttgccatc ttttgacggc agattacctg aaacaaccac 1620
agaccccgcc ggaacgcgac cgtagtggat ttgcgcggtt tcacggtcgt aaatacgggt 1680
gctctgacca atgtatacgc ccatggaaat gacggaacct tcttcgacaa tcaccccttc 1740
aaccacttca gagcgcgcgc cgatgaagca attatcttca atgatgggtg ggtagcctg 1800
cagcggttcc agcacgccgc cgatgccac gccaccggaa aggtggacgt ttttaccaat 1860
ctgcgcacaa gaaccgacgg tcgcccaggt atcaaccatg gtgccttcat caacatatgc 1920
gccgatgttg acgtaagacg gcatcagcac ggtgttacgg gcaataaacg caccctgacg 1980
taccgccgct ggtggcaciaa cgcggaagcc ttctttctgg aaacgtgctt cgtcgtagtc 2040
ggcgaatttc atcggcactt tgtogaagta gcggctttct gcccttcga tcacctgatt 2100
atcattaata cggaaagaga gcagcaccgc ttttttcaac cactgatgcg tcacccactg 2160
accgtcaatt ttttcogcta cacgcagtgc gccggaatcc agcagggcga tcacctgatt 2220
taccgcttcg cgggtaacgg tgtctgcatt ggctggcgtg atctcggcac ggcgttcaaa 2280
agcggtttca ataatgttct gtaactgctg cattgttaaa ctcttttcat atcagtaaac 2340
acatcacctt ttatcgtttg gattgagggc ctctgtcaac cgctgatgca cttcctgctg 2400
caactcgtta ttaagcgcac gccggtcagc ggtggcaatt atgaataaat cttctactcg 2460
ctcgccaatg gttgtaattc gggcaccatg aagcgaaatt ccagatcgg caaaaatttt 2520
cccgactcgc gccagcagtc caggttggtc gagggcgatc agttcgagga acgatttgcg 2580
gtcggtatgg gtcggcaaaa acgttacttc ggtttcaaca gtaaaatggc gtaatttggc 2640
gggttggcga cggggctgcg gtggctgcca gctactttgc gtcagtactt gctccagacc 2700
aaaccgaata acctcatgac gatctgcgga cagcgggttg ccatcgggtt ccagcacgat 2760
aaaggatatcc atcgccatac cgtcgcgagt ggtgaaaatt tgtcgcgtcg gaacacttaa 2820
attgcggcgg tctaattcgg cacagacggc ggcaaacaga taagggcggg ccgggctcca 2880
gataaaaatc tcggtgcctc cagcgtagc ctgcgggcta agcaatacca gcggtttgct 2940

- 126 -

taaatcatgc tgtaataaat ggcgggcatg ccaggccagt tgatttgggc tatggcggac 3000
aaaatagtta gcacgacagc gtgaccaaatt ttggtgcagc gcctcttcgt cgatgttatc 3060
catgcgcagt agtgccagt cctggagttg gtgatggcga acccgttcgc gcataatccgg 3120
cgtgttttgc attccgcgtc gtagctgctt ttcggtggca aagtagagct cagcaaacag 3180
actttgcttc cagctattcc acagcgtttc gttggtggcg caaatgtcag ccacagtcag 3240
gcataccaga tagcgcagac gattttccgt ttgcacttct tcggcaaaact gcttgatgac 3300
ttccgggtcc tgaatatcgc ggcgltgggc ggtcacgcag atcaacaggt gctggcgaac 3360
cagccaggcg accagctgtg tttcgcgtga gttcagcccg tggagttcgg caaaatgcac 3420
tacatcctga gcaccgagaa tggagtggtc gccgccgcgt cctttggcga tatcgtgaaa 3480
cagcgcggcg atgaaaatca gctcagttga cggcaggcgc ggccagacgt ccacacacaa 3540
cggatggcgc tggcgcgttt cttcactggc aaaactctcc agtttcagca tcacgcggat 3600
agtatgttca tccaccgtgt aggcgtggaa cagatcaaac tgcattctgc cgacgatatg 3660
cgaccattgc ggcatatacg cgccgagcac gctatggcga tgcattggca atagcccgcg 3720
ccgcaccgct ccgggggtgac gcagaatgct caaaaacagt tttcgtgctt ccggaatatt 3780
acacagcggg tgttgcagat ggcgacgggc atggcgtaac tggcgcagcg tgggtggagta 3840
aatgccggtg atcgcactgt tgtgcaccat ggtgtagaac atacgcaaga tggcttccgg 3900
ctggcgcata aatagtgttt catcacgcag gtcgattagc gtaccgcgta gctgaaactc 3960
atcgtcgatt ggacgtggtt tttcgtcggc gggaagggcg aggatggctt catcgaacag 4020
ttgcagcagc atctggttga gttcactgac gcggcggtga acgcggaagt aatccttcat 4080
catccgctcg accggttcgt taccttcacc actgtaattc agacgctggg cgacgctaag 4140
ctggcgatcg aataacaggc gattatcgta acggctgacg accagatgca gggcaaagcg 4200
aatacgccac aatatatgca gacattcgtt taattccgcc cgctccgctg aggttaagaa 4260
gccaaaccgg accatttcat ccagcgatgt tgcgccaaaa tgacggcggg ccaccattg 4320
cagagtgtgg atatcgcgca agccgccagg gctgcttttg atgtctgggt caaggttgta 4380
gctggtgcca tggtaacgct gatggcgctg gttctgttct tcaactttcg ccgcgtagaa 4440
cttgtcggaa ggccagaatc cttcgctgaa aatatgtttt tgcagttcga ggaacagcgc 4500
aacatcgcca attaataagc gggattcgat taaattgggt gcgacgggta aatccgataa 4560
cccttcagc atgcactctt caagcgtgcg cacgctatga ccgacttcca gctttacatc 4620
ccagagcagc gttaacagct cgcccacttt ttgcgcctga tcgtccggga gctttttacg 4680
gcttaaaatc agtaaatega cgtctgaaag tggatgcagc tcgccacgac cgtagccacc 4740

- 127 -

gacggcgacc aatgccaggt cggcaatctg gctgaatccc gcttcaatcc ataatcgttg 4800
caggagctgg tcgataaact cgggtgcgcgc ctcaatcaac tgttctgcag agatcccatt 4860
gtcaaaggca tcacccagcc aacgctggaa agtatcgata tgggctttta tcccaccgac 4920
ggttaattca tcacggggcc agacgcatgg attttgcggt tgaccgggca ggggtggggag 4980
agcgggtgttt gcgtactgtt ctggaagggt attcattgtg cgccacccat aagattaaat 5040
tatcgcattha aaaaagccgg cattcgccgg cttcatctta ttcgtcgtgc gagattatcg 5100
ccgggatggg gtcattccttg cgtagcgtca gaatttcgca gccgttatca gtcaccacaa 5160
tagtatgctc atattgtgca gacaagctgc gatctttggg ttttaccgtc cagccatctt 5220
tcatgggtgcg gatctctttt ttacccgcgt tgaccattgg ctcgatgggtg aacgtcatcc 5280
caggtttcag tacgacgttg gtttcacggg agtcatagtg cagcacctgc ggttcttcat 5340
ggaagccgcg accaataccg tgtccgcaat attcacgaac gacggagaag ccttctgctt 5400
cgacaaatth ctgaatcgcc gcaccgattt cgcgcagatt aatgcctggg tttaccatgc 5460
gtagcgccag gtacaggctt tcttgcggtg tgcggcacag acgttcgccc atgatgggtc 5520
gcttaccgac gataaacatt ttcgaggtat cgccgtggaa accatctttg attacggtga 5580
catcaatgtt aacgatatcg ccatctttca gcagcttagc atcgctccggg ataccgtggc 5640
acaccacttc attaatagag atgcaaacgg atttcggata gccgtgatag ccgaggcagg 5700
cagaaaccgc gtgttggttca ttaacaatgt aatcattaca gatgcgatcc agctcgccgg 5760
tgctgacgcc cggtttaaca tacggttcga tcatctccag cacttcggca gccagtcggc 5820
cagcgacgcg catthttttcg atatcttctg gggctcttgat tgagatagcc attaatctg 5880
tccatcagcg tcggtgatac cgacaatata tatgtaagtg ccgtcaatgg tatcacaccc 5940
gggcaaattg agaatcattc tgaatttcgc caaacgtgcc actgaagggtt ttctataata 6000
gaaaattcga cgtctgatgc tgtacacagc gccacaatt attgggtgtcc acgacgtatt 6060
tgtgggtataa agcgcgcggg acttccgata catttcgtat acacagactg gacggaagcg 6120
acaatctcac tttgtgtaac aacacacagc tatcggcaca tattccgggg tgccctttgg 6180
ggtcggtaat atgggatacg tggaggcata accccaactt ttatatagag gttttaatca 6240
tggcaactgt ttccatgcgc gacatgctca aggctgggtg tcaacttcggg caccagaccc 6300
gttactggaa cccgaaaatg aagccgttca tcttcgggtg gcgtaacaaa gttcacatca 6360
tcaaccttga gaaaactgta ccgatgttca acgaagctct ggctgaactg aacaagattg 6420
cttctcgcaa aggtaaaatc cttttcggtg gtactaaacg cgctgcaagc gaagcgggtga 6480

- 128 -

aagacgctgc tctgagctgc gaccagttct tcgtgaacca tcgctggctg ggcggtatgc 6540
tgactaactg gaaaaccgtt cgtcagtcca tcaaacgtct gaaagacctg gaaactcagt 6600
ctcaggacgg tactttcgac aagctgacca agaaagaagc gctgatgcgc actcgtgagc 6660
tggagaaact ggaaaacagc ctgggcggtg tcaaagacat gggcggtctg ccggacgctc 6720
tgtttgtaat cgatgctgac cacgaacaca ttgctatcaa agaagcaaac aacctgggta 6780
ttccggtatt tgctatcgtt gataccaact ctgatccgga cgggtgttgac ttcgttatcc 6840
cgggtaacga cgacgcaatc cgtgctgtga ccctgtacct gggcgctgtt gctgcaaccg 6900
tacgtgaagg ccgttctcag gatctggctt cccaggcgga agaaagcttc gtagaagctg 6960
agtaataagg cttgataact ccccaaaaat agttcgagtt gcagaaaggc ggcaagctcg 7020
agaattcccg ggagcttaca tcagtaagtg accgggatga gcgagcgaag ataacgcac 7080
tgcggcgcga aatatgaagg gggagagccc ttatagacca ggtagtacac gtttggttag 7140
ggggcctgca tatggcccc tttttcactt ttatatctgt gcggtttaat gccgggcaga 7200
tcacatctcc gaggatttta gaatggctga aattaccgca tccctggtaa aagagctgcg 7260
tgagcgtact ggcgcaggca tgatggattg caaaaaagca ctgactgaag ctaacggcga 7320
catcgagctg gcaatcgaaa acatgcgtaa gtccggtgct attaaagcag cgaaaaaagc 7380
aggcaacgtt gctgctgacg gcgtgatcaa aaccaaatac gacggcaact acggcatcat 7440
tctggaagtt aactgccaga ctgacttcgt tgcaaaagac gctggtttcc aggcgttcgc 7500
agacaaagtt ctggacgcag ctgttgctgg caaatcact gacgttgaag ttctgaaagc 7560
acagttcgaa gaagaacgtg ttgcgctggt agcgaaaatt ggtgaaaaca tcaacattcg 7620
ccgcgttgct gcgctggaag gcgacgttct gggttcttat cagcacggtg cgcgtatcgg 7680
cgttctgggt gctgctaaag gcgctgacga agagctgggt aaacacatcg ctatgcacgt 7740
tgctgcaagc aagccagaat tcatcaaacc ggaagacgta tccgctgaag tggtagaaaa 7800
agaataccag gtacagctgg atatcgcgat gcagtctggt aagccgaaag aaatcgcaga 7860
gaaaatgggt gaaggccgca tgaagaaatt caccggcgaa gtttctctga ccggtcagcc 7920
gttcgttatg gaaccaagca aaactggttg tcagctgctg aaagagcata acgctgaagt 7980
gactggcttc atccgcttcg aagtgggtga aggcacgag aaagttgaga ctgactttgc 8040
agcagaagtt gctgcgatgt ccaagcagtc ttaattatca aaaaggagcc gcctgagggc 8100
ggcttctttt tgtgcccatc ttgtaaattc agctaaccct tgtggggctg cgctgaaaag 8160
cgacgtacaa tgtcgctagt attaattcat ttcaatcgtt gacagtctca ggaaagaaac 8220
atggctacca atgcaaaacc cgtctataaa cgcattctgc ttaagttgag tggcgaagct 8280

- 129 -

ctgcagggca ctgaaggctt cggtattgat gcaagcatac tggatcgtat ggctcaggaa 8340
atcaaagaac tggttgaact gggatttcag gttggtgtgg tgattggtgg gggtaacctg 8400
ttccgtggcg ctggtctggc gaaagcgggt atgaaccgcg ttgtgggcga ccacatgggg 8460
atgctggcga ccgtaatgaa cggcctggca atgcgtgatg cactgcaccg cgcctatgtg 8520
aacgctcgtc tgatgtccgc tattccattg aatggcgtgt gcgacagcta cagctgggca 8580
gaagctatca gcctgttgcg caacaaccgt gtggtgatcc tctccgccgg tacaggtaac 8640
ccgttcttta ccaccgactc agcagcttgc ctgcgtggta tcgaaattga agccgatgtg 8700
gtgctgaaag caaccaaagt tgacggcgtg tttaccgctg atccggcgaa agatccaacc 8760
gcaaccatgt acgagcaact gacttacagc gaagtgtggt aaaaagagct gaaagtcattg 8820
gacctggcgg ccttcacgct ggctcgtgac cataaattac cgattcgtgt tttcaatatg 8880
aacaaccgg gtgcgctgcg ccgtgtggta atgggtgaaa aagaaggagc tttaatcacg 8940
gaataattcc cgtgatggat aaataagggt aagattccgc gtaagtatcg cgggggcgta 9000
agtctggtta taaggcgta ttgttgcagg cagtttggtc acggccagcg cgcagcaacc 9060
ggagcgtaca aaagtacgtg aggatggcga gcactgcccg gggccaaaat ggcaaataaa 9120
atagcctaata aatccagacg attaccgta atatgtttaa tcagggctat acttagcaca 9180
cttccactgt gtgtgactgt ctggtctgac tgagacaagt tttcaaggat tcgtaacgtg 9240
attagcgata tcagaaaaga tgctgaagta cgcattggaca aatgcgtaga agcgttcaaa 9300
acccaaatca gcaaaatacg cacgggtcgt gcttctccca gcctgctgga tggcattgtc 9360
gtggaatatt acggcacgcc gacgccgctg cgtcagctgg caagcgtaac ggtagaagat 9420
tcccgtaacac tgaaaatcaa cgtgtttgat cgttcaatgt ctccggccgt tgaaaaagcg 9480
attatggcgt ccgatcttgg cctgaaccgg aactctgcgg gtagcgacat ccgtgttccg 9540
ctgccgccgc tgacggaaga acgtcgtaaa gatctgacca aaatcgttcg tggtagaagca 9600
gaacaagcgc gtgttgcagt acgtaacgtg cgtcgtgacg cgaacgacaa agtgaaagca 9660
ctgttgaaag ataaagagat cagcgaagac gacgatcgcc gttctcagga cgatgtacag 9720
aaactgactg atgctgcaat caagaaaatt gaagcggcgc tggcagacaa agaagcagaa 9780
ctgatgcagt tctgatttct tgaacgacaa aaacgccgct cagtagatcc ttgcggatcg 9840
gctggcggcg ttttgctttt tattctgtct caactctgga tgtttcatga agcaactcac 9900
cattctgggc tcgaccggct cgattggttg cagcacgctg gacgtggtgc gccataatcc 9960
cgaacacttc cgcgtagttg cgctggtggc aggcaaaaat gtcactcgca tggtagaaca 10020

- 130 -

gtgcctggaa ttctctcccc gctatgccgt aatggacgat gaagcgagtg cgaaacttct 10080
 taaaacgatg ctacagcaac agggtagccg caccgaagtc ttaagtgggc aacaagccgc 10140
 ttgcgatatg gcagcgcttg aggatgttga tcaggatgat gcagccattg ttggcgctgc 10200
 tgggctgtta cctacgcttg ctgcgatccg cgcgggtaaa accatthttgc tggccaataa 10260
 agaatcactg gttacctgcg gacgtctgtt tatggacgcc gtaaagcaga gcaaagcgca 10320
 attgttaccg gtcgatagcg aacataacgc catthtttcag agthttaccgc aacctatcca 10380
 gcataatctg ggatacgctg accttgagca aaatggcgtg gtgtccattt tacttaccgg 10440
 gtctgggtggc cthttccgtg agacgccatt gcgcgatttg gcaacaatga cgcgggatca 10500
 agcctgccgt catccgaact ggtcgatggg gcgtaaaatt tctgtcgatt cggctaccat 10560
 gatgaacaaa ggtctggaat acattgaagc gcgttggctg tttaacgcca gcgccagcca 10620
 gatggaagtg ctgattcacc cgcagtcagt gattcactca atgggtgcgt atcaggacgg 10680
 cagtgttctg gcgcagctgg gggaaccgga tatgcgtacg ccaattgccc acaccatggc 10740
 atggccgaat cgcgtgaact ctggcgtgaa gccgctcgat thttgcaaac taagtgcgtt 10800
 gacatthtgc gcaccggatt atgatcgtha tccatgcctg aaactggcga tggaggcggt 10860
 cgaacaaggc caggcagcga cgacagcatt gaatgccgca aacgaaatca ccgttgctgc 10920
 thttcttgcg caacaaatcc gctthtacgga tatcgctgcg ttgaatthtat ccgtactgga 10980
 aaaaatggat atgcgcgaac cacaatgtgt ggacgatgtg ttatctgttg atgcgaacgc 11040
 gcgtgaagtc gccagaaaag aggtgatgcg tctcgcaagc tgaggataat ccggctacag 11100
 agagtcgcgc tathttgttag cgtagggtt cagtgatata gtctgcgcca tctgatcgta 11160
 agtagttggc thtataaggt cagatatgcc gtggthtttac acgg 11204

<210> 31

<211> 264

<212> PRT

<213> Escherichia coli

<220>

<223> complement of position 5078-5872 of seq id 30

<400> 31

Met	Ala	Ile	Ser	Ile	Lys	Thr	Pro	Glu	Asp	Ile	Glu	Lys	Met	Arg	Val
1				5				10					15		

Ala	Gly	Arg	Leu	Ala	Ala	Glu	Val	Leu	Glu	Met	Ile	Glu	Pro	Tyr	Val
			20					25					30		

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

- 131 -

Ile Val Asn Glu Gln His Ala Val Ser Ala Cys Leu Gly Tyr His Gly
 50 55 60
 Tyr Pro Lys Ser Val Cys Ile Ser Ile Asn Glu Val Val Cys His Gly
 65 70 75 80
 Ile Pro Asp Asp Ala Lys Leu Leu Lys Asp Gly Asp Ile Val Asn Ile
 85 90 95
 Asp Val Thr Val Ile Lys Asp Gly Phe His Gly Asp Thr Ser Lys Met
 100 105 110
 Phe Ile Val Gly Lys Pro Thr Ile Met Gly Glu Arg Leu Cys Arg Ile
 115 120 125
 Thr Gln Glu Ser Leu Tyr Leu Ala Leu Arg Met Val Lys Pro Gly Ile
 130 135 140
 Asn Leu Arg Glu Ile Gly Ala Ala Ile Gln Lys Phe Val Glu Ala Glu
 145 150 155 160
 Gly Phe Ser Val Val Arg Glu Tyr Cys Gly His Gly Ile Gly Arg Gly
 165 170 175
 Phe His Glu Glu Pro Gln Val Leu His Tyr Asp Ser Arg Glu Thr Asn
 180 185 190
 Val Val Leu Lys Pro Gly Met Thr Phe Thr Ile Glu Pro Met Val Asn
 195 200 205
 Ala Gly Lys Lys Glu Ile Arg Thr Met Lys Asp Gly Trp Thr Val Lys
 210 215 220
 Thr Lys Asp Arg Ser Leu Ser Ala Gln Tyr Glu His Thr Ile Val Val
 225 230 235 240
 Thr Asp Asn Gly Cys Glu Ile Leu Thr Leu Arg Lys Asp Asp Thr Ile
 245 250 255
 Pro Ala Ile Ile Ser His Asp Glu
 260

<210> 32

<211> 14820

<212> DNA

<213> Escherichia coli

<400> 32

cactttgtta acaacttta ctactcttta atgcagtatt aaagattaat cggtacaaca 60
 gtgagctgtt atgactgata aaaccattgc gttttcgcta ctcatctgg ccccatctcc 120
 cgaagggttct tcagcgcgag aagcattctc ccactctctc gatctcgccc gtctggctga 180
 aaagcgcggc tatcatcgct actggctggc agaacaccac aatatgactg gcattgccag 240
 tgctgccacg tcggtattga tcggctatct ggcggcgaat accaccacgc tgcattctggg 300

- 132 -

gtctggcggc gtgatgttgc ctaaccactc accgttggtc attgcagaac agttcggcac 360
gcttaataca ctctatccgg ggcgaatcga tttggggctg ggtcgtgctc cgggtagtga 420
ccaacggaca atgatggcgc tacgtcgtca tatgagcggc gatattgata atttcccccg 480
cgatgtggcg gagctgggtg actggtttga cgcgcgcgat cccaatccgc atgtgcgccc 540
ggtaccaggc tatggcgaga aaatccccgt gtggttggtta ggctccagcc tttacagcgc 600
gcaactggcg gcgcagcttg gtctgccgtt tgcgtttgcc tcacacttcg cgccggatat 660
gctgttccag gcgctgcac tttatcgcag caacttcaaa ccgtcagcac ggctggaaaa 720
accatacgcg atgggtgtgca tcaatattat cgccgccgac agcaaccgcg acgctgaatt 780
tctgtttacc tcaatgcagc aagcctttgt gaagctgcgc cgtggcgaaa ccgggcaact 840
gccgccgccg attcaaaata tggatcagtt ctggtcaccg tctgagcagt atggcgtgca 900
gcaggcgctg agtatgtcgt tggtaggtga taaagcgaaa gtgcgtcatg gcttgcagtc 960
gatcctgcgc gaaaccgacg ccgatgagat tatgggtcaac gggcagattt tcgaccacca 1020
ggcgcggctg cattcgtttg agctggcgat ggatgttaag gaagagttgt tgggatagtg 1080
tgtcttaacg cgggaagcct tatccgagct ggcaacgctg tcctacatag acctgataag 1140
cgaagcgcat caggcattgt gtaggcagca gaaatgtcgg ataaggcacc gctgattact 1200
gatacaccgg cagtaaatta aagctcgata aaatatgcac cagtgcgttg ccgacgcaa 1260
acaccagaat cagcgcaatc atcggcttgc caccocagac gcggaatttc gggctgcaa 1320
agcgtttacg cgatgcacgg gctaacagcg ccggaacaat tgccgccag atggtagccg 1380
ctaaaccagc ataaccaatg gcgtacagga atccgttcgg gaacaacagc cccccacaa 1440
ctggcggggc aaaggtcagc aatgccgttt tcaagcggcc cacagccgag tcgtcgaaac 1500
caaacagatc tgccagatag tcaaacaac ccagcgttac gccgaggaac gaactcgcta 1560
ccgcaaagtt tgagaacacg accagcagca gatccagact acggctgttc agtacgccgc 1620
ttaacgcctg taccagcaca tcaatattac cgcccttctc tgcaataccg ataaactccg 1680
gacgcgggat gttaccatc gtgcgcagca accagatggg atacagcgcc agcgccatca 1740
gcgtaccgta caccagacat ttcacgatgg ttttcggatc tttgccgtaa tacttcatca 1800
ggcttggcac gttaccgtga taaccaaacg atgccagaca gaacggcagg gtcataca 1860
gatacgggtg ataagacgca ttgctttcgg cgacgttgaa caatgtcgca ggctgcacat 1920
gccccagcag gctaccaaag gtgaggaaga aggtaatgac tttcgcccc agcacaatcg 1980
ctgtcatgcg actgacggct ttagtgctca accacaccac aaacgctacc agcaatgcaa 2040
aaccaaaacc cgccgccgt gccgggacgt ttagtgacat ctctgcgaag gtgtgatgca 2100

- 133 -

gaatcgaacc actggcagaa atataggcat aggtcaggat atagagcaca aaggcaatgg 2160
aaatgccgtt gaccacgttc cagcctttgc ccagcaaatac tttgggtgatg gtgtcaaaac 2220
tcgaaccgat tctgtaattc aggttagctt ccagaatcat caagccggaa tgcagcatac 2280
agaaccaggt aaagatcagc gccgccattg accagaaaaa ccacgccccg gacatgacca 2340
ctggcagaga aaacatccct gcgccaataa tgggtgccgc gataatcacc acgccgcca 2400
gcagcgacgg tgacgtttgg gtgggtggta gtgttgccat gagggcttct ctccagtga 2460
aaatagtgcg actgcgttgt tatgcattgc actgtaccag tacacgagta caaaagacag 2520
aaaaaaagcc ccgatggtaa aaatcggggc tgtatatatt attttacaga ttgtgttcgc 2580
tgttcagcga tgattacgca tcaccaccga aacgacgacg accggtagaa tcatcacgac 2640
gcggagcgcg gccttcacga cgttcgcccgc taaaacgacg accatcacca cggccacctt 2700
cacggcggttc accgctgaag ttacgaccgc cttcacgacg ttccgccaccg aaaccacgac 2760
caccgccacg acgctcaccg ccagtatgcg gctgtgcac gccagtaac tgcattgttca 2820
tcggcttggt gagaatgcga gtgcgcgtaa agtggttcag cacttcaccc ggcatacctt 2880
tcggcagttc gatggtggag tgagaagcaa acagcttgat gttaccaatg taacggctgc 2940
tgatgtcgcc ttcgtttagca atcgcaccaa cgatatgacg aacttcaaca ccatcatcgc 3000
ggcccacttc aatgcggtac agctgcatat cgccaacatc acgacgttca cgacgcggac 3060
gatcttcacg gtcaccacgc gggccacggc cgttacgac gccgcggacca cggtcatcac 3120
ggtcacggaa ttcacgtttc ggacgcatcg gcgcatctgg cggtagacg agagtacgtt 3180
caccctgtgc cattttcagc agtgccgcag ccagagtttc gagatccagc tcttcacctt 3240
cagcagtcgg ctgaattttg ctacgagtg cgcggtattg atccagatcg ctgctttcca 3300
gctgctgctg tacttttagcg gcgaattttt ccagacggcg tttgcctagc agttctgcgt 3360
tcggcagttc tacttccgga atagtcagct tcatagtacg ttcaatgttg cgcagcagac 3420
gacgctcgcg gttctcaacg aacagcagcg cgcggccagc acgacccgca cgaccggtac 3480
gaccgatacg gtgaacgtaa gactcagaat ccatcgggat atcgtagtta actaccaggc 3540
tgatacgctc aacgtccagg ccacgggctg caacgtcggc cgcaatcagg atgtccagac 3600
gaccatcttt caggcgttcc agtgtctgtt cacgcagcgc ctggttcatg tcaccgttca 3660
gcgcggcgct gttgtagccg ttacgctcaa gagcttcagc cacttccaga gtcgcgtttt 3720
tggtacgaac gaagataatc gccgcatcaa aatcttccgc ttccaggaaa cgtaccagtg 3780
cttcgttttt gcgcataccc cagacagtcc agtagctctg gctgatgtca ggacgggtag 3840

- 134 -

tcacgctgga ctgaatgcg ctttcctgcg gctctttcat aaagcggcgg gtaatgcgac 3900
gaatcgcttc cggcatgggt gcagagaaca gagcgggtctg atgaccttcc gggatctgcg 3960
ccataatggg ttcaacgtct tcgatgaagc ccatgcgag catttcgtca gcttcatcca 4020
gaaccagacc gctcagttta gagaggcca gaggccacg tttcagggtg tccagcagac 4080
ggcccggagt accgacaacg atctgcggcc cctgacgcag ggcgcgtaat tgcacgtcat 4140
aacgctggcc gccgtacaga gcaaccacat ttacgccgcg catgtgttta gagaaatccg 4200
tcattgcttc agcaacctgt accgccagtt cgcgggtcgg tgccagcacc agaattctgtg 4260
gtgctttcag ctgaggatca agattctgca acagaggtaa agagaatgct gcagtttttc 4320
cgctccccgt ctgggccata ccagaaacgt cgcggccatt cagcagatgt ggaatacact 4380
ctgcctgaat tggagatggg ttttcgtaac ccagatcggt aagggttca aggataggag 4440
ccttcaggcc cagatctgca aaagtgggtt cgaattcagc catgtagtac gtgtgcctca 4500
aaattaatgg cggccagtct acataactca tcatgaaatt gatcagcaat tttcattgaa 4560
aagtgtgaac cggctcaaag taggtgtatt aacgaacaac aacgccctca cccgttaagg 4620
tgatggcaat caaaaagat tacgggctga tgtgtacgtc agctattgct ggtccgattc 4680
tgccaggcca tcttggtcct ggcccaggag cgataattcc aacaatgcgt atcgggtgctc 4740
aacaagttta tgaacgttgt tggcaaccgc cagtttgaac agtgccgtgg cgctgtccaa 4800
atccccaga cttaggtagt acttacctaa atagaagttg gtttactga gatgctcagc 4860
gagcgagggtg ttatccgttg cgtccgcctt gagcctttcc attaacgttt gttcgctaata 4920
gttgcccagg tagaactcga caatgttcca tccccactgt tccttatccg atttttcgaa 4980
gtgctgtttc aacacttctt tagcctgctt ctcatcgagc ttctgctcgg cgagataaag 5040
ccacagacta cggaaaggat cattgggatc gtcttgataa aacgccagca gatcatcttg 5100
cgctaacttg tcacgaccgc cgtaataata tgcgatcccg cgattcaagt gcgcgtagtt 5160
gtaagttgga tcaagctcaa gtacagaatc aaacgcttca taggcagcat caaaattgcc 5220
tgctgctgtt aaatatatgc ctaagtaatt gaatacttca ggcatatccg gtcggattgc 5280
cagcgcttgc gaaaaatcgt tacgcgctaa tgccctcaga ccgagactat catacaacac 5340
tccgcgctca tataaaagct gtgcgcgttc gtcacgggtt aaagcccagc tggcaaggat 5400
ttgttccata cgtgccagaa tcacttcctg ctgtaaagtc ggttgcaatg gtaccgcgag 5460
gacttcactt ttacgccagg aagtattact gcacctgca agcgtaagtg ctgtcgcaac 5520
gaaacaccag cgcaaaaaag gcttcatttc ccactcccga agaccacggg tgaatgaacg 5580
tcctgttccc ggttgctaac aaggcgtcct gcccggttaa aagcccccg ccgcagcgga 5640

- 135 -

gggcaaattgg caaccttact cgccctgttc agcagccgga gcttccggtg ctgcagcagg 5700
ttgagactgc tcagtcgctt ctttaattgct cagacggata cggccctggc gatcaacttc 5760
cagaactttc accggtactt cctgacccat ctgcaggtaa tcggtcactt tctcaacgcg 5820
tttgtcagcg atttgagaga tgtggaccag accttcttta ccgccgccga tggcaacaaa 5880
tgcgccaaag tcaacgatac gggtcacttt accagtgtag acgcggccca cttcgatttc 5940
tgcaagtgatc tcttcgatac gacgaatagc atgtttcgct ttctcgccgt cggtcgctgc 6000
aatcttcaca gtaccgtcat cttcgatttc gatggtggtg ccagtttctt cggtcagagc 6060
acggattaca gaaccgcctt taccgataac atctttgatc ttgtccgggt tgatcttgat 6120
ggtatggata cgcggtgcga actcagagat atcgccacgc ggcgcgttga tcgcctgttc 6180
cattaogccc aggatatgca gacgcgcacc tttagcctgg ttcagcgcaa cctgcatgat 6240
ctctttggtg ataccttcaa ttttgatata catctgcagt gcagagatac cgtcgcggga 6300
acctgcaact ttgaagtcca tatcgcccag gtgatcttcg tcgccccaaa tgtcagacag 6360
tacaacgtag ttgtcgctt ctttcaccag acccattgcg ataccgcaa cggcagcttt 6420
gatcggcaca cctgcgtcca tcagcgccag agacgcgccg cacacggaag ccatagaaga 6480
ggaaccgttg gattcagtga tttcagacac aacacgtacg gtgtacggga atttgtccat 6540
atccggcatg actgccagca cgccgcgctt cgccagacga ccgtgaccaa tttcacgacg 6600
cttcggagaa ccgaccatgc cggtttcgcc tacggagtag ggagggaagt tgtagtggaa 6660
caggaaggta tcggtacgtt cgcccatcag ttcatacaaga acctgcgcgt cacgagcagt 6720
accagcgtt gcggtaacca gcgcctgcgt ttcaccacgg gtgaacagcg cagaaccgtg 6780
agtacgcggc agcacgccag tacgcacatc cagaccacgg atcatatctt tttcacgacc 6840
gtcgatacgc ggttcgcctg ccagtacgcg gctacgaaca acgtttttct cgatcgcgtg 6900
cagaatttca ccagttcgt tttcgtccag ggtttcgtct tcagcaagca gcgtcgcgat 6960
ggtttcagat ttgatgacat caacctgcgc ataacgctct tgtttgtcgg tgatgcggta 7020
agcatcgctc aggcgagctt cagccagtgc agcaacgcgc gcgttttagcg cttcgtttac 7080
cggctccggc tgccagtccc aacgcgggtt accggcttct ttcaccagtt cattgatgtt 7140
ctgaataaca acctgctggt gtcatgacc gaacactact gcgcccagca tctggtcttc 7200
gctcagcagt tgagcttcag attcaaccat cagtacagcg gcttcagtag cggcaacaac 7260
cagatccagt ttgctctctt tcagctcgtc ctgagtcggg ttcagtagct actggtcatt 7320
gatgtaacct acgcgggcag caccaatcgg gccattgaac ggaataccag acagagacag 7380

- 136 -

cgctgcggaa gcaccaatca tcgcgacgat atccgggtta acttgcgggt taacagaaac 7440
cacggtggcg ataacctgaa cttcgttgac gaagccttcc gggaacagcg ggcgaatcgg 7500
gcggtcaatc agacgcgcga tcagggtttc gccttcgctt gggcggcctt cacgacggaa 7560
gaagctaccc gggatacgac cagcagcgta ggtacgctcc tgatagttaa cggtcagtgg 7620
gaagaagtcc tgacctgggt tggctttttt ctggccaaca acggtaacga ataccgcggt 7680
gtcatccatg ctaaccataa cagcggcagt agcctgacga gccatcatgc cggtttccag 7740
agtcacggtg tgttggccgt actggaattt acgaacgatc ggattaagca atgtaatatc 7800
ctttctcttt cttagacagt accttacggc actgggtgtta ataccgatc ttctgcgcat 7860
cctcgcgact aatgacaacc ctaaccacgc tctatgtggg taaagcctct cattagccgc 7920
gcgaacctct gcaacggaag atcattcata gcaacaatac attagtttcc agtgaattgc 7980
tgccgtcagc ttgaaaaaag gggccactca ggcccccttt tctgaaactc gcaagaatta 8040
gcgacgcaga ccagggcgct cgatgagctg ggtgtaacgt gctacgtctt tacgtttcag 8100
gtagtcgagc agtttacgac gctgagaaac catgcgcagc agaccacgac ggctgtggtg 8160
atcttttttg tgctctgcaa agtggccctg cagggtggtt atctgtgcag tcagcagtgc 8220
tacctgaact tcggtagaac cgggtgcgtt tgcgtcacga ccaaactcag aaacgatattt 8280
agctgttgct tcagtactta gagacatttt aaaactccaa agtatataga atgaaaggac 8340
gccgatctct aattcagcga tcccagtgtg cgttacgcaa agtgttaaac aatttacgcg 8400
acgttaagcg gcagtattct actcgtagcg acctgttata gcaagacggg taacattacg 8460
ccgggtattc aaccaccagg cgacgaggcg caacgcggcc ttcatcgtca atttcgcca 8520
taccgataaa ttgccgttc tcaccttccg tgacgcgaac cagtccttcc agtggcgcac 8580
cagatgtacg aaccgggtta ccatttttga agtaaacaga agacgttaac ggaagattca 8640
ccaccggata gtccgaagct ggactgtcca ttggcatcag taatggatca agtaactccg 8700
cggctggaat atcctgctgt tcagcttgct caacaagttc acgcaggtgc tccaggggtca 8760
ccatccgttc aaccggatat ttacttacgg ccagacggcg caggtaaata acatgcgcgc 8820
cacagccgag tttttcaccc aggtcatcaa tgatggtgcg gatataagt ctttttgagc 8880
agtgaatttc cagctccagc tcattgcctt catggcgaat aaacagcaat tcataaacgg 8940
taatcggacg cgcttcacgc ggaacttcaa tgccctgacg cgcatattcg tacagttttt 9000
tgccctgata tttcagtgt gaatacatcg aagggatctg ttcgatatcg ccacggaaag 9060
tatccagtgc cgctgccagc tgctctgcgc taaaggttac cggacgttct tcaacgatct 9120
gtccgtcggc atcagaagta tcggtacgct gtccaagacg cgcaatgacc cgatagcggt 9180

- 137 -

tgtcggagtc cagcagatac tgggaaaact tcgtcgcttc cccgaggcaa atcggcaaca 9240
tgccggtcgc cagcgggtcc agcgcaccgg tatgcccggc acggttggcg ttatatatac 9300
gtttcacttt ttgcagcgca tcgttgctgg acataccctg aggtttatcc agcaacaaaa 9360
cgccgttaat gtgcgcaccg cgacgacgag gacgactcat tagtcctcct tgctgtcgtc 9420
cgggttaaca cgacgttctt cgtcatgttt gaccacgctg gtcaccaggt ttgacatgcg 9480
catcccttca accagagagt tgtcgtagaa gaaggtcagt tccggcacga tacgcaggcg 9540
catcgctttc cccagcaggc tgcggatgaa accagaagct tcttgcaacg ctttgatgcc 9600
cgctttaacc gcgtcttcat ctttgtcgtt gaggaacgtc acatatactt tggcatacgc 9660
caggtcgcga gacatttcga caccggaaac ggtggtcata atgcccaggc gaggatcttt 9720
aatttcacgc tgcaggatga gagcgatctc tttttgcatt tcctgcgcta cgcgctgcgg 9780
gcgaccaa at tctttcgcca taataaatc tcctgacaaa aaaggggctg ttagcccctt 9840
tttaaaatta atttcagggtg gaagggctgt tcacgttgac ctgataagac gcgccagcgt 9900
cacatcaggc aatccatgcc ggatgcagcg taaacgcctt atcccgcata gaaccctaaa 9960
aaccttaagc aatggtacgt tggatctcga tgatttcgaa tacttcgata acatcgccag 10020
tgccggacgtc gttgtagttc ttaacgccga taccacattc cataccgtta cggacttcgt 10080
taacgtcata tttgaagcgg cgcagggact ccagctcgcc ttcgtagata accacgttgt 10140
cacgcagaac gcggatcggg ttgtgacgtt taaccacacc ttccgtaacc atacagcctg 10200
cgatggcacc aaatttcggc gatttgaaca cgtcacgaac ttccgccaga ccgataatct 10260
gctgtttcag ttccggagac agcataccgc tcatacgccg tttcacttcg tcaatcagg 10320
tatagatgac ggagtagtaa cgcagatcca ggctttccgc ttcaatcact ttacgtgcag 10380
aggcatcagc acgtacgtta aagccaacca ggatggcgtt ggacgccgca gccaggggtg 10440
cgtcggtttc ggtgatacca cctacgccag aaccgatgat cttcacttta acttcgtcag 10500
tagacagttt cagcaaggag tcggagatcg cttcgacaga accctgtacg tctgccttca 10560
ggacgatatt cacttcgtga acttcgcctt cggtcatgtt ggogaacatg ttctcgagtt 10620
tagatttctg ctgacgcgcc agtttaactt cgcggaattt accctgacga tagagtgcaa 10680
cttcacgcgc tttcttctcg tcacgtacaa cggtaacttc atcaccgcga gccggtagc 10740
cggacaggcc gaggatttcc accggaatgg acggaccgc ttccagcact tcctgacca 10800
gttcgttacg catcgcacga acacgaccgt attcgaagcc acacagaacg atatcgccct 10860
tgtgcagagt accttcacgt accagaacgg tagcaaccgg accacgacct ttatcgagga 10920

- 138 -

aggattcgat aaccgcaccg ctgccatac ctttacgtac cgctttcagc tccagaactt 10980
ccgcctgcag caggatagcg tccagcagtt catcgatacc ggtacccgct ttcgcagata 11040
cgtgtacgaa ctggctttca ccgccccact cttccggcag gatgccgtac tgggagagtt 11100
cgttcttaac gcgatccgga tcagcttctg gtttatcgat cttgttctact gcaaccacca 11160
ccggtacctg cgccgctttc gcgtgctgga ttgcttcgat ggtctgcggc atcacaccgt 11220
cgtcggcagc aacaaccagg actacgatgt ccgttgctg cgccaccaga gcacgcattg 11280
aagtaaacgc ggcgtgcccc ggggtgtcca ggaaggatgat catgccgttt tcagtttcaa 11340
cgtggatatgc accaatgtgc tgggtaatgc cgcccgcttc gccagaggcc actttcgttg 11400
aacgaatgta gtccagcaga gaggttttac cgtggccaac gtgacccatg atggtcacaa 11460
ccggcgcgcg cggttcagcc gcagcaccg tgtcacggtc gctcattacc gcctcttcca 11520
gctcgttttc acgacgcagg ataactttat ggcccatctc ttcagcaacc agctgtgcgg 11580
tttctgatac gataacctgg ttgatggttg ccattgcgcc cagtttctac atcgctttga 11640
tgacctgaga gcctttaacc gccatcttgt tcgccagttc gccaacgggtg atagtctcgc 11700
cgatcacaac gtcacggtta acggcctgag caggcttctg gaagccttgc tgcagcgaag 11760
aacctttacg ttttccgcct ttaccgccac gtactgctgc gcgtgcttct tcacgatcag 11820
cttttgattc agcgtgtttg ttgcctttct tcggacgcgc tgctttcgcg ttacgaccac 11880
ggccacggcc gccttcgact tcacgatcgc tttcgtcttc tgccctggcga gcatgttgag 11940
aagtagtgac gtgataatcg ctggaatctt cagtcggttc cgcgttatca gtccatttgt 12000
tttcttccgc catacgacgt gcttcttcag caacgcgacg tgcttcttct tcgagtttac 12060
gacgcgcttc ttcttcagct ttacgcttga gctctgcagc ttctgctca cggcgggctt 12120
tttcagcctg ggcgttttta gtcataatcg cttgttgatt gctcactttg tctttttccg 12180
cagcttcacg tttcgcttgt tcagcagctt cacgcttagc ttgttctgcg gcctcacgtt 12240
cagctttttg ttgcgcctcg cgttttagccg attcttctgc ctcacgacgg gcttgctctt 12300
ccgcttcacg ctgcgcttgc tcttccgctg caaggcgttc agcctcttgc ggatcgcggt 12360
tcacaaaggt gcgtttcttg cggacttcga tttgtaccga tttgcttttt ccaccggtac 12420
caggaatggt aagggtgctg cgtgttttac gttgcagcgt caatttgtcc gggcctgaat 12480
ttttctgatt cagggtggtca atcaaagtct gtttctcttg tgcagacaca gagtcgtcag 12540
cagacttccg gataacctga tcagcaaatt gctgtaccag gcgttccacg gaggtctgtc 12600
gctctgcggc cagcgtttta atcgttacat ctgtcatgct gttccttctt gctacagttt 12660
attacgcttc gtcaccgaac cagcaaatat tacgggcagc cataatcagt gctccggctt 12720

- 139 -

tttcgtcggg caacccttcg atatcagcca gatcatcaat gccctgttcg gcgagatctt 12780
ccagcgtaca aacgccacgg gcggccagtt tgaatgccaa atcacgatct accccttcaa 12840
ggttcagcag atcgtcagcc ggtttggtat caccgaggct ttcttcctgg gcctgtgcaa 12900
tggtggccag tgcattttta gcacgctcgc gcagtgttc aacggtcggc tcatcaaggc 12960
cttcgatttc caacagctct ttcacgcca cataggccaa ttcttcacgc gtcgagaagc 13020
cttcttctac cagaacagtc gcgaagtctt cgtcgtatgc gagatatctg gtgaagggtg 13080
cgatcgctgc gtgcgcttcc gcctgatgct tagcctgcag gtcgtcaacg gtcatcacgt 13140
tgagttccca accgctcagc tgcgaagcca gacgcacgtt ctgaccgtta cggccaatcg 13200
cctgcgccag gttaccggct tcaacggcga tatccatggg gtgtttatct tcatccacca 13260
cgatagaagc aacgtctgcc ggtgccattg cgttaatcac gaactgcgcc gggttatcat 13320
cccacaggac gatatcgata cgctcgccac ccagttcagt agacaccgcc tgaacacgcg 13380
cgccacgcat acctacgcaa gcacctaccg gatcgatacg ttatcgttg gttttcaccg 13440
cgattttcgc acgagaaccg ggatcgcgag ccgctgcttt aatttcaatc acttcttcgc 13500
cgattttctg cacttcaata cggaacagtt cgatcagcat ttccggcttg gaacgagtga 13560
cgaacagttg cgcgccacgc gcttccgggc gaacggaata gagcacgcca cgaacgcggg 13620
cgccagggcg gaagttttca cgcggcagca tatcttcgcg caggatcacg gcttcagcgt 13680
tggtgcccag atccagagag atgttgctgc gggttacttt ttaccacgc ccggtgatga 13740
tttcaccttc gtgttcacgg aactgatcaa ccaccatgc acgttcggct tcacgcactt 13800
tctgcacgat aacctgtttt gccgtctggg tagtgatacg gtcaaaggta acagactcaa 13860
tctgatcttc aacgtaatcg ccaggttca ggctttcatc ttcataacgt gcggcttcaa 13920
gggtgatctt cttgggtcggc tgggtgactt catcaacaac taaccagcga cggaaagtgt 13980
caaatcacc gcttttgcca tcgatctgta cgcgacgctc gatctcttgt tcatattttt 14040
tctttgttgc tgcgccagc gcgctttcca atgcttcgaa aatcttctcg cgaggtagcg 14100
ccttttcatt ggatacggct tcaactacag ccaaaatttc ttgttcatc gcgggctttt 14160
cacctcatcc agactattaa aagtggggaa ccaggttcgc cttctggata ttactcagcg 14220
cgaacacttc atctttacct tcgacggtaa ctgtgatcat ttaccgtct accgctttga 14280
taacgccctg ccatttacga cggttttgta ccgccatacg gagaaccaga gtcacctctt 14340
ctccgacaaa acgggcgtag tggtcagccg tgaacagtgg gcgatcgaga ccggtgagg 14400
agacttcag gttataagca acggtgatgg gatcttcaac atccagcaca gcacttacct 14460

- 140 -

ggtggctcac atcagcaciaa tcatcaacat tgatgccatc ttcactatca atatagatgc 14520
 gcagtgtgga tgtgcgaccg cgaataaatt cgatgccaac cagttcaaaa cccagggcct 14580
 caactggcgc agtaatcatc tctgttaatt tttgctctaa tgtggacaag cccacccccca 14640
 agacataaaa aaagggccta aagcccagtt attctgtagt cagataacaa aaaaccccga 14700
 taaatcgggg ctttatataa ctgaacccta taaccgcaac tgcggtctgg agcactttcc 14760
 agaaggattt tttcaaattc cactacgaag gccgaagtct tcacagtata tttgaaaaag 14820

<210> 33
 <211> 414
 <212> PRT
 <213> Escherichia coli

<220>
 <223> complement of position 1196-2440 of seq id 32

<400> 33
 Met Ala Thr Leu Thr Thr Thr Gln Thr Ser Pro Ser Leu Leu Gly Gly
 1 5 10 15
 Val Val Ile Ile Gly Gly Thr Ile Ile Gly Ala Gly Met Phe Ser Leu
 20 25 30
 Pro Val Val Met Ser Gly Ala Trp Phe Phe Trp Ser Met Ala Ala Leu
 35 40 45
 Ile Phe Thr Trp Phe Cys Met Leu His Ser Gly Leu Met Ile Leu Glu
 50 55 60
 Ala Asn Leu Asn Tyr Arg Ile Gly Ser Ser Phe Asp Thr Ile Thr Lys
 65 70 75 80
 Asp Leu Leu Gly Lys Gly Trp Asn Val Val Asn Gly Ile Ser Ile Ala
 85 90 95
 Phe Val Leu Tyr Ile Leu Thr Tyr Ala Tyr Ile Ser Ala Ser Gly Ser
 100 105 110
 Ile Leu His His Thr Phe Ala Glu Met Ser Leu Asn Val Pro Ala Arg
 115 120 125
 Ala Ala Gly Phe Gly Phe Ala Leu Leu Val Ala Phe Val Val Trp Leu
 130 135 140
 Ser Thr Lys Ala Val Ser Arg Met Thr Ala Ile Val Leu Gly Ala Lys
 145 150 155 160
 Val Ile Thr Phe Phe Leu Thr Phe Gly Ser Leu Leu Gly His Val Gln
 165 170 175
 Pro Ala Thr Leu Phe Asn Val Ala Glu Ser Asn Ala Ser Tyr Ala Pro
 180 185 190
 Tyr Leu Leu Met Thr Leu Pro Phe Cys Leu Ala Ser Phe Gly Tyr His

- 141 -

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

<210> 34

<211> 10731

<212> DNA

<213> Escherichia coli

<400> 34

ctacatccga caaaacgatg tcaaccatcc gaaaccgctc tcatccattc gatgagagcg 60

gttttttttaa ttactgctta aatgcacccg ccagagagcg aatatcattg ccggttggcg 120

actgatgaag tcgcagacca aactcttcga caatcgcaaa tatgtgatcg aatatatcag 180

cctgaatgct ttcatatcc agccacacca cggtgttagt aaacgcgtag atctcgagcg 240

gtaaaccggtt atcacctgga gccagttggc gtaccattaa ggtcatatct ttacgaatcc 300

gcggatggtt acgcagatat tcgttcagat aggcacgaaa ggttccaata ttggtcattc 360

- 142 -

ggcgcagatt taataccgac tccgtagaac cttgctggcg attccactca ttaatttcct 420
gatggcggct ggttaaataa ggctttaaca aatgcgcttt attcagacgt tgcatttcgt 480
cttcatcaag aaaacgaata ctggtgacat caatactgat actgcgctta atacgtcgcc 540
cgccagatgc tgacatcccc ctccagtttt taaaggagtc agaaaccaga gaccaggtgg 600
gaatagtggc aatgggtattg tcccagttac gcactttgac ggtgggttaac ccaatatcga 660
tcaccgcgcc atccgcgcgc tatttcggca tctccagcca gtgcgccagt ttcagcatat 720
cgttcgcgga aagctgaata cctgccacca gaccaagaat cggatcttta aataccaaca 780
tcagcacggc agccattgca ccaagaccgc tgatcagaat cgctggcgac tgaccaatca 840
gcagcgagat catcaaaatg ccgaccagaa tcgcgccgat cagtttaatc ccctgaaata 900
tccctttcag cggtaactga gatgctgccg ggaatttctg cgccagattc aaaataacat 960
ccagcaacga gaagactgaa agcagcgcat acatcatgat ccacaactgc gcgcaggtag 1020
tcagaatatc tgccgcttcg gtgccttttt gcagccagaa taccgcctga atattgacga 1080
taatcccctg cagggtaaaa gctaaacggc ggaagagttt attctgggta atgatttgca 1140
accaaagccg tgaactggcg atggcacgtt tttcgaaggc ccgcagtacc acccaatgca 1200
aaataatatg caccacgacg gcggtcagaa aaataatacc gaaaatcatc accatcgagg 1260
tggtgtgatc gatctcaata cccgctaaat cttcaacctg ggatattaaa tcctgcataa 1320
cgtctccttt atacaacagc agcctatgat gacggctgaa acagggttat gcaaatacag 1380
agaatctgag aggaaatagc cgggcagatg cccggcaaga gagaattaca cttcgggtta 1440
ggtgatgttt tgccgcagac gagatttcg cagcgtagcg ttaaaatctt caacgctgtg 1500
atgacctacc ggaacaacca ccagactggc gtagcctttc tctttcagac caaattctgc 1560
atcgaggatg gcggcgtaa aaccttcgat gggtagcgcg tccagacca gagccgccac 1620
gccgagcagg aagttaccga cgttgagata aacctgtttt gccatccact ctgcatcatc 1680
atgcagatct ttacggtgca tatcagcgaa gaacttgca cctttatcgt tcgcggcttt 1740
cgcttccggc gtggcaaagc ggccatcggc atcttcctgg tcaacaacca gcttcagcca 1800
gacatcgtcc atcgcggttt ttgcacagaa caccacgacg tgcgaggcat caagcatttt 1860
acgctcgttg aacacgtaat taccggcagc ggatttgga acacgcgctt taccttcttc 1920
cgtgctggca acaataaaat gccacggctg ggagttggc ctggatgggc tgtattgcag 1980
tagcgttttg atctgctcgg cctgttccgg ggtaagtttt ttgctggcat caaatgcctt 2040
agtggaatga cgctttaagg cgacagaaat gatatccata aagactccat gtgaaagtaa 2100

- 143 -

ttttgCGTgc cagcagatta caaggTtcaa cgagaaatgg taagcgagaa aaatgcgcta 2160
tagattttccg ctttagcctg gacgcactct tttttgatcg cgttttagcca gaccatcaac 2220
caccagattc cacgaatcgt tgatcagatc gcgaagtaac gcttcggaga tttcttcgcc 2280
gggatacact gaaatccagt gcttttttatt catgtgatac cctggccttaa tgcttgggta 2340
tatttgctga tttaacaggg atttttTgtg atcggacttc agattgataa aggggacgcc 2400
gcgtagctcc gacgacagca taaaaatctt gccgccaatt ttaaaaacat cgaactccgg 2460
gccaaaaggc cagcaaagct cgacaaaggg taactcaagg gccaggcgTt tcgccgtttc 2520
gtgcagtgat tgcttatcca taaacgttcc tttaggcgaa ggagaataag caaagtatgc 2580
cgcgaaGTac ggcgataatc gacgtttaat ccgccagcga gaaccagcgt cgccagataa 2640
agcgcagaac aaaatactca atagcgccca gcactaaaaa ccacagacaa aacaataaag 2700
tgtaaagctg actaagatcc atcagatgga acatggTcac cagtttttgt gccagcgcca 2760
gccccagtgc gggggcgggc agcagcaggc aagagaggaa agccatcagc aaaatgcctg 2820
ctgcggtggt caaggTttct aaagggtgTt tcataacatt gttaaTgta agttaaaaca 2880
ccattgtcag ggatattctt ctgtaaggca attcccggct tagtcaccgg cccagatctc 2940
acaatgcttt ttcaccagcc caatcagcga gccgccatcg gcgacgaaat cgcgcattag 3000
ctgcgcttca ttcagaccgc tgacgacctg gcgatgcagg gcctcaatcg cgctcgatgc 3060
accaatttta tgtgcggaag gggcgatttt ttcagcaat cgcaaggTat cttccgttag 3120
cggtcgacga tctccagtgt gcggatcggt gatgacgcct tcaagcccat agcgacaggc 3180
ctggaaacgg ttgaatttat acagcaggta atctttttcc tgatgtttaa acgggcgttc 3240
cgtcagtaac cagtgggcgg tagcctgaat taatcccgcc atatttactg cgtggctaag 3300
ggttaacggg gtatccatca cccgaacctc caccgtgcca aaatgaggac tggggcgaaT 3360
atcccagtgc agatctttaa tgctgtcgat catcgtggTg taactcagac agcgaaacag 3420
ggcttcaaTt tgttgccagt tactgacca cggcatcggg ccattatcag gaaaggcgga 3480
aaaaatattc ggtcgtgagg aggcaaaacg cgtatccgtt ccctgcatat atggcgacgc 3540
ggcggaagg gcgataaagt gcggcacaaa tcgtgacaag ccgtgcagca aataaatggc 3600
gtcatcgcca ctggcgcgag caacatggac atgctgacca aaaacggTcg cctgctgaat 3660
gagataacca aagttttcca gcgtgcgttg atagcgTtcg ttatcgcata cctcctgacg 3720
ctgccatttc tgaaacgggt gcgtgccacc gccgcaaatt tccagatgat ggtctgtggc 3780
tgcttgcaat acgactttct gcatcgctga aaactgcccg gcagcctggt tgatatcacg 3840
gcaaacatcc gtcgccagct ccagcatact ttcggtgata tcgtgcttta cctctccggc 3900

- 144 -

cgtgatctta tttttaaccg cgtcaatcag cattgaagag tcctggctta agtcatagcc 3960
cggcggatta accacctgca tttccagttc aataccgagg gtaaaagggt cagaaacatg 4020
aaaatcgggt aatggcatag gtttctctta tggtggcggt ttctattcag tatagaagtc 4080
ggagcggctg ggcgagatgc ggaagtctg gaatgtttct tttttttggt gatgggtgaac 4140
tgatgggtgcc tgaagcaatt tggctacttt tgcaatgtga caagttatgg cagtgtgac 4200
tggtggcgaa gaattttgac gattgagtca tgcagaaaaa aacgggttca gcttcagttg 4260
attaaatgag gagtgagaag tccgaaacag gactcactgt ataaataaac agctattttg 4320
ttgaggaagg gtaagataac ggcgggtgcc tgaagctttc cggtttcagg tttactctga 4380
ggtctggaaa gatgaagccc caggagatat ttctatcaac cctggggctg ccactccaaa 4440
cccgaacaat ttggatggta gtcccttctt cgcattggagg caatataaac atgctgacga 4500
aatatgccct tgcggcagtc atagtgtgt gttaaaccgt gctgggattt acgcttctgg 4560
tcggagactc gctgtgtgag ttacgggtga aggaacgtaa tattgagttt aaggctgttc 4620
tcgcttacga accgaagaag tagccgttgt gcggggagta atcccataag cgctaactta 4680
agggttgtgg tattacgcct gatatgattt aacgtgccga tgaattactc tcacgataac 4740
tggtcagcaa ttctggccca tattggtaag cccgaagaac tggatacttc ggcacgtaat 4800
gccggggctc taaccgccc cgcgaaatt cgtgatgctg caactctgct acgtctgggg 4860
ctggcttacg gccccgggg gatgtcatta cgtgaagtca ctgcatgggc tcagctccat 4920
gacgttgcaa cattatctga cgtggctctc ctgaagcggc tgcggaatgc cgccgactgg 4980
tttggcatac ttgccgcaca aacacttgct gtacgcgccg cagttacggg ttgtacaagc 5040
ggaaagagat tgcgtcttgt cgatggaaca gcaatcagtg cccccgggg cggcagcgt 5100
gaatggcgac tacatatggg atatgatcct catacctgtc agttcactga ttttgagcta 5160
accgacagca gagacgctga acggctggac cgatttgccg aaacggcaga cgagatacgc 5220
attgctgacc ggggattcgg ttgcgtccc gaatgtatcc gctcacttgc ttttgagaa 5280
gctgattata tcgtccgggt tcaactggcg ggattgogct ggttaactgc agaaggaatg 5340
cgctttgaca tgatgggttt tctgcgcggg ctggattgcg gtaagaacgg tgaaaccact 5400
gtaatgatag gcaattcagg taataaaaaa gccggagctc cttttccggc acgtctcatt 5460
gccgtatcac ttctcccga aaaagcatta atcagtaaaa cccgactgct cagcgagaat 5520
cgtcgaaaag gacgagtagt tcaggcggaa acgctggaag cagcgggcca tgtgctattg 5580
ctaacatcat taccggaaga tgaatattca gcagagcaag tggctgattg ttaccgtctg 5640

- 145 -

cgatggcaaa ttgaactggc ttttaagcgg ctcaaaagtt tgctgcacct ggatgctttg 5700
cgtgcaaagg aacctgaact cgcgaaagcg tggatatttg ctaatctact cgccgcattt 5760
ttaattgacg acataatcca gccatcgctg gatttcccc ccagaagtgc cggatccgaa 5820
aagaagaact aactcgttgt ggagaataac aaaaatggtc atctggagct tacagggtggc 5880
cattcgtggg acagtatccc tgacagccta caaaacgcaa ttgaagaacg cgaggcatcg 5940
tcttaacgag gcaccgaggc gtcgcattct tcagatgggt caacccttaa gttagcgctt 6000
atgggagtaa tccccgcata tccggttgct aggtcaggat ggtaaggcac ctgctttaca 6060
ctttcgcccg tggtcagtga tggctgcggg cgaatcgta cagatgttgt caattaatcg 6120
tgttggcaca gcgttatgac tatcttttct tttatctgcc agtgcacagc aaacatctca 6180
ttctcacgat gaatgatgac ctgctgttta ttccagctaa ttatctgata gtccagaaaa 6240
cctgcatcag tttggatctc acttgcccta aatgcgctct ctttggcgga aaatgccagt 6300
gtcagcgcca gagaaaaggc taaaccgcag tctgcgagtc gctcgtgttc cgctggtgta 6360
ataatgttgt ctgtcaattc tcttgcggtt tgtacagaaa aaatttcttc tatatcaatg 6420
ccaatcggtt gacgagatac cacggctaata gccgtagtcc cacagtggct aatactgccg 6480
tatacctccg caggccagac aggttggcgt agctcgccga ttgcgggcac acatttatag 6540
ccatattccc gcaaagcata aacagcagcg atccgtccgg ctaaagtctc tgttttacgt 6600
ttacgtccag cgtgttgacg ttgtgcgtag tgcggcagcc agagtaaata ctgctcacia 6660
aaattcgccg gatcgaaact aacaaaatgc agcgtatgtc cggcaaaggg gagggaggta 6720
tgcgtagttt tcatatcgac catattcgag actgatgaca aacgcaaaac tgcctgatgc 6780
gctacgctta tcaggcctac atggtcctgc aatatattga attggcaaga tttttgtagg 6840
ccggataagg cgttcacgcc gcatccggca tgaacgacgc gcactttgtc aacaatctga 6900
cgtttagcatc agaagtgggt gtttacgctc atataccacg tacgtcccgg ctcgttatag 6960
gtatacgccc cggcaccggc gatatagttg gccctgccca aatcgcccgt ggtctgggca 7020
ttaccgcac gccacaaacg tttgtcgaac agattgtcca cgccgccggt cagactgaca 7080
ttcttcgtca catcccaggt cgcgctcagg ccaacaatgc tgtaaggact aatttctttg 7140
gtttccggtc caaccgctgg ctgaccttta tagttgtact tcttcggctg ctgcttgccg 7200
taccagggtga aggtcgtttg catcgacaaa tcttcccgtg cctgccagct cagcgttgag 7260
ttcaacgtat actccgggat gatcgacaaa cggtcgcccg tggttttgtt ttcactcttc 7320
agcatataag tgatgttatt ggtccacatc accgtttcgc taaccggtac gtttaacgat 7380
ccttccagac cttcaaccac cgctttcggc acgttatccc actgatagag atcgggtgccg 7440

- 146 -

actgcgtttt gccctacagc cacatagccc gcttcaatct tattgcgata atcgttacgg 7500
aaccaggtga cgcccgccag ccacccgtcg cgtttgaact ccagaccaat ctctttgttg 7560
atgctggttt ctgctttcag gtcacgttta ccttgcagat agcagccgcc cgcgctggca 7620
tagcaaccct gacctttact gtagagaatg tagttcgggt tagtctggta caggctcggc 7680
gctttataag cacgggcgat gccattttc agcgtgaagt catcgccata accttgcgat 7740
atgttcagcg ccgggctcca gttattgccg acaatactgt gatgatcgaa acgcagcccc 7800
ggcgttacga tgggtgctgtc agtcagctcc atgttggttt cggcaaacag cgagaaaatt 7860
tctgcttttg aatacgggct acggtcgggt gtactcacgc catcaatagc gccaccggta 7920
ttcgttccgg tcagtgcctg ggtgttgga ctttaagtcct tcatccgttg ctgattccac 7980
tccgtacca gcgtcagcgt ctggttaacg aggaaatcaa tcggcagggt aacttcgctg 8040
tgcagcatca cgtcatcaag atcgatatcg acgaaatcct gtgtcgcttt ttcgttaaatt 8100
ttcccttcgg taccgcccgc cagaccttc ggaatacgcg agttacgggt gtgttcgtac 8160
tgcacccagt tgctgggtgg cagccgtta tcccagccac cgttccagggt cagcgcgtag 8220
ttctggcgat acagacgggt ggtttcatcg ccatatttcg agcgggtata ggaatcggag 8280
ttggtattct ggggtgtcgc cgcatacagg ttaccctggc ggctgtaacc tgcttccagt 8340
tccagcgatt gcaatggcgc gaaatcccag cgcaccacgc cattaatatc tttgttgatt 8400
acccttcgc gcccggtgg taacgtcgtg gcatacgttc cggcacgcgc ggactgatgg 8460
ccctgggtga tatcccacgc gtcagcctgg gttttgtcga gggtgccata caaacggaag 8520
ctgaattcgt cgcccagcgg accggtcagg ctaaagttag tgcgtttgggt ggcaccttc 8580
tctttatggt ctggcgcatt gaaatatgcg tcccaggagc cgtgccactc gccgctgcct 8640
tttttggtaa tgatgttaac cagccgccc gccgcgccgt tgccataacg cgcagctgcc 8700
ggaccacgca gaacttcaat acgttcaatc atttcagggt gcacccagga agtatcacca 8760
cgggtatcgc gctcgccacg ccagccctga cgcaccgagt tacggctgct taccggcttg 8820
ccgtcaatca aaatcagcgt gttttccgga cccataccgc gaatatcaat ctgtcggtta 8880
tcccacgct gaccactggt ggagttaccg gtcagggtta cgcctggcat ggtacggatg 8940
atcttcgaca catcgccggc aaccgggttt ttgcggattt catctgcggt gatggtcgaa 9000
acgccaggcg cctgtaagtt ctgctcggcg gcggtaacga caatagtatc gtcattgtgaa 9060
acaggagtat cggtcggctc ttgtgcctgc gctaccccat aaatccccag attgaccaac 9120
aaggccaggg aatgaatctt cttgttcatt gttttattcc tgcatttttg ccacgaattg 9180

- 147 -

caactgtcgg gcatggtcgt catcaacacg acgcatcccg ctaccgcgaa aacctttgat 9240
cctgaaagac acgcagtgca gttgggttaat taatgtccgc gcttcccacg gcgcgccatt 9300
acgctattgc aaatgcaaat agttatcaat aatattatca atatatttct gcaatcaatg 9360
aaaaattgca cagtaaacad ggggttatgg tgtgacggcg ttaaaagtag gaagtgagag 9420
ctggtggcag tcgaaacatg gcccggaatg gcagcgtctg aatgacgaaa tgtttgaggt 9480
cactttctgg tggcgtgatc cccaagggtc tgaagaatac tcgacgataa agcgcgtatg 9540
ggtctacatc actggtgtga ccgatcacca tcagaacagc cagccccagt cgatgcagcg 9600
aattgcaggc actaacgtct ggcagtggac gacacaactc aatgccaaact ggcgcggcag 9660
ctactgcttt attcccaccg aacgcgatga ctttttttct gtaccatccc ccgatcgctt 9720
cgaattgcgc gaaggctggc gaaaactatt accccaggcg atagccgatc cgctgaacct 9780
acaaagctgg aaaggcgggc gagggcacgc tgtttctgca ctcgaaatgc cgcaagcgcc 9840
tctgcaaccg ggatgggatt gtccgcaagc gccagaaata cctgccaaag aaattatctg 9900
gaaaagtga cggttgaaaa agtcacggcg tgtatggatt tttaccaccg gcgatgcaac 9960
agcagaagaa cggccgctgg cagttttgct cgatggcgaa ttttgggcgc aaagtatgcc 10020
cgtctggcca gtgctgactt cgctgacca tcgtcagcaa cttcctcccg ccgtgtatgt 10080
gttgatcgac gctatcgaca ccacgcaccg cggccacgaa ctgccgtgta atgcggatgt 10140
ctggctcgca gtacagcaag agttattacc cctggtgaaa gctattgccc ctttttagcga 10200
tcgtgccgat cgcaccgtgg ttgccgggca gagttttggg gggctttccg cgctgtatgc 10260
cggactgcac tggcctgaac gctttggctg tgtattaagc cagtcaggat cgtactgggtg 10320
gccgcacgag ggccgggcagc aagagggcgt gttacttgaa aagctaaaag ctgggtgaagt 10380
tagcgccgaa ggtctgcgca ttgtgctgga agcgggtatt cgcgagccga tgatcatgcg 10440
ggccaatcag gcgctgtatg cgcaattaca ccccataaaa gaatccattt tctggcgtca 10500
ggttgacggc ggacatgatg cgctttgttg gcgcgggtggc ttgatgcagg ggctaatacga 10560
cctctggcaa ccacttttcc atgacaggag ttgaatatgg cattcagtaa tcccttcgat 10620
gatccgcagg gagcgtttta catattgcgc aatgcgcagg ggcaattcag tctgtggccg 10680
caacaatgcg tcttaccggc aggctgggac attgtgtgtc agccgcagtc a 10731

<210> 35

<211> 217

<212> PRT

<213> Escherichia coli

<220>

- 148 -

<223> complement of position 1426-2079 of seq id 34

<400> 35

Met	Asp	Ile	Ile	Ser	Val	Ala	Leu	Lys	Arg	His	Ser	Thr	Lys	Ala	Phe
1				5					10					15	

Asp	Ala	Ser	Lys	Lys	Leu	Thr	Pro	Glu	Gln	Ala	Glu	Gln	Ile	Lys	Thr
			20					25					30		

Leu	Leu	Gln	Tyr	Ser	Pro	Ser	Ser	Thr	Asn	Ser	Gln	Pro	Trp	His	Phe
		35					40					45			

Ile	Val	Ala	Ser	Thr	Glu	Glu	Gly	Lys	Ala	Arg	Val	Ala	Lys	Ser	Ala
	50					55					60				

Ala	Gly	Asn	Tyr	Val	Phe	Asn	Glu	Arg	Lys	Met	Leu	Asp	Ala	Ser	His
65					70					75					80

Val	Val	Val	Phe	Cys	Ala	Lys	Thr	Ala	Met	Asp	Asp	Val	Trp	Leu	Lys
				85					90					95	

Leu	Val	Val	Asp	Gln	Glu	Asp	Ala	Asp	Gly	Arg	Phe	Ala	Thr	Pro	Glu
			100					105					110		

Ala	Lys	Ala	Ala	Asn	Asp	Lys	Gly	Arg	Lys	Phe	Phe	Ala	Asp	Met	His
		115					120					125			

Arg	Lys	Asp	Leu	His	Asp	Asp	Ala	Glu	Trp	Met	Ala	Lys	Gln	Val	Tyr
	130					135					140				

Leu	Asn	Val	Gly	Asn	Phe	Leu	Leu	Gly	Val	Ala	Ala	Leu	Gly	Leu	Asp
145					150					155					160

Ala	Val	Pro	Ile	Glu	Gly	Phe	Asp	Ala	Ala	Ile	Leu	Asp	Ala	Glu	Phe
				165					170					175	

Gly	Leu	Lys	Glu	Lys	Gly	Tyr	Thr	Ser	Leu	Val	Val	Val	Pro	Val	Gly
			180					185					190		

His	His	Ser	Val	Glu	Asp	Phe	Asn	Ala	Thr	Leu	Pro	Lys	Ser	Arg	Leu
		195					200					205			

Pro	Gln	Asn	Ile	Thr	Leu	Thr	Glu	Val
	210					215		

<210> 36

<211> 14674

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (469)..(2118)

<400> 36

acgggtgggc gcagagcgat acttcgctac tattttcacc cagaagtgtc ccaccacttg 60

cgaaacgccc gactgcgaac gcttctgggtg acaaccagg ggattcagcc cctgtagccg 120

- 149 -

atgatgaacg tggccagccg ttcaatcacc tcggcgatgc accccctcag gtgttatcac															180
aggactggct cctccaacac cgttacttgg gcaacgcgcc tcttctggcc tgcgctagcg															240
caggtagtac atttataaat aaaggggtgag cggggcggtt gtcaacgatg gggtcatgcg															300
gatttttcat ccactcctgg cggtcagtag ttcagctaataaatgcttca ctgcgctaag															360
ggtttacact caacattacg ctaacggcac taaaaccatcacatttttct gtgactggcg															420
ctacaatctt ccaaagtcac aatttctcaaa atcagaagag tattgcta atg aaa aac															477
															Met Lys Asn
															1
atc aat cca acg cag acc gct gcc tgg cag gca cta cag aaa cac ttc															525
Ile Asn Pro Thr Gln Thr Ala Ala Trp Gln Ala Leu Gln Lys His Phe															
5 10 15															
gat gaa atg aaa gac gtt acg atc gcc gat ctt ttt gct aaa gac ggc															573
Asp Glu Met Lys Asp Val Thr Ile Ala Asp Leu Phe Ala Lys Asp Gly															
20 25 30 35															
gat cgt ttt tct aag ttc tcc gca acc ttc gac gat cag atg ctg gtg															621
Asp Arg Phe Ser Lys Phe Ser Ala Thr Phe Asp Asp Gln Met Leu Val															
40 45 50															
gat tac tcc aaa aac cgc atc act gaa gag acg ctg gcg aaa tta cag															669
Asp Tyr Ser Lys Asn Arg Ile Thr Glu Glu Thr Leu Ala Lys Leu Gln															
55 60 65															
gat ctg gcg aaa gag tgc gat ctg gcg ggc gcg att aag tcg atg ttc															717
Asp Leu Ala Lys Glu Cys Asp Leu Ala Gly Ala Ile Lys Ser Met Phe															
70 75 80															
tct ggc gag aag atc aac cgc act gaa aac cgc gcc gtg ctg cac gta															765
Ser Gly Glu Lys Ile Asn Arg Thr Glu Asn Arg Ala Val Leu His Val															
85 90 95															
gcg ctg cgt aac cgt agc aat acc ccg att ttg gtt gat ggc aaa gac															813
Ala Leu Arg Asn Arg Ser Asn Thr Pro Ile Leu Val Asp Gly Lys Asp															
100 105 110 115															
gta atg ccg gaa gtc aac gcg gtg ctg gag aag atg aaa acc ttc tca															861
Val Met Pro Glu Val Asn Ala Val Leu Glu Lys Met Lys Thr Phe Ser															
120 125 130															
gaa gcg att att tcc ggt gag tgg aaa ggt tat acc ggc aaa gca atc															909
Glu Ala Ile Ile Ser Gly Glu Trp Lys Gly Tyr Thr Gly Lys Ala Ile															
135 140 145															
act gac gta gtg aac atc ggg atc ggc ggt tct gac ctc ggc cca tac															957
Thr Asp Val Val Asn Ile Gly Ile Gly Gly Ser Asp Leu Gly Pro Tyr															
150 155 160															
atg gtg acc gaa gct ctg cgt ccg tac aaa aac cac ctg aac atg cac															1005
Met Val Thr Glu Ala Leu Arg Pro Tyr Lys Asn His Leu Asn Met His															
165 170 175															

- 150 -

ttt gtt tct aac gtc gat ggg act cac atc gcg gaa gtg ctg aaa aaa	1053
Phe Val Ser Asn Val Asp Gly Thr His Ile Ala Glu Val Leu Lys Lys	
180 185 190 195	
gta aac ccg gaa acc acg ctg ttc ttg gta gca tct aaa acc ttc acc	1101
Val Asn Pro Glu Thr Thr Leu Phe Leu Val Ala Ser Lys Thr Phe Thr	
200 205 210	
act cag gaa act atg acc aac gcc cat agc gcg cgt gac tgg ttc ctg	1149
Thr Gln Glu Thr Met Thr Asn Ala His Ser Ala Arg Asp Trp Phe Leu	
215 220 225	
aaa gcg gca ggt gat gaa aaa cac gtt gca aaa cac ttt gcg gcg ctt	1197
Lys Ala Ala Gly Asp Glu Lys His Val Ala Lys His Phe Ala Ala Leu	
230 235 240	
tcc acc aat gcc aaa gcc gtt ggc gag ttt ggt att gat act gcc aac	1245
Ser Thr Asn Ala Lys Ala Val Gly Glu Phe Gly Ile Asp Thr Ala Asn	
245 250 255	
atg ttc gag ttc tgg gac tgg gtt ggc ggc cgt tac tct ttg tgg tca	1293
Met Phe Glu Phe Trp Asp Trp Val Gly Gly Arg Tyr Ser Leu Trp Ser	
260 265 270 275	
gcg att ggc ctg tcg att gtt ctc tcc atc ggc ttt gat aac ttc gtt	1341
Ala Ile Gly Leu Ser Ile Val Leu Ser Ile Gly Phe Asp Asn Phe Val	
280 285 290	
gaa ctg ctt tcc ggc gca cac gcg atg gac aag cat ttc tcc acc acg	1389
Glu Leu Leu Ser Gly Ala His Ala Met Asp Lys His Phe Ser Thr Thr	
295 300 305	
cct gcc gag aaa aac ctg cct gta ctg ctg gcg ctg att ggc atc tgg	1437
Pro Ala Glu Lys Asn Leu Pro Val Leu Leu Ala Leu Ile Gly Ile Trp	
310 315 320	
tac aac aat ttc ttt ggt gcg gaa act gaa gcg att ctg ccg tat gac	1485
Tyr Asn Asn Phe Phe Gly Ala Glu Thr Glu Ala Ile Leu Pro Tyr Asp	
325 330 335	
cag tat atg cac cgt ttc gcg gcg tac ttc cag cag ggc aat atg gag	1533
Gln Tyr Met His Arg Phe Ala Ala Tyr Phe Gln Gln Gly Asn Met Glu	
340 345 350 355	
tcc aac ggt aag tat gtt gac cgt aac ggt aac gtt gtg gat tac cag	1581
Ser Asn Gly Lys Tyr Val Asp Arg Asn Gly Asn Val Val Asp Tyr Gln	
360 365 370	
act ggc ccg att atc tgg ggt gaa cca ggc act aac ggt cag cac gcg	1629
Thr Gly Pro Ile Ile Trp Gly Glu Pro Gly Thr Asn Gly Gln His Ala	
375 380 385	
ttc tac cag ctg atc cac cag gga acc aaa atg gta ccg tgc gat ttc	1677
Phe Tyr Gln Leu Ile His Gln Gly Thr Lys Met Val Pro Cys Asp Phe	
390 395 400	
atc gct ccg gct atc acc cat aac ccg ctc tct gat cat cac cag aaa	1725
Ile Ala Pro Ala Ile Thr His Asn Pro Leu Ser Asp His His Gln Lys	
405 410 415	

- 151 -

ctg ctg tct aac ttc ttc gcc cag acc gaa gcg ctg gcg ttt ggt aaa	1773
Leu Leu Ser Asn Phe Phe Ala Gln Thr Glu Ala Leu Ala Phe Gly Lys	
420 425 430 435	
tcc cgc gaa gtg gtt gag cag gaa tat cgt gat cag ggt aaa gat ccg	1821
Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly Lys Asp Pro	
440 445 450	
gca acg ctt gac tac gtg gtg ccg ttc aaa gta ttc gaa ggt aac cgc	1869
Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu Gly Asn Arg	
455 460 465	
ccg acc aac tcc atc ctg ctg cgt gaa atc act ccg ttc agc ctg ggt	1917
Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe Ser Leu Gly	
470 475 480	
gcg ttg att gcg ctg tat gag cac aaa atc ttt act cag ggc gtg atc	1965
Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln Gly Val Ile	
485 490 495	
ctg aac atc ttc acc ttc gac cag tgg ggc gtg gaa ctg ggt aaa cag	2013
Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu Gly Lys Gln	
500 505 510 515	
ctg gcg aac cgt att ctg cca gag ctg aaa gat gat aaa gaa atc agc	2061
Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Asp Asp Lys Glu Ile Ser	
520 525 530	
agc cac gat agc tcg acc aat ggt ctg att aac cgc tat aaa gcg tgg	2109
Ser His Asp Ser Ser Thr Asn Gly Leu Ile Asn Arg Tyr Lys Ala Trp	
535 540 545	
cgc ggt taa tcacgtcga tatgtaggcc ggataaggcg ttcacgcgc	2158
Arg Gly	
550	
atccggcaac cgatgcctga tgcgacgcgg tcgcgtctta tcaggcctac aggtcgatgc	2218
cgatatgtac atcgtattcg gcaattaata catagcacga ttgattaaat aaccttaata	2278
acaatgccga cgttatgtcg gcattttttt atcagataaa tccccttgtc tgtaatttaa	2338
cggaaatcat accgtgaggt taatcctaaa atagattttt aatcgttggt tatttcggaa	2398
aatacgcaga ttaattgctt ttgtttttat tttaagttta tgatttttat tgttatttaa	2458
atataagttg aaacttatat ttgatattca ttccaattat cctaaaacgc catcgcta	2518
tccccgcgcc gtaattcgca tgcttttagtt gtgtatactc gatccccgcc gaaatgtttt	2578
tgggtaaadc tccattcatt caatgaaggg aaattgttat gaaaaagtt ctgtatggca	2638
tttttgccat atctgcgctt gcggcgactt ctgcgtgggc tgcacctgta caggtggg	2698
aagcggcagg gtcggcagca acgtcggttt cggcggggag ttctccgcg accagcgtca	2758
gcaccgtaag ctcggcggtg ggtgtcgcgc ttgcggcaac cgggtggcgt gatggttcta	2818

- 152 -

ataccgggac caccacaacc acgaccacca gtaccacagta ataaagtatg tatccccaaa 2878
ataattcgag tcattgcatc tgtggctaga agtatgaagg gattaacccat aaccacactc 2938
cgggtgtggtt attctgcccc tctggagaag agtcgtgaag cgacctgcac tcattcttat 2998
ctgcctgtta ttacaggcct gttcagccac gactaaagag ctgggcaatt cactgtggga 3058
cagtctgttc ggcacgccag gcgtacagct gacggatgat gatattcaaa atatgcccta 3118
cgccagccag tacatgcagc ttaatggcgg gccgcagtta tttgtggtgc tggccttcgc 3178
tgaagacgga caacaaaaat gggtcactca ggatcaggct actctcgtta cccaacatgg 3238
ccgtctggtg aagactctgc tcggcggcga caacctgatt gaagtgaata accttgctgc 3298
cgacccgctg attaaacccg cacaaatcgt tgatggcgca agctggacgc gcacgatggg 3358
ctggaccgag taccagcagg tacgctacgc caccgcacgc tcagtcttca aatgggatgg 3418
caccgatacc gtcaaagtcg gcagcgatga aaccccggtt cgcgtgctgg acgaagaagt 3478
ctccaccgac caggcgcgct ggcataaccg ctattggatc gacagcgaag ggcaaattcg 3538
ccagtcggaa cagtatctcg gcgcggatta ttttccggtg aaaaccactc tcatcaaggc 3598
ggcaaaacaa tgattaaaca aactattgtc gcgttgcttt tgagtgtggg agcgtcatcg 3658
gtcttcgcgg caggaaccgt caaggtgttc agcaatggca gcagtgaggc caaaacgcta 3718
acgggcgag agcatttaat cgatctggta ggccaaccgc ggctggcaaa cagctggtgg 3778
cccggtgcgg tgattagcga agagctggca acggcggcag cattgcgtca gcagcaggcg 3838
ttgctgacac ggctggcaga acagggcgca gattccagcg ccgacgatgc cgctgcgatt 3898
aacgccttac gccagcaaatt tcaggcggtg aaagtacgg gcaggcaaaa aatcaatctt 3958
gatcccgata tcgtccgctg tgccgaacgc ggtaaccgc cgttgcaggg caactacacg 4018
ctgtgggtcg gaccaccgcc gtccacggtc acgttggtcg ggcttatcag ccgtcctggc 4078
aagcagccat tcactccagg tcgcgacgtg gcgagctatc tctctgacca gagcctgctc 4138
agcggtgcgg atcgcagcta cgcgtgggtg gtttaccgg acggacgcac gcaaaaagcg 4198
ccggtggctt actggaacaa gcgtcacgta gagccgatgc ccggcagcat tatttatgtt 4258
ggcctcgcgg actccgtctg gagtgagacg cctgatgcc ttaacgccga cattcttcag 4318
actctgacgc agcggatacc tcaataatga aaaaaagaca tctgcttagc ttactggcgc 4378
tgggcattag cacagcttgc tacggcgaaa catatcctgc gccattggg ccgtcgcagt 4438
cggatttcgg tggcgtagga ttattacaaa cgcccaccgc gcgtatggca cgggaagggg 4498
agttgagtct gaactatcg gataacgatc agtaccgtta ttactcagct tcagtgaac 4558
tcttcccggtg gctggaaaca acgctgcgct acaccgacgt gcgcaccgg cagtacagca 4618

- 153 -

gcgtcgaagc gttctctggc gatcaaactg ataaagataa agccttcgat ctcaaactgc 4678
gtttgtggga agagagttac tggctgccgc aagtggcggg tgggtgcgcgg gatatcggcg 4738
gtacgggggt gtttgatgcg gaatatcttg ttgccagcaa agcctggggg ccgttcgatt 4798
ttacgctcgg cctgggctgg gggatatttg gcaccagcgg taatgtgaaa aatccgctct 4858
gttcagccag tgataaatat tgctatcgcg ataacagcta caaacaggcg ggatctatcg 4918
acggtagcca gatgttccac ggtcctgcct cactgttttg cggcgtggaa taccagacgc 4978
cctggcaacc gctgcgcctg aaactggagt atgaaggcaa taattatcag caggattttg 5038
ccgggaagct ggagcaaaaa agtaagttaa acgtcgggtg gatattatcg gttaccgatt 5098
gggcccagct taaccttagc tatgaacgtg gcaacacctt tatgtttggc gttacgttgc 5158
gcaccaactt taacgatctg cggccgtctt acaacgataa cggccgcccg caatatcaac 5218
cgcagccgca ggatgccatt ttgcagcatt cgggtgggtg gaatcagtta acgctgttga 5278
aatacaatgc cggacttgcc gatccacaga tccaggcgaa aggcgatacg ctgtatgtta 5338
ccggcgagca ggtgaaatat cgtgatctgc gcgaagggat catccgtgcg aatcggatcg 5398
tgatgaacga tctgccggat gggatcaaaa cgatccgcat tacggaaaat cgccttaaca 5458
tgccgcaggt gacgacggaa accgatgtcg ccagcctgaa aaatcatctc gccggagagc 5518
cgttggggca cgaaacgacg ctggcgcaaa aacgcgtcga gccagtgggt ccgcagtcca 5578
ccgagcaggg ctggtatatc gacaaatcac gctttgattt ccatatcgat ccggtgctga 5638
accagtcggt cgggtggccc gaaaactttt acatgtatca gctgggcgtg atgggaacgg 5698
cagatttgtg gctgacggac catctgctga ccaccggcag cctgtttgca aatcttgcca 5758
acaactacga caagtttaac tacactaatc ctccgcagga ctgcactta ccgcgcgtgc 5818
gtacccatgt gcgcgagtat gtgcagaacg atgtctatgt gaataacctg caagccaact 5878
acttccagca tctgggcaac ggcttctacg gtcagggtta cgggtggttat ctcgaaacca 5938
tgtttggcgg tgcgggggca gaagtgttgt atcgcccgtg ggacagcaac tgggcgtttg 5998
gtctggatgc caactacgtt aaacagcgcg actggcgtag tgcaaaagat atgatgaaat 6058
tcaccgacta cagcgtgaaa accggacatc tgaccgccta ctggacgcca tctttcgctc 6118
aggatgtgtt agttaagcc agcgtcgggc agtatctggc aggggataaa ggcggcacgc 6178
tgagatcgc caaacgcttt gatagcggcg tgggtgggtgg tggctatgcc acgatcacta 6238
atgtttcgaa agaggagtac ggcgaagggg acttcaccaa aggcgtgtat gtctcggtac 6298
cgttggatct cttctcgtct ggcccagac gcagccgtgc ggcgattggc tggacgccgc 6358

- 154 -

tgacgcgtga cgggtggtcag caacttgggc gtaagttcca gttgtatgac atgaccagcg 6418
accgtagcgt caatttccgc taagtcattg gaaaggtgcc agttttcgca ctactggca 6478
ccttcattct taatggcctg tcgagccatc aggaacgcag ttactcccg gtttgcgatc 6538
tggtgtgtaa aggcgaaaac attctgtctt cctgtccgtc gatatcggtt gcaggtcttg 6598
ttgcctgagc gactgggcca tcattgggct ggagacattg ctttgcggtta ggtcggcttt 6658
atcagcagta ccgagcggac cagcataagc aggaagaaca gagactgata acatcaaagc 6718
agcaaaataa ggcttcattt ttaccacctt tatcaggtta cgtttcattt gttccagagg 6778
aacattgtcg atttttcgcg cattgctggt ggctgggaat cacctgaatg ggtgattttt 6838
gaattaccgg ctttggtgcg gtttgtcttg ccgatgcgc cgccaggcgc ggcttatccg 6898
gcctacgggt aggtatatcc ggctttggtg gaggcgcgct ccaaaccag gttgaacaaa 6958
acatacacia aaaatataga tctccgtcac atttttgctg tatacaggaa gctcgccact 7018
gtgaaggagg tactgctatg acgtcactct ctggtccgcg cgtggagttt atctccacca 7078
ttttgcagac cgtactcaat cttggcctgc tgtgcctcgg cctgattttg gttgtcttcc 7138
tcggcaaaga aacggtgcat ctggctgatg tgctgttcgc gccagaacaa accagcaaat 7198
atgagctggt agaaggactg gtggtttact ttctctatct cgaatttatc gcgctgattg 7258
tgaagtactt tcagtccggt ttactcttcc cgctacgtta ctttgtctat attgggatca 7318
ccgccattgt gcgcttgatc atcgtcgatc ataaatcgcc actggatgtg ttgatctact 7378
cggctgcgat cctgttactg gtgatcacc tgtggctgtg taattcgaaa cggctgaagc 7438
gggagtaaaa agtcagcacg ccgaaatggc gcggcgtgct ggacaggaag attacagcgt 7498
agcagtttgt tgtgttttct tcgtttccgg ttcccagagc gcttccagct cctcaagggt 7558
tttacctttg gtttccggga caaatttcca cataaacagt gctgccagaa cgccataca 7618
accgtaaata cagtaggaga aaccgttgtg gaaatgggcc accagccagg agtttttgtc 7678
catcatcggg aaggtccagg agacgaagta gttcgccagc cactgggccc ccaccgcgat 7738
tgccagcgct ttaccacgaa tagcattcgg gaagatttcc gacagcagta cccagcatac 7798
cggaccccag gacatggcaa aggcggcaac atagaacagc atcgacagta gcgccacaat 7858
acccggtgcc tgagtgtaaa acgcggtacc gaggctaaac ataccgattg ccattccgag 7918
tgccgcgata atttgagtg gcttacgacc aaatttatcc accgtcataa ttgccagaac 7978
ggtgaagggt aggttgataa ctccgacaat aatggtctgc aacagcgca tatccgtgct 8038
ggccccagc gttttgaaca cttccggcgc gtagtacagc accacattga tgccgacaaa 8098
ttgctggaag atggagagca ttacgccgat tacaatcacg ccacgccaac acatcagcag 8158

- 155 -

acgaccaccg gttttgcggc catgatccag ggagtgttta atttcctgta ctgcctgagt 8218
tgcaagcgtg ttgcccataa ttttgcgcag gataccttcc gcctgttctt gcttgccgcg 8278
cgacatcagc cagcgaggac tttctggcac ggtatacagc agcattaaga acagcagtgc 8338
agggatacat tccgaggcaa acatataacg ccagccgtca gtattcagcc agctggcatc 8398
accggaacgg gcaataaaat agtttacgca gtaaactaaa agttgcccgga aaataatcgc 8458
aaactgggtta aaagagacca gtttcccgcg aatatgagct ggagccagtt ccgcaatata 8518
cattggcgag agcattgagg ctaaaccaac gccaataccg ccaataatgc gataaataac 8578
aaattccggg acataacctg ccagataaac aggcacagtg ttgtccgggt ttatagaggt 8638
aaaaccaagt tctggccagg cagaacctac accagaaata aaaaacagga cagcagcaat 8698
cttaagtga tccagacgac cgaagcgggt actgcaataa ccaccgaggg caccgccgat 8758
gatgcaacca atcagagcgc tggccacgca aaacctaac agggagttgg cagcggattc 8818
acttaagttt tgtggagcaa caaagacggt attgagtac tcaacagtac cggaaataac 8878
ggcgggtgtc tagccaaata ataaaccacc taatgtagcg actaaggtaa tcgaaaatat 8938
ataactggaa ttatactggg tattcattca gacctgcctt agaccattct gatcttttat 8998
cagtgcattg atgttggaat tgtcaggcag ttaaggcaat tatcattttt tacaatgcgt 9058
ttacgtaatg tcttttctgt gatcttaatt gtgataatta tccaaaattg ataaaaacaa 9118
tactattgcc gtgactcaga gcacgaaaga gaattatcgt aagtgygaaa acaataaacg 9178
taaaaataat aagctctatg atggaaatat taaccggcga acgattcaga ttgcagacga 9238
aagaaaaaaa ggcgtccgt ggagcgccga ataacagtca caagttggga taacgtaagt 9298
tgaggggtgca gcggcataac attggcagaa caacatcttt aacctttcac accacctgcc 9358
gtcaggccgt tgaccagcca gcgttgagcc agcaagaaga cgatggatgat cggtaatgca 9418
gacatcacgg cagcggcggc aaagtcaccc cacaggtagt tttgcgggtt gaggtattgc 9478
tgcatcccca cggccagggt gtagctgttt acgtcacgca gtaacagcga cgcgaccgga 9538
acttcagtaa tggcagcgat aaacgacagg atgaatacca ccgccagaat cggactgac 9598
agcggcaaca ggacaaggcg gaaggcctgc cacggtgtcg caccatccag cgcagcagct 9658
tcttccagcg aactgtcgat ggtttcgaaa tagcctttga tgggtccagac atgcagcgca 9718
atcccacca gatacgcgaa aattacgccg ccgtgagtat tcaggccaat gaatggaatg 9778
tactacca gaagatcaaa caacgcatac aacgcgacca gtgaaagtac tgccgggaac 9838
atctggaaaa tcagcattcc tttcagcagc gtcgctttgc ctggaaagcg catacgggcg 9898

- 156 -

aaagcgtagg cgcaggtggt ggagagcgcc acaatgccaa tcgcggaaat cccggcgacc 9958
tttaccgagt tccacagcca cagcagtacc gggaatggcg gtggcgtaat gcgaccatca 10018
gcctgttcaa cgctaaaacc taacgccagt ttccagtgat cccaggagat ttgctccggg 10078
atcaggctgc cggtcgcaaa gtttccctga cgcagcgaga tagcgacgac catcagcagc 10138
gggaacataa tcgctgcgat aaaaagtagc agtagcaggt gagtaataaa taaacgagct 10198
ttttgcgatt tcggttggac cattgccatt tttgttatct cccttaatca aacttcattc 10258
gcgtggcttt caggttcact atcgccagcg caccaccag caggaagatc agcgtggcaa 10318
ttgctgccgc cagaccgaag tcctgacccc cgccgccttc aaaagcgatg cggtaggtgt 10378
agttaacaag caggtcggta taaccggctg gcgtggtcgt gccaagacga tccgggcccgc 10438
cgttgggttaa cagttgaatc agcacgaagt tggttaaagtt aaaggcgaag ctggcgatca 10498
tcagcggcgt cagcggttta atcagcagcg gcagcgtaat cttaaagaag ttctggaacg 10558
gacctgcgcc atccattgct gaggcttcat acaaatcgtc cggaatcgct ttcagcaagc 10618
ccatgcagag gatcatcatg tacggataac ccagccaggt attgacgata attagcatcg 10678
tgccggcggt ggtcggatcg ctgaaccagg cgggcttcac gccaaacagc gcgctcaaca 10738
tcatgttgat ttcaccgaag ctctgggttaa acaacccttt gaaaatcaag attgaaatga 10798
acgatggcac cgcgtagggc agaatcagca ggacgcgata gaccgctttg ccgcgcaacg 10858
cttcccactg caccagacac gccagaacca tgccgaccgc caccgttaaa aagacagtga 10918
tcagcgagaa caccacggtc cagacgaaaa tggcgaggaa cggtttctga atgccttcgt 10978
cggtaaagac gcgggtaaag tttttccagc cggtggtcac ggtgtaaccg gggcttagct 11038
tttcatcacc ccagttgccg tcggcggtaa tggactggta aaagccaatt tggttattcg 11098
gacgatatth caccgctc tgattattcg tcaacgtgcc gtcaccgtcg agtgtgtaga 11158
gcggctgcgt gccagaaaac tggcgaggag agtcatcat cactttgttg ccatccggca 11218
gaatggcggt aatgtcactc agcgctgac gattctgggt aatcacgcgc agattcgcgc 11278
gttcgccttc gggctgggag gtcgtttctt tcagttgcag tttttgctcg ccgccaatt 11338
taaaagcgtc ggagaggtaa tttttgccgg tttcgccgtc gctgagcgcc agttgccact 11398
catcgccgc cgggtaaaga ccaaagttat aggttttgcc tgcttgccag gacgatcta 11458
acaacacttc ctgcgcacgt tcaaaagtca gctggtagt gctgctgtag ttggtgaagg 11518
caatggcgat ggtgcagacc agagggaaga ggacgaataa tccattcca gccattcccg 11578
ggtaaacata gcgccaggcg taggctttac gattggcgaa aatatacagc cccgctgaac 11638
tcaatatcag cgtggtaatg gcgaacaggt attccccttg tgcgtacatt aaaacaaca 11698

- 157 -

ggtaaccac cagcaggccg agcagaccta gcactgacca tttcagcgcg tcgctttgcc 11758
accaatgttt ctttttaatg acatccatgg ggttcttcct cattccagga cggataaggc 11818
tttcacgcct tatcogacaa caactgcctg atgcgacgct gacgcgtctt atcaggccta 11878
catacgtttc ggttttgtag gccggacaag gcgttcacgc cgcattccggc atttcacagc 11938
attacttggg gatacgagtc tgcgcgtctt tcagggttc atcgacagtc tgacgaccgc 11998
tggcggcggt gatcaccgca gtacgcacgg cataccagaa agcggacatc tgcgggatgt 12058
tcggcatgat ttcacctttc tggggcgttt ccatgggtggc ggcaatacgt ggatctttcg 12118
ccaactcttc ctcgtaagac ttcagcgcta cggcaccag cggtttgtct ttattaaccg 12178
cttccagacc ttcattcagtc agcagatagt tttcgaggaa ctctttcgcc agctctttgt 12238
tcggactggc ggcgttaata cctgcgctca gcacgccaac gaacggtttg gatggttgac 12298
ccttgaagggt cggcagtacc gttacaccat aattcacttt gctgggtgtcg atgttggacc 12358
atgcccacgg gccgttgatg gtcattcgctg tttcgctttt attaaaggca gcttctgcga 12418
tggagtaatc ggtgtctgca ttcattgtgtt tgtttttaat cagggtcaacc aggaagggtca 12478
gacctgcttt cgcgccagcg ttatccacgc ccacgtcttt aatgtcgtac ttgccgtttt 12538
catacttgaa cgcataacc cgcgcagcag caatcagcgg ccagggtgaag tacgggtctt 12598
gcagggtgaa catcagcgcg ctcttacctt tcgctttcag ttctttatcc agcgccggga 12658
tctcttccca ggtttttggc ggggttcggca gcagatcttt gttataaatc agcgataacg 12718
cttcaacagc gatcgggtaa gcaatcagct tgccgttgta acgtacggca tcccaggtaa 12778
acggatacag cttgtcctgg aacgctttgt cgggggtgat ttcagccaac aggccagatt 12838
gagcgtagcc accaaagcgg tcgtgtgccc agaagataat gtcagggtcca tcgccagttg 12898
cgcgaacctg tgggaatttc tcttccagtt tatcoggatg ctcaacgggtg actttaattc 12958
cggatatctt ctogaatttc ttaccgactt cagcgagacc gttatagcct ttatcgccgt 13018
taatccagat taccagttta ccttcttcga ttttggcgag agccgaggcg gaaaacatca 13078
tcgtcgtaa tgccgataat gcgaggatgc gtgcacctgt ttttatcttc ataactatg 13138
gtccttggtg gtgaagtgc cgtgaaaaca cctaaacgga ctctagtttc tttatacggc 13198
aacctctttc catcctcctt gccctacgc cccaccgtcg ctttgtgtga tctctgttac 13258
agaattggcg gtaatgtgga gatgcgcaca taaaatcgcc acgatttttg caagcaacat 13318
cacgaaattc cttacatgac ctcggttttag ttcacagaag ccgtgttctc atcctccgc 13378
ctcctcccc ataaaaaagc caggggggtgg aggatttaag ccatctcctg atgacgcata 13438

- 158 -

gtcagcccat catgaatgtt gctgtcgatg acaggttggt acaaagggag aagggcatgg 13498
 cgagcgtaca gctgcaaaat gtaacgaaag cctggggcga ggctgtggta tcgaaagata 13558
 tcaatctcga tatccatgaa ggtgaattcg tgggtgtttgt cggaccgtct ggctgcggta 13618
 aatcgacttt actgcgcatg attgccgggc ttgagacgat caccagcggc gacctgttca 13678
 tcggtgagaa acggatgaat gacactccgc cagcagaacg cggcgttggt atggtgtttc 13738
 agtcttacgc gctctatccc cacctgtcag tagcagaaaa catgtcattt ggcttgaaac 13798
 tggctggcgc aaaaaagag gtgattaacc aacgcgttaa ccaggtggcg gaagtgtctac 13858
 aactggcgca ttgtctggat cgcaaaccga aagcgtctc cgggtggtcag cgtcagcgtg 13918
 tggcgattgg ccgtacgctg gtggccgagc caagcgtatt ttgtctcgat gaaccgctct 13978
 ccaacctcga tgctgcactg cgtgtgcaaa tgcgtatcga aatctcccgt ctgcataaac 14038
 gcctggggccg cacaatgatt tacgtcaccc acgatcaggt cgaagcgatg acgctggccg 14098
 acaaaatcgt ggtgctggac gccggtcgcg tggcgcaggt tgggaaaccg ctggagctgt 14158
 accactatcc ggcagaccgt ttgtctgccg gatttatcgg ttcgccaaag atgaacttcc 14218
 tgccggtaaa agtgaccgcc accgcaatcg atcaagtga ggtggagctg ccgatgccaa 14278
 atcgtcagca agtctggctg ccagttgaaa gccgtgatgt ccaggttgga gccaatatgt 14338
 cgctgggtat tcgcccggaa catctactgc cgagtatat cgctgacgtc atccttgagg 14398
 gtgaagttca ggctgtcgag caactcggca acgaaactca aatccatata cagatccctt 14458
 ccattcgtca aaacctgggtg taccgccaga acgacgtggg gttggtagaa gaaggtgcc 14518
 cattcgctat cggcctgccg ccagagcgtt gccatctgtt ccgtgaggat ggcactgcat 14578
 gtcgtcgact gcataaggag ccgggcgttt aagcaccoca caaacacac aaagcctgtc 14638
 acaggtgatg tgaaaaaaga aaagcaatga ctcagg 14674

<210> 37

<211> 549

<212> PRT

<213> Escherichia coli

<400> 37

Met	Lys	Asn	Ile	Asn	Pro	Thr	Gln	Thr	Ala	Ala	Trp	Gln	Ala	Leu	Gln
1				5					10					15	

Lys	His	Phe	Asp	Glu	Met	Lys	Asp	Val	Thr	Ile	Ala	Asp	Leu	Phe	Ala
			20					25					30		

Lys	Asp	Gly	Asp	Arg	Phe	Ser	Lys	Phe	Ser	Ala	Thr	Phe	Asp	Asp	Gln
		35					40					45			

Met	Leu	Val	Asp	Tyr	Ser	Lys	Asn	Arg	Ile	Thr	Glu	Glu	Thr	Leu	Ala

WO 01/70776

PCT/US01/07478

- 159 -

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

- 160 -

Asp Tyr Gln Thr Gly Pro Ile Ile Trp Gly Glu Pro Gly Thr Asn Gly
 370 375 380
 Gln His Ala Phe Tyr Gln Leu Ile His Gln Gly Thr Lys Met Val Pro
 385 390 395 400
 Cys Asp Phe Ile Ala Pro Ala Ile Thr His Asn Pro Leu Ser Asp His
 405 410 415
 His Gln Lys Leu Leu Ser Asn Phe Phe Ala Gln Thr Glu Ala Leu Ala
 420 425 430
 Phe Gly Lys Ser Arg Glu Val Val Glu Gln Glu Tyr Arg Asp Gln Gly
 435 440 445
 Lys Asp Pro Ala Thr Leu Asp Tyr Val Val Pro Phe Lys Val Phe Glu
 450 455 460
 Gly Asn Arg Pro Thr Asn Ser Ile Leu Leu Arg Glu Ile Thr Pro Phe
 465 470 475 480
 Ser Leu Gly Ala Leu Ile Ala Leu Tyr Glu His Lys Ile Phe Thr Gln
 485 490 495
 Gly Val Ile Leu Asn Ile Phe Thr Phe Asp Gln Trp Gly Val Glu Leu
 500 505 510
 Gly Lys Gln Leu Ala Asn Arg Ile Leu Pro Glu Leu Lys Asp Asp Lys
 515 520 525
 Glu Ile Ser Ser His Asp Ser Ser Thr Asn Gly Leu Ile Asn Arg Tyr
 530 535 540
 Lys Ala Trp Arg Gly
 545

<210> 38

<211> 11692

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (529)..(1632)

<400> 38

aaagcgttgc gcctttgttg tatcgtcagt tcagggtaaa atagatttcc gttaaccacc 60
 tggtcaggac gccgtatgca ttgccattc tgtttcgccg tggacactaa ggtaattgac 120
 tctcgtctcg tgggcgaggg ttcattccgta cgccgccgtc ggcagtgtct ggtgtgtaat 180
 gaacgtttca ccacctttga agtggcgagg ctggttatgc cgcgtgttgt aaaaagcaac 240
 gacgtgcgtg aaccgtttta tgaagagaaa ttgcgtagcg gaatgctgcg ggcgctggaa 300
 aaacgtccgg tgagttccga tgacgtcgaa atggcaatca atcatattaa atcgcagctg 360

cgcgccaccg gtgagcgcgga agtgccgagc aagatgattg gcaatctggt gatggagcaa															420
ttgaaaaagc tcgataaagt cgcctatatc cgtttttgct ctgtctaccg cagtttcgaa															480
gatatcaaag aatttggcga agagatcgcg cgcctggagg actaagcc															537
gtg cag gac Val Gln Asp 1															
gag tat tac atg gcg cgg gcg cta aag ctg gcg caa cga gga cgt ttt															585
Glu Tyr Tyr Met Ala Arg Ala Leu Lys Leu Ala Gln Arg Gly Arg Phe 5 10 15															
acc acg cat ccc aac ccg aat gtc ggg tgc gtc att gtc aaa gat ggc															633
Thr Thr His Pro Asn Pro Asn Val Gly Cys Val Ile Val Lys Asp Gly 20 25 30 35															
gaa att gtc ggt gaa ggt tac cac caa cgt gcg ggt gaa cca cat gcc															681
Glu Ile Val Gly Glu Gly Tyr His Gln Arg Ala Gly Glu Pro His Ala 40 45 50															
gaa gta cac gcg ttg cgt atg gcg ggt gaa aaa gcc aaa ggt gcg acc															729
Glu Val His Ala Leu Arg Met Ala Gly Glu Lys Ala Lys Gly Ala Thr 55 60 65															
gcc tat gtc aca ctc gaa ccc tgt agc cat cat ggt cgt acg cca ccg															777
Ala Tyr Val Thr Leu Glu Pro Cys Ser His His Gly Arg Thr Pro Pro 70 75 80															
tgc tgt gac gca ctc atc gcc gct ggc gta gcg cgc gtg gtt gcc tcg															825
Cys Cys Asp Ala Leu Ile Ala Ala Gly Val Ala Arg Val Val Ala Ser 85 90 95															
atg caa gat cct aac ccg cag gtc gct ggg cgt gga ctt tac cgt ctg															873
Met Gln Asp Pro Asn Pro Gln Val Ala Gly Arg Gly Leu Tyr Arg Leu 100 105 110 115															
caa cag gct ggc att gac gtc agc cac ggc ctg atg atg agt gaa gcc															921
Gln Gln Ala Gly Ile Asp Val Ser His Gly Leu Met Met Ser Glu Ala 120 125 130															
gag caa ttg aat aaa ggc ttt ctc aag cgg atg cgc acc ggc ttt cct															969
Glu Gln Leu Asn Lys Gly Phe Leu Lys Arg Met Arg Thr Gly Phe Pro 135 140 145															
tat att cag tta aaa ctt ggc gca tcg ctt gat ggt cgc acg gcg atg															1017
Tyr Ile Gln Leu Lys Leu Gly Ala Ser Leu Asp Gly Arg Thr Ala Met 150 155 160															
gcg agc ggc gaa agc cag tgg atc act tcg ccc cag gcg cgg cgt gat															1065
Ala Ser Gly Glu Ser Gln Trp Ile Thr Ser Pro Gln Ala Arg Arg Asp 165 170 175															
gta caa cta ctg cgc gcg caa agt cat gcc att tta acc agc agc gcc															1113
Val Gln Leu Leu Arg Ala Gln Ser His Ala Ile Leu Thr Ser Ser Ala 180 185 190 195															
acg gtg ctg gcg gat gat cct gcc tta acg gtg cgt tgg tct gaa ctg															1161
Thr Val Leu Ala Asp Asp Pro Ala Leu Thr Val Arg Trp Ser Glu Leu 200 205 210															

- 162 -

gat gaa caa act cag gcg ctc tat ccg caa caa aat ctc cgt cag ccg	1209
Asp Glu Gln Thr Gln Ala Leu Tyr Pro Gln Gln Asn Leu Arg Gln Pro	
215 220 225	
ata cgt att gtg att gat agc caa aat cgc gtg acg ccg gta cat cgc	1257
Ile Arg Ile Val Ile Asp Ser Gln Asn Arg Val Thr Pro Val His Arg	
230 235 240	
att gtg cag cag ccc gcc gaa acc tgg ttc gcg cgt acg cag gaa gat	1305
Ile Val Gln Gln Pro Gly Glu Thr Trp Phe Ala Arg Thr Gln Glu Asp	
245 250 255	
tct cgt gag tgg ccg gaa acg gtg cgt acc ttg ctg att cca gag cat	1353
Ser Arg Glu Trp Pro Glu Thr Val Arg Thr Leu Leu Ile Pro Glu His	
260 265 270 275	
aaa ggt cat ctg gat ctg gtt gta ctg atg atg caa ctg ggt aaa cag	1401
Lys Gly His Leu Asp Leu Val Val Leu Met Met Gln Leu Gly Lys Gln	
280 285 290	
caa att aac agc atc tgg gtg gaa gcg ggg cca acg ctc gct gcc gca	1449
Gln Ile Asn Ser Ile Trp Val Glu Ala Gly Pro Thr Leu Ala Gly Ala	
295 300 305	
ttg ctg cag gcg ggt tta gtc gat gag ctg att gtc tat atc gca cct	1497
Leu Leu Gln Ala Gly Leu Val Asp Glu Leu Ile Val Tyr Ile Ala Pro	
310 315 320	
aaa cta tta gcc agc gac gcc cgc gga tta tgc acg ctg cca ggg ctt	1545
Lys Leu Leu Gly Ser Asp Ala Arg Gly Leu Cys Thr Leu Pro Gly Leu	
325 330 335	
gag aaa tta gcc gac gcc ccc caa ttt aaa ttc aaa gag ata cgt cat	1593
Glu Lys Leu Ala Asp Ala Pro Gln Phe Lys Phe Lys Glu Ile Arg His	
340 345 350 355	
gta gcc ccg gat gtt tgc ctg cat tta gtg ggt gca tga tctcccggct	1642
Val Gly Pro Asp Val Cys Leu His Leu Val Gly Ala	
360 365	
cgaaagggaa gcagcgcacg aaatattatg ctaaaatccg cccccctgcg gggccatact	1702
cgaacccgaa ggaagaaaat gaacattatt gaagctaacg ttgctacccc ggacgctcgc	1762
gtcgccatca ccacgcgcg tttcaacaac tttatcaatg acagcctgct ggaagggtgca	1822
attgacgcac tgaaacgtat cggtcaggta aaagatgaaa acattaccgt tgtttgggtg	1882
cctgggtgcct atgagctgcc gctggcggcg ggtgcactgg ctaaaaccgg taaatacgac	1942
gcggtgattg cgctgggtac ggttattcgt ggtggcactg ccactttga atatgtcgcct	2002
ggtggtgcaa gcaacggcct ggcgcatgtt gcccaggaca gcgaaattcc ggttgctttt	2062
ggggttctga ccactgaaag cattgaacaa gcgatcgaac gtgctggcac caaagctggc	2122
aacaaagggtg cagaagctgc actgaccgcg cttgaaatga ttaatgtatt gaaagccatc	2182

- 163 -

aaggcctgaa attagtaagg ggaaatccgt gaaacctgct gctcgtcgcc gcgctcgtga 2242
gtgtgccgtc caggcgctct actcctggca gttgtcccag aacgacatcg ctgatgttga 2302
ataccagttc ctggctgaac aggatgtaaa agacgttgac gtcctgtact tccgtgagct 2362
gctggccggg gtggcgacta ataccgcata cctcgacgga ctgatgaagc catacctgtc 2422
ccgcctgctg gaagaactgg gacaggtaga aaaagcagta ctgcgcattg cgctgtacga 2482
actgtctaaa cgtagcgatg tgccatacaa agtggccatt aacgaagcga tcgaactggc 2542
gaaatcgttc ggcgcagaag acagccataa gttcgtcaac ggcgtaactcg ataaagcagc 2602
acctgtgatt cgccctaaca aaaagtgata tccaggccgg tagattcacg gaagaccgtt 2662
ccatgatcgc cggccttttc ttttttacct gctgaggcat aacgtatggc atgtggcgag 2722
ttctccctga ttgcccgtta ttttgaccgt gtaagaagtt ctcgctcttga tgtcgaaactg 2782
ggcatcggcg acgattgcgc acttctcaat atccccgaga aacagaccct ggcgatcagc 2842
actgatacgc tgggtggcggg taaccatttc ctccctgata tcgatcctgc tgatctggct 2902
tataaagcac tggcggtgaa cctaagcgat ctggcagcga tggggggccga tccggcctgg 2962
ctgacgctgg cattaacctt accggacgta gacgaagcgt ggcttgagtc cttcagcgac 3022
agtttgtttg atcttctcaa ttattacgat atgcaactca ttggcggcga taccacgcgt 3082
gggccattat caatgacgtt gggatatccac ggctttgttc cgatgggacg agccttaacg 3142
cgctctgggg cgaaaccggg tgactggatc tatgtgaccg gtacaccggg cgatagcgcc 3202
gccgggctgg cgattttgca aaaccgtttg caggttgccg atgctaaaga tgcggactac 3262
ttgatcaaac gtcactctcg tccatcgccg cgtattttac aggggcaggc actgcgcgat 3322
ctggcaaatt cagccatcga tctctctgac ggtttgattt ccgatctcgg gcatatcgtg 3382
aaagccagcg actgcggcgc acgtattgac ctggcattgc tgccgttttc tgatgcgctt 3442
tctcgccatg ttgaaccgga acaggcgctg cgctgggcgc tctctggcgg tgaagattac 3502
gagttgtgtt tcaactgtgc ggaactgaac cgtggcgcgc tggatgtggc tctcggacac 3562
ctgggcgtac cgtttacctg tatcgggcaa atgaccgccg atatcgaagg gctttgtttt 3622
attcgtgacg gcgaacctgt tacattagac tggaaaggat atgaccattt tgccacgcca 3682
taaagatgtc gcgaaaagtc gcctgaagat gagtaatccg tggcatctac ttgctgtcgg 3742
attcgggaagt ggattaagcc cgatcgttcc tgggacgatg ggctcgctgg cagcgattcc 3802
gttctggtat ctgatgacct ttttgccctg gcagctctac tcgctggtgg tgatgctggg 3862
gatctgtatc ggcgctctatc tttgtcatca aacggcgaaa gacatgggtg tgcacgatca 3922
tggcagcatt gtctgggacg aatttattgg tatgtggatc acgctcatgg cgctgccgac 3982

- 164 -

caatgactgg cagtgggttg ccgccgggtt tgtgattttc cgtattctgg atatgtggaa 4042
gccgtggccg atccgctggt ttgatcgcaa tgtgcatggc ggcatgggga tcatgatcga 4102
cgatattgtc gccggggtga tttccgcagg catcctgtat tttatcggtc atcactggcc 4162
gctgggtatt ctgtcgtagt tgtgcaccga tgcctgatgt gacgcttgtc acgtctcatc 4222
aggcctggac tcttatttaa atcctacgac aggatgcggt ttatacggcg tttccagttc 4282
ggcaatctgt tccggcttca aagtgatatc caccgcgttc aatagctcat caagctgttc 4342
ttcgcgcgaa gttccgataa tcggtgcggc aatgcccgggt ttactcaaca accaggccag 4402
cgcaacttgt gctcgtgtcg ccccagttc ttactgacg cctgttaacc gctctgcgat 4462
ctgcgcgtca ttttcatcgc tttctttata gagatttttc cccacctcat cagacaccag 4522
tcgtgcggta gtttctcccc acggacgcgt cagacggccc cttgccagcg ggctccatgg 4582
aattaccgcc acgccctcct gataacacag tggtagcatc tcgcgctctt cttcacgata 4642
aatcagattg tagtgatcct gcatactgac aaactgcgcc cagccgtgct gtttttggag 4702
ttccagtgcc tgagcaaact gcgaagcgtg cattgatgac gcgccgatat aacgcgcttt 4762
cccggctttt accacgtcgt tgagggcttc cagcgtctct tcgatcggcg tgttgtaatc 4822
ccagcgatga atttgcagga tatcgacata atccatgccg agacgtcgca ggctgtcgtc 4882
gatagagcgc aaaattttgcg cacgggataa tccttccgggt aaatcaccaa cgcgatggaa 4942
cactttggtc gcaacgacca cgtcttcacg acggggcgaat tcccgcagtg cgcgaccgac 5002
gatctcttcg ctgctgccgt cagaataact gttggcggta tcaaagaaat ttatgccgcc 5062
ttccagtgca cgtttaatta tgggacggct gctttcttcc ggcagtgtcc atgcgtgatt 5122
accgcgatct ggctcgccaa aggtcataca gccgaggcaa agtcgggaaa cgcgagggtc 5182
ggtttttctt aaggggttgt attgcatgct gccactcctg ctatactcgt catacttcaa 5242
gttgcatgtg ctgcggctgc attcgttcac ccagtcact tacttatgta agctcctggg 5302
gcttcactcg ttgcccgcct tcctgcaact cgaattatth agagtctatg aataatttct 5362
taagcatagc aggagtggag tagggattat gccagccagg ccttgatttt ggcttccata 5422
ccagcggcat cgaggccgag ttcggcgcgc atttcttctt gagttccttg cggaataaag 5482
aagtccggca ggccaatgtt cagcacgggt actggtttac gatgggcat cagcacttcg 5542
ttcacgccgc tgccctgcgc gcccataatg gcgttttctt ctacggtgac cagcgcttca 5602
tggtggcg ccatthccag aattaacgct tcatcaagcg gtttcacaaa acgcatatcg 5662
accagcgtgg cgttcagcga ttcggcgact ttcgcgcgtt ctggcatcag cgtaccaaag 5722

- 165 -

ttaaggatcg ccagtttctc gccacgacgc ttcacaatgc ctttgccaat tggtagtttt 5782
tccagcggcg tcagttccac gccgaccgcg ttgccacgcg ggtagcgcac cgctgacggg 5842
ccatcgttat agtgatagcc ggtatagagc atctggcgac attcgttttc atcgctcggg 5902
gtcataatga ccatttccgg tatgcagcgc aggtaagaga gatcaaaagc accctgatgg 5962
gtttgaccgt cagcaccaac aatgcccgcg cggtcgatgg cgaacaggac cggaagcttt 6022
tgaatcgcca cgtcatgcag cacctgatca taggcgcgtt gcaggaaagt ggagtaaate 6082
gcgacaatgg gtttgtaccc accaatcgcc agacccgcag caaaggtcac cgcgtgttgc 6142
tcggcaattg ccacgtcgaa gtagcgatcc gggaatttac gtgaaaactc gaccatgccg 6202
gaaccttcac gcatcgccgg agtaatcgcc atcagcttgt tgtctttcgc tgccgtttcg 6262
cacaaccagt cgccaaagat ttttgaatag ctcggaacac cgccgctact tttcggcaaa 6322
caaccgctgg agggatcaaa tttaggcacg gcgtggaaag tgatcgggtc tttttctgcc 6382
ggttcataac cacgaccttt tttggtcatg atatgcagga actgcgggcc tttcaggctc 6442
cgcatgttct ttagcgtggg gataagcccc agcacatcgt gaccgtccac cgggccgatg 6502
tagttaaagc ccagctcttc aaacaacgtg ccaggcacta ccatgccttt aatatgttct 6562
tcgggtgcgtt tgagcagctc ttttaattggc ggcacgccag agaaaacttt tttcccgcct 6622
tcgcgcagtg aagagtaaag cttaccggaa agcagctgtg ccagatgggt gttgagcgcg 6682
ccgacatttt cggaaatcga catttcattg tcgttgagaa tcaccagcat atcaggacgg 6742
atatcgcccc cgtgattcat cgcttcaaac gccatgcctg cggtaatcgc gccatcgcca 6802
atgacacaga cggtgccggc atttttgcct tctttttcgg cagcaaccgc aataccaatt 6862
ccggcactga tggagggtga tgaatgcccg acgcttaata cgtcatattc gctttcgccg 6922
cgccacggga acgggtgcag accgcctttc tgacggatgg tgccgatttt gtcgcggcgt 6982
ccggtcaaaa ttttatgcgg ataagcctga tgccccacat cccaaatcaa ttgggtcaaac 7042
ggggtgttgt agacatagtg cagcgccacg gtcagttcga ccgtgcccag cccggaggcg 7102
aagtgccgcg tggaacggct cacgctgtcg agtaaatagc ggcgcagttc gtcgcagagt 7162
ttcggtaaac tctctttcgg caacagtcgt aactcctggg tggagtcgac cagtgccagg 7222
gtcgggtatt tggcaatata aaaactcatc aggggcctat taatacttat tgtttattta 7282
ttacgctgga tgatgtagtc cgctagcgct tccagtgccg aggtatcgag tgactgttca 7342
gccagttggt tcagcgactg acgggcacgc tcgatcagat cccgggcttt cttccgggct 7402
tgctcaagac ccagaagtgc agggtaggta cttttaccaa gttgctggtc ggcaccctgg 7462
cgttttccca acgttgcagt atctcccacc acatccagga tgatcctcgt aacctggaag 7522

- 166 -

gcaaggccga tgctctctgc atacttgctg agtaccggca gagcacgacg tcctttatct 7582
ccggcgctta atgcaccaag gcgaacggcg gcgcgaatca atgcgccggt tttatgacga 7642
tgaatacgct caagcgcgtc cagaggtaag tgtttgctt ccgcgtctaa atctaagcc 7702
tgaccaccgc acattccggc aataccactg gcgctcgcca gttcagaaat catcgaaatt 7762
ctgtcgcggt ccgacacttc cggcatatcg gcacgcctta aaatcgagaa cgccagcggt 7822
tgtaaagcgt cgccagcgag aatcgcgttt gcttcgcca acttcacatg gcagggtggc 7882
aaaccgcgac gcagatcgtc atcatccatt gccggtaaatt catcatgaat taatgagtaa 7942
gcgtggatac actcaacggc ggcagcgggt gcgtccagcg tgtttgctgct aacgccgaac 8002
atatgaccgg tggcataaac caggaaagggt cgcaggcgct taccaccta taatgcgcca 8062
tactgcatgg ttctgaccac gggagtgttc tgaaagggca gtggggcgat aaaacggctc 8122
agcgctgggt tggcctgctt aacgcaggct tcgagttgct gcggaaagtc cattactcat 8182
tgtccggtgt aaaaggggtt agagaggcgt ctcatgttc agacagcaga atttgtacgc 8242
gctgttcggc ttgttgtaat ttggcctgcc cctgacgtgc cagctgcacg ccgcgttcga 8302
actcgttcag cgctcttcc agcggcagggt cgccacttcc cagacgggtt acaatctgtt 8362
ccagctcgct cagcgctttt tcaaagctgg cgggcgcctc atttttcttc ggcataatga 8422
atgtctgact ctcaatatct ttgcggcgt catggtaacg gactcagggc aaatagcaaa 8482
taacgcgcaa tggtaagggt atgtgcacag caaagcgatg ttagtggtat acttccgcgc 8542
ctggatgcag ccgcagggtg gggctgctgt atttttccct atacaagtcg ctttaaggctt 8602
gccaacgaac cattgccgcc atgaagttta tcattaaatt gttcccggaa atcaccatca 8662
aaagccaatc tgtgcgcttg cgctttataa aaatccttac cgggaacatt cgtaacgttt 8722
taaagcacta tgatgagacg ctgcgtgtcg tccgccactg ggataacatc gaagttcgcg 8782
caaaagatga aaaccagcgt ctggctattc gcgacgctct gaccogtatt ccgggtatcc 8842
accatattct cgaagtcgaa gacgtgccgt ttaccgacat gcacgatatt ttcgagaaag 8902
cgttgggtca gtatcgcgat cagctggaag gcaaaacctt ctgcgtacgc gtgaagcgcc 8962
gtggcaaaca tgatttttagc tcgattgatg tggaacgtta cgtcggcggc ggttttaaatc 9022
agcatattga atcgcgcgcg gtgaagctga ccaatccgga tgtgactgtc catctggaag 9082
tggaagacga tcgtctcctg ctgattaaag gcgctacga aggtattggc ggtttcccga 9142
tcggcaccca ggaagatgtg ctgtcgctca tttccgggtg ttcgactcc ggtgtttcca 9202
gttatatgtt gatgcgtcgc ggctgccgcg tgcattactg cttctttaac ctcggcggcg 9262

- 167 -

cggcgcacatga aattggcgtg cgtcaggtgg cgcattatct gtggaaccgc tttggcagct 9322
cccaccgcgt gcgttttgtc gctattaatt tcgaaccggt cgtcggggaa attctcgaga 9382
aaatcgacga cggtcagatg ggcgttatcc tcaaacgtat gatgggtgcgt gccgcaccta 9442
aagtggctga acgttacggc gtacaggcgc tggtcaccgg cgaagcgctc ggccaggtgt 9502
ccagccagac gctgaccaac ctgcgcctga ttgataacgt ctccgacacg ctgatcctgc 9562
gtccgctgat ctcttacgac aaagagcaca tcatcaacct ggcccggcag attggcaccg 9622
aagactttgc tcgcacgatg ccggaatatt gtgggtgtgat ctccaaaagc ccgacggtga 9682
aagcagttaa atcgaagatt gaagcgggaag aagagaagtt cgacttcagc attctcgata 9742
aagtggttga ggaagcgaat aacgttgata tccgcgaaat cggccagcag accgagcagg 9802
aagtgggtga agtggaacc gtcaatggct tcggcccgaa cgacgtgatc ctcgatatcc 9862
gttctatcga tgaacaggaa gataagccac tgaaagtcga agggattgat gtggtttctc 9922
tgccgttcta taaactgagc accaaatttg gcgatctcga ccagaacaaa acctggctgc 9982
tgtgggtgtga gcgcggggtg atgagccgtc tgcaggcgct ctatctgcgc gagcagggt 10042
ttaacaatgt gaaggtatat cggccgtaat ttgtggtttt tacgtcgcat ctggtcagat 10102
gcgacgtttg ccgcatccga cactactcgt aataattata aatccctgcc gccatcacca 10162
gttgtgatgc cacttcatgg gctttttcac gcccaaccaa caggtcgata attttcagac 10222
caaagtcgat agctgtacc ccgcccctggc tggtcagcaa ttttaccgt gcacccaga 10282
cgacgcgctt gtccagccat tgttcggcgg gaattttgtc tttcagcgct gggaagccgg 10342
tcatattacc aatcgggaag atatcgtgcg gcaccagcac ggtggctggc gcggcgcaaa 10402
tagccgcgac gatacggccg gaacgggtga actgtttaac ggtttcaacc agcagagtgc 10462
tatcgcgaaa acactccgcg cctttaatgc caccaggcag cacgatcacg tcatattcgc 10522
catcagccac ttcgaccagc ggcgcatccg ccagcagctt cagccgcgc gagcaggtaa 10582
tcgccaggtt accatcgctg gcgacgctgg cagtgggtgac tttgataccg ccgcgaacca 10642
gcagatcgat agtgggtgac gcttcagtct cttcactacc aggggcgagg caaaccagtgc 10702
ccgatgcgct catattcact ctcttttctt tttaccattt caaacaggcg ggtgttttcc 10762
ggtacggcaa tcccatgcgc gcgggcgcg cgtaagagaa aaccattgat atagtcgatt 10822
tcagtgtggc gcagcgcgcg gatatacctgc aacatcgac agatattttc cgctgtggca 10882
tcaatcacct gcatcacgta atcacgcaaa tcttctgctg aagtatgatg cccttcgcgt 10942
tcgatcaccg ccgcgacttc ttgcgatatc tgcataattt cttgcggatg atgacgtaat 11002
tcaccgttcg ggcaattcca gatggcagtc agtggattaa tcacgcagtt gactgccagc 11062

- 168 -

ttgcgccaca gctcggcgcg aatattgtta tgccaggcaa cgtcaggcaa cacggtttgc 11122
 aaaatatccg ccagataact gtaatccccg tcctgttgcc gtgccggggc aatatgcgtg 11182
 ataccgtttg ccacatgaat aatgacattg ccgtcgcggc gggctgcatg ggtggtggtg 11242
 cccatcagta atggctgctg aatgttttgc aactcttcga tggtgcccat gccgttgtga 11302
 attaacagta ttggcgtagt tacaggcagt gtggacgcga ggcttttgac ggcatacgaa 11362
 acctgccatg ctttcagcgt caccaggagc agatcgctgg tggcgagaaa atcgggatcg 11422
 ttggcgggtca gcgattcgtt aaatatcgaa ccatctgtct caaccagatt cacgctacaa 11482
 taaggttgcy gtacgcgcag ccagccctga acttcatgac cctgtttgca aagtgcgtga 11542
 agccataatt gccctaaggc accgcatccc aatacggtaa ttttcattgt tcctcctcac 11602
 ccgcaaccac tccgggtgtt caataaggct atcccttaat tgtgcatgct gttgcgacta 11662
 tgcacaatta agggatacgt cctgggtgcag 11692

<210> 39
 <211> 367
 <212> PRT
 <213> Escherichia coli

<400> 39
 Val Gln Asp Glu Tyr Tyr Met Ala Arg Ala Leu Lys Leu Ala Gln Arg
 1 5 10 15
 Gly Arg Phe Thr Thr His Pro Asn Pro Asn Val Gly Cys Val Ile Val
 20 25 30
 Lys Asp Gly Glu Ile Val Gly Glu Gly Tyr His Gln Arg Ala Gly Glu
 35 40 45
 Pro His Ala Glu Val His Ala Leu Arg Met Ala Gly Glu Lys Ala Lys
 50 55 60
 Gly Ala Thr Ala Tyr Val Thr Leu Glu Pro Cys Ser His His Gly Arg
 65 70 75 80
 Thr Pro Pro Cys Cys Asp Ala Leu Ile Ala Ala Gly Val Ala Arg Val
 85 90 95
 Val Ala Ser Met Gln Asp Pro Asn Pro Gln Val Ala Gly Arg Gly Leu
 100 105 110
 Tyr Arg Leu Gln Gln Ala Gly Ile Asp Val Ser His Gly Leu Met Met
 115 120 125
 Ser Glu Ala Glu Gln Leu Asn Lys Gly Phe Leu Lys Arg Met Arg Thr
 130 135 140
 Gly Phe Pro Tyr Ile Gln Leu Lys Leu Gly Ala Ser Leu Asp Gly Arg
 145 150 155 160

- 169 -

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

<210> 40

<211> 10944

<212> DNA

<213> Escherichia coli

<400> 40

tgcactccta ctattttaata tgtacgttcc atgctgaaaa gcccgttttc aggataactca 60
 aatggaaacg cacagacata caaaagattg gctggctaata ctagccagct caacccaact 120
 ttgcaagaaa aatatgcgaa aaaatcagcc ttggcgctgt ttatgcttcg gctcggcact 180
 gcaaatacaca cggatgacac catcacgctt aacgattttg cagttacggc ataattttctt 240
 gacggaagca cgaactttca tttttactct ccgtaacttc tcgggcgacc aattatcggc 300
 cgtagccttt cagggttcgcc ttcttcaatg cagactcata ctgactggac atcatcagag 360

- 170 -

tttgcacttg agccataaag tccataatca cgacaacaac gataagcagt gaggtccac 420
 cgaagtagaa cgggtactttc attgcatcac gcatgaactc cgggatcagg cagataaagg 480
 taatatacag cgcaccaacc agggtcaggc gggtcattac tttatcgata tacttcgccg 540
 tttgctctcc cggacgaatt cctgggtacaa atgcaccgga cttcttcagg ttatctgctg 600
 tttcacgcgg gttgaaaacc aacgccgtgt agaagaaaca gaagaagatg attgcagacg 660
 catagagtaa cacataaagc gggtgcccag gctgcaaata cagcgaaatt gttgtcagcc 720
 agttccaacc agtaccgccc ccgaaccatg acgcgatggc cgccgggaac agaataatac 780
 tggaagcgaa gattgccggg attacccccg ccatattcac tttcagcggc aaatgtgtgc 840
 tctgtgcagc atagacacga cgaccttgct gacgtttcgc gtagtttacc acaatgcggc 900
 gttgaccacg ctcaacaaat acaacaaaga acgtcactgc aaataactaat actgcaacca 960
 acagcaacac gaggaagtgc aggtcgcctt gacgcgcttg ctcgatagta tgggcaatgg 1020
 ctggcgggag tcccgcgaca ataccggcgga agataatgat tgaaataaccg ttgccgatac 1080
 ctcgttcagt aatctgttcg cccaaccaca tcaggaacat ggttcctgtg accagactta 1140
 caacagcggc gaagtagaat gcaaagcccc ggttaatcac caggccttgc ataccaggca 1200
 tatteggcag accggtagca ataccgatcg actggaatat tgccagcacc agagtaccgt 1260
 agcgggtgta ctggctgac ttacgacgac cagactcccc ttctttctta atttctgcca 1320
 acgttgggtg aaccaccgtc agcagctgga taatgatcga cgccgaaata tacggcatga 1380
 tccccagagc aaagatagaa gcacggctga gagcaccacc agagaacatg ttaaacaatc 1440
 caatgatggc gcctcgctgt tgctcaagca gtttggcaag tacagcggca tcaataccag 1500
 ggatcggaat aaaagagcca atacggaaca caatcagcgc accgataaca aacagcagtc 1560
 tgcgtttcag ctgcctaag ccacctttgg cactttgaaa atctaatacc gggtgtttag 1620
 ccatctgcta cttattcctc gatcttaccg ccagcagctt cgatagcagc acgagcgcct 1680
 ttagtaacac gcaggccacg aacagttacc ggagtcgtta cttcgccagc caggatcact 1740
 ttcgcgaact cgatctggat accgataatg ttagccgctt tcagcgtggt caggcttact 1800
 acaccgcctt ctactttagc caggtcagac agacgaatct cggctgtaat cgctgcttta 1860
 cgagaagtga agccgaatct cggcagacga cggtagagag gcatctgacc accctcgaaa 1920
 ccgcgacgta cgccaccgcc agaacgagac ttctgacctt tgtgaccacg accaccgggt 1980
 ttaccgaggg cagaaccgat accacgaccc aggcgtttac ccgccttttt ggagccttcg 2040
 gccggagaca gagtatttaa acgcatctct tactcctcaa ctttaaccat gaaggaaacc 2100
 gcgttgatca taccgcgaat agcaggagta tcctcgcgct ctacgggtgtg accaatacga 2160

- 171 -

cgagaccca ggccaagcag cggtgccttg tgtttcggca gacgaccgat tgcactgcgg 2220
gtttgagtaa ttttaatagt ctttgccatg gtttatattcc ccagaatttc ttcaacggat 2280
ttaccacgct tggcagcgac catttctgga gaattcatat tttccaggcc atcaatagtt 2340
gcacgaacca cggtgatcgg gttggtggaa ccataggctt tagccagaac gttatgaacc 2400
ccagcgactt ccagaacggc gcgcattgca ccaccggcga tgataccggt accttcggaa 2460
gccggctgca tgaatacgcg agaaccctg tgaacacctt taacagggtg ttgcagagtg 2520
ccgttattca gcgcgacggt aatcatattg cgacgggctt tttccatcgc tttctggatc 2580
gctgctggaa cttcacgcgc tttaccgtaa ccaaaaccaa cgcgaccggt accatcgcca 2640
actacagtca gagctgtgaa ggagaaaata cgaccacctt taacggtttt agatacgcgg 2700
tttaccgcga tcagcttttc ctgcagttcg ccagcttggt tttcgatgtg agccatctta 2760
cacctctacc ttagaactga aggccagctt cacgggcagc atctgccagt gcctggacac 2820
gaccatgata ttggaaccgc gaacgggtcaa aggatacatc tttgatgcct ttttccagag 2880
cgcgttcagc gacagcttta cccacagctg cagccgcgctc tttgttaccg gtgtacttca 2940
gttgttcagc gatagctttt tctacagtag aagcagctac cagaacttca gaaccgttcg 3000
gtgcaattac ctgtgcgtaa atgtgacgcg gggtagcatg taccaccagg cgagttgcgc 3060
ccagctcctg gagcttgccg cggtgcgcggg tcgcacgacg gatacgagca gatttcttat 3120
ccatagtgtt accttacttc ttcttagcct ctttggtacg cacgacttcg tcggcgtaac 3180
gaacaccctt gcctttataa ggctcaggac gacggtaggc gcgcagatcc gctgcaacct 3240
ggccgatcac ctgcttatca gcgcctttca gcacgatttc agtctgagtc ggacattcag 3300
cagtgatacc cgcaggcagc tgatggtcaa caggatgaga gaaaccaga gacaggttaa 3360
tcacattgcc ttttaaccgct gcacggtaac ctacaccaac cagctgcagc ttcttagtga 3420
agccttcggt aacaccgata accattgagt tcagcagggc acgcgcggta ccagcctgtg 3480
cccaaccgctc tgcgtaacca tcacgcggac cgaaggtcag ggtattatct gcatgtttta 3540
cttcaacagc atcgttgaga gtacgagtca gctcgccgtt tttacctttg atcgtaataa 3600
cctgaccggt gattttttacg tcaacgccgg caggaacaac gaccggtgct ttagcaacac 3660
gagacathtt ttcctccgat taggctacgt agcagataat ttgccacca agaccagcct 3720
ggcgcgctgc acgatcagtc ataacacctt tagaggtaga aacaactgcg ataccagac 3780
ccgccataac tttcggcagc tcattctttac gtttatagat gcgcagacct gggcggtga 3840
cacgctgaat gctttctaca acagctttgc cctggaaata cttcagagta agttccagtt 3900

- 172 -

caggcttggg gtcgccttca actttaaaat cttcaataaa accttcttcc ttcagcacgt 3960
tggcgattgc cactttcagc ttggaggaag gcatgggtgac cgcagctttg ttcgcggcct 4020
gaccgttacg gatacgggtc agcatatccg cgatcggatc ttgcatgctc atctgtcttt 4080
actcccgtga ttcaattggg gacaattacc agctagcctt tttcagaccc gggattttcac 4140
cgcgcatagc ggcttcacgg accttaatac ggctcaaccc gaacttccgc aggaaaccat 4200
gcggacgacc tgtttgacgg cagcgggttac gctgacgaga cgggctggaa tcacgcggca 4260
gagtctgcag cttgagaaca gcgttccaac gatcttcgtc ggaagcgttc acatcagaga 4320
tgatcgcttt cagttcagcg cgtttcgcga agtatattatc agctaaagct acgcgtttta 4380
cttcgcgtgc tttcattgat tgcttagcca tttagtaacc ctaccttact tgcggaacgg 4440
gaagtcaaag gcagccagca gagcgcggcc ttcttcgtca gatttcgcag tagtggtaat 4500
ggtaatatcc aaaccacgaa cgcggtcgac tttatcgtag tcgatttctg ggaagatgat 4560
ctgctcacgg acacccatgc tgtagttacc acgaccgtcg aaagacttag cggacaggcc 4620
acggaagtca cggatacgag gtacagcaat agtgatcagg cgctcaaaga actcccacat 4680
gcgttcgcca cgcagagtta ctttacagcc gatcggatag ccctgacgga ttttgaagcc 4740
tgcaacagat ttgcgtgctt tggatgatcag cggtttttga ccggagattg ctgccaggtc 4800
tgctgctgcg ttatccagca gttttttgtc agcgatcgct tcaccaacac ccatgttcag 4860
ggtgatcttc tcgacccgag ggacttgcac gacagaattg tagttaaact cagtcatgag 4920
ttttttaact acttcgtctt tgtagtaatc atgcagtttc gccatcgtac tactccaaat 4980
tacttgatag tttcgctggt agacttgaag aaacggactt ttttacgctc ttcgaatcta 5040
aagcctacac ggtcagcctt gccgggtgcc gcattgaaga ttgctacgtt ggaaacctga 5100
atagcggctt ctttttcaac gatgccaccc ggttggttca gggccggaac cggcttctga 5160
tgtttcttaa ccaggttgat accttcaaca atgaccttgc cggaagacag gacattctta 5220
actttaccgc gtttaccttt atctttaccg gttaacacga taacttcgtc atcacgacgg 5280
attttcgctg ccatgattcg ctcttagag tacttctggg gccagagaga taattttcat 5340
gaacttctca ctacgaagct cacgagttac cggcccaaaa atacgcgtac cgataggctg 5400
ctcgctgttg ttgttcagaa gaacacaagc attaccatcg aagcgaatga cagaaccgtc 5460
cgggcgacga acacccttct tgggtgcgcac cactaccgcc ttcagcacat cacctttttt 5520
gaccttacca cgcggaattg cttctttgat ggtgatcttg atgatgtcgc ctacgcctgc 5580
gtagcgacgg tgcgagccac ccagaacctt gatacacatt acgcgacgtg caccggagtt 5640
gtcggcgacg ttcagcatag tctgttcttg gatcatttta gtgctccgct aatgtcaact 5700

- 173 -

actactgaga cccgaaaatc aggtcgttaa aaatccccat atcgagggcg cggcattata 5760
acaccgcttc aaggatatgg gtagaaaaaa taaacggctc atttctgagc cgtttattcg 5820
tattgagaga gtgtactgta ttacagaacc gctttctcta caacgcgaac cagcgtccag 5880
gatttagtct tggacagcgg acggcattcg cggatttcaa ccacgtcacc gataccgcat 5940
tcgttggtct cgtcatgtac gtgcagtttg gtcgtacgct tgatgaattt accgtagatc 6000
gggtgtttca caaacgctc gatagcaaca acaatggatt tctccatttt gtcgctaaca 6060
acgcgacctt gcagagtacg gattttatcg gtcattacgc accgccttc tcgttcagta 6120
aagtcttaac gcgtgcgaca tcgcgacgca cttgcttcaa caggtgagac tgttgcagct 6180
ggccacttgc agcctgcata cgcaggttga actgctcacg cagcaggttc agcagctcgg 6240
tgttcagctc ttcaacgctc ttctcacgca gctcttttgc tttcattaca tcaccgtctt 6300
agttacaaag gtggttttaa tcggcagttt cgctgctgcc agcttgaatg cttcacgggc 6360
cagctcttcc ggaacaccgt ccatttcata caggacttta ccgggtgaa tcaaggcaac 6420
ccaatactcc acgttacctt tacctttacc catacgcact gccagcggct tttcagtgat 6480
cggtttgtcc gggaaacacac ggatccagat cttaccttga cgcttaactg cacgggtcat 6540
agcacgacgt gctgcttcga tctgacgggc agtcagacga ccacggccaa cagctttcag 6600
accgaagctg ccgaagctaa catccgtacc ctgcgccaga ccgcgggttac ggcctttgtg 6660
cattttacgg aattttgtac gctttggttg taacatcagc gacgctcctt atttacggcc 6720
tttacgctgc tgcttttttag gctgagcagc cggtttttcc ggttggttcaa cagcagccat 6780
accaccaggg atctcgcctt tgaagatcca cactttaacg ccgattacac cgtaagtggc 6840
gtgcgcttca gaggtggtgt agtcgatgtc agcacgcaga gtgtgcagcg gtacgcgacc 6900
ttcgcggtag cattcggtac gtgcgatttc cgcgccgcc agacggccgc taacttcaac 6960
tttaatacct ttagcgccca gacgcattgc gttctgtaca gcacgcttca tagcacgacg 7020
gaacataacg cgacgttcca gctgagaagt gatgctgtca gcaaccagtt ttgcgtccag 7080
ttcaggctta cgaacttcgg cgatgttgat ctgtgcagga acgccagcga tgtccgctac 7140
gaccttacgc agtttttcta cgtcttcacc ttttttaccg ataacgatac ccgggcgagc 7200
agtgtgaatg gttacacgga tgctcttagc cggacgctcg ataacgatac gagatacggg 7260
cgcttttagc agttccttag tcaggtactg acgtacttta aaatcgctgt ccaggttgtc 7320
agcgaattct ttggtgttcg caaaccaggt agagttccat ggttttacia taccagggcg 7380
aataccatta ggatgtactt tctgacccat tgctagtctc cagagtctca gcgatcggac 7440

- 174 -

acaaccacag tgatgtggct ggtgcgcttc aggatgcgat ctgcacgacc ttttgcacgc 7500
 ggcataatgc gcttcatgct cgggccttcg tctacgaaaa ttttcgtaac tttcagatcg 7560
 tcaatgtcag cgccatcggt gtgttcagcg ttagcaatgg cagattccag aactttcttg 7620
 accagtacag ccgctttctt gttggtgtag gtcaaaatat ccagagcctg cgacactttc 7680
 ttaccgcgaa tcaggtcagc aacaaggcga accttctgag cagaagaacg agcatggcga 7740
 tgttttagcga tagtttccat ctcttcctcc taccttattt cttcttcgct tttttatcag 7800
 cagcgtggcc gcgataagta cgagtcggtg cgaattcacc cagtttgtga ccaaccattt 7860
 cgtcggttac aaataccgga acgtgctgac gaccattatg gacagcgatg gtcaaaccga 7920
 tcatgttagg aaagatcggt gaacgacggg accaagtgcg caggggcttc ttgtctccgc 7980
 tttccaccgc tttctctacc ttcttcagca agtgcaggtc aataaaagga cctttcttga 8040
 gagaacgtgg catggcttat cctctaaaat tatgtgctac ggcgacgtac gatgaattta 8100
 tcagtacgct tgttgctgcg ggtcttctta cctttggtct gaacgccccca cggagttacc 8160
 ggggtgcttac caaagttacg accttcacca ccaccatgtg ggtggtctac cgggttcac 8220
 gcggtaccgc gaacggtcgg acgaacacca cgccagcgtg cagcacctgc tttaccaga 8280
 acgcgcagca tatgctcagc attgccaact tcgcccagag ttgcacggca gtctgcttct 8340
 actttacgca tttcaccaga acgcagacgc agggtgacat aagcaccatc acgagcaacg 8400
 atctgaacgt aagtaccagc ggaacgtgcc agctgaccgc ctttacctgg tttcatttct 8460
 acgttatgaa cagtagaacc aaccgggatg ttgcgcacgc gcagggtggt acctgggttg 8520
 attgcagcat caacgccaga ctgaatctgg tcgccagctt tcaggccttt aggggccagg 8580
 atgtaacggc gttcaccgtc tttgtacaga accagcgcga tgttcgcgga acggttcgga 8640
 tcgtactcaa gacgttcaac aactgccggg ataccgtctt tgttgcgttt gaagtcaaca 8700
 atacggtaag cctgcttggt gccaccaccg atatgacgag tggtgatacg gccattgttg 8760
 ttacgaccac cggatttgct gtttttttcc agcaacggag caaaagggtt gcccttgtgc 8820
 agctcagggt taaccacttt aactacgtgg cgacgaccgc gagatgtcgg tttacattta 8880
 acaactgcca ttgtattact cctccgactt actcagcgcc gccaacgaag tccagattct 8940
 ggccttcttt cagggtgacg taagcttttt tccagtcgct acgacgaccg atacgctgtc 9000
 cgtgacgttt aactttccct ttaactacca ggggtgtaac gacttcgact tcgacttcaa 9060
 acagtttctg cacagcagct ttgatttctg ctttggtcgc gtcttttagca actttgagta 9120
 cgatgggtgtt ggatttttcc atcgcagtag acgctttttc agaaacgtgc ggtgcacgca 9180
 gcacctcag cagacgttct tcacgaatca tgccagcatc tcctcaactt gcttaacagc 9240

- 175 -

atcagcagtc attacgactt tgtcgaaggc gatcaggcta accgggtcga taccagttgc 9300
 atcgcgtacg tcaaccttgt gcagggttgcg cgcagccagg aacagggtttt cgtccagctc 9360
 accggtgatg atcagcacat cttccagagc catgtctttc agtttctgtg ccagcagctt 9420
 agttttcggc gcttctacag agaacttctc gacaacgata agacgatacct gacgtaccag 9480
 ttcggacagg atgcttttca gcgcgccgcg gtacatcttc ttgttaactt tttgactgtg 9540
 gtcctgcgga cgagcagcaa aggtcacgcc accagaacgc cagatcgggc tcttgataga 9600
 accagaacgc gcacggccgg tgcctttctg gcgccacggg tttttaccgg aaccagttac 9660
 ttcagcacga gtcttctgag cacgagtacc ctgacgagca ccagctgcat aagcaacaac 9720
 aacctggtga accagcgctt cggtgaaatc acgaccgaag gtagtttcgg aacagtcag 9780
 cgcgctctgc gcgtctttca atactaattc cattgctatc tccttacgcc ttcacagctg 9840
 gtttaacgat caggtcgcta ccggttgac ccgggacagc acctttaacc agcagcagg 9900
 tgcgctcagc gtcaacgcgt actacgtcaa ggctctgaac gggttacacgt tcgttacc 9960
 tctgacctgc cattttcttg cctttgaaca ctttgcccgg agtctgggtc tgaccgatag 10020
 aaccggaac gcggtgagac aaggagttac cgtgagtagc gtcctgggta cggaagtcc 10080
 agcgcttaac ggtacctgcg aaacctttac ctttagagggt gccagttacg tcaacttttt 10140
 taacgtcagc aacagttca acgctaattg tctgacctac agtgaactct tcgccttcag 10200
 ccaggcgga tttccacaga ccacggccag cttctacgcc agcttttagcg aagtggccag 10260
 cttcaggctt ggtcacacgg ttagcttttt tagcaccggg ggtcacctga atagcacgg 10320
 agccatcggt agccagggtt ttaacctgag taacgcgggt tgcttcaact tcgattacgg 10380
 ttactgggat agaaacgccg tcttctgtga agatacgggt catacccact tttttaccga 10440
 ctaaaccaat cattgtttca acctctcaat cgctcaatga cctgattaac ccaggctgat 10500
 ctgcacgtct acaccggcag ccagatccag acgcatcaga gcatcaacgg ttttctcgg 10560
 tggctcaacg atgtcaacca gacgcaagtg agtacggatt tcgtactgat cgcgcgcgtc 10620
 tttgttgacg tgcggggaga tcagaacagt gaagcgctct ttgcgtgtcg gcagcgggat 10680
 cggaccacgg acctgcgcac cagtgcgctt ggcagtctcg acgatttccg cggttgcttg 10740
 atcgatcaga cgatgatcaa acgctttcag gcggatacgg attctttggg tctgcatgag 10800
 accagagctc caattatttt ataaacgaaa atgattactc ctacagacca ttacgattga 10860
 tgggagagtg taaccgttct tacgtagctc cccgattggg agcattgtta ggtagccaaa 10920
 ttcggctaac tgaggttcag attg 10944

WO 01/70776

PCT/US01/07478

- 176 -

<210> 41
 <211> 179
 <212> PRT
 <213> Escherichia coli

<220>
 <223> complement of position 4426-4965 of seq id 40

<400> 41
 Met Ala Lys Leu His Asp Tyr Tyr Lys Asp Glu Val Val Lys Lys Leu
 1 5 10 15
 Met Thr Glu Phe Asn Tyr Asn Ser Val Met Gln Val Pro Arg Val Glu
 20 25 30
 Lys Ile Thr Leu Asn Met Gly Val Gly Glu Ala Ile Ala Asp Lys Lys
 35 40 45
 Leu Leu Asp Asn Ala Ala Ala Asp Leu Ala Ala Ile Ser Gly Gln Lys
 50 55 60
 Pro Leu Ile Thr Lys Ala Arg Lys Ser Val Ala Gly Phe Lys Ile Arg
 65 70 75 80
 Gln Gly Tyr Pro Ile Gly Cys Lys Val Thr Leu Arg Gly Glu Arg Met
 85 90 95
 Trp Glu Phe Phe Glu Arg Leu Ile Thr Ile Ala Val Pro Arg Ile Arg
 100 105 110
 Asp Phe Arg Gly Leu Ser Ala Lys Ser Phe Asp Gly Arg Gly Asn Tyr
 115 120 125
 Ser Met Gly Val Arg Glu Gln Ile Ile Phe Pro Glu Ile Asp Tyr Asp
 130 135 140
 Lys Val Asp Arg Val Arg Gly Leu Asp Ile Thr Ile Thr Thr Thr Ala
 145 150 155 160
 Lys Ser Asp Glu Glu Gly Arg Ala Leu Leu Ala Ala Phe Asp Phe Pro
 165 170 175

Phe Arg Lys

<210> 42
 <211> 10390
 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 <222> (432)..(1862)

<400> 42
 aaaattctta cgtaatttat aatctttaaa aaaagcattt aatattgctc cccgaacgat 60

tgtgattcga	ttcacattta	aacaatttca	gaatagacaa	aaactctgag	tgtaataatg	120										
tagcctcgtg	tcttgcgagg	ataagtgc	tatgaata	ttacatat	gtgtgacctc	180										
aaaatggttc	aatattgaca	acaaaattgt	cgatcaccgc	ccttgatttg	cccttctgta	240										
gccatcacca	gagccaaacc	gattagattc	aatgtgatct	atttgtttgc	tatatcttaa	300										
ttttgccttt	tgcaaaggtc	atctctcg	tatttacttg	ttttagtaaa	tgatgggtgct	360										
tgcatatata	tctggcgaat	taatcggtat	agcagatgta	atattcacag	ggatcactgt	420										
aattaaata	a	atg	aag	gat	tat	gta	atg	gaa	aac	ttt	aaa	cat	ctc	cct	470	
		Met	Lys	Asp	Tyr	Val	Met	Glu	Asn	Phe	Lys	His	Leu	Pro		
		1				5					10					
gaa	ccg	ttc	cgc	att	cgt	gtt	att	gag	cca	gta	aaa	cgt	acc	act	cgc	518
Glu	Pro	Phe	Arg	Ile	Arg	Val	Ile	Glu	Pro	Val	Lys	Arg	Thr	Thr	Arg	
	15					20					25					
gct	tat	cgt	gaa	gag	gca	att	att	aaa	tcc	ggt	atg	aac	ccg	ttc	ctg	566
Ala	Tyr	Arg	Glu	Glu	Ala	Ile	Ile	Lys	Ser	Gly	Met	Asn	Pro	Phe	Leu	
	30				35					40					45	
ctg	gat	agc	gaa	gat	gtt	ttt	atc	gat	tta	ctg	acc	gac	agc	ggc	acc	614
Leu	Asp	Ser	Glu	Asp	Val	Phe	Ile	Asp	Leu	Leu	Thr	Asp	Ser	Gly	Thr	
				50					55					60		
ggg	gcg	gtg	acg	cag	agc	atg	cag	gct	gcg	atg	atg	cgc	ggc	gac	gaa	662
Gly	Ala	Val	Thr	Gln	Ser	Met	Gln	Ala	Ala	Met	Met	Arg	Gly	Asp	Glu	
			65					70					75			
gcc	tac	agc	ggc	agt	cgt	agc	tac	tat	gcg	tta	gcc	gag	tca	gtg	aaa	710
Ala	Tyr	Ser	Gly	Ser	Arg	Ser	Tyr	Tyr	Ala	Leu	Ala	Glu	Ser	Val	Lys	
		80					85					90				
aat	atc	ttt	ggt	tat	caa	tac	acc	att	ccg	act	cac	cag	ggc	cgt	ggc	758
Asn	Ile	Phe	Gly	Tyr	Gln	Tyr	Thr	Ile	Pro	Thr	His	Gln	Gly	Arg	Gly	
	95					100					105					
gca	gag	caa	atc	tat	att	ccg	gta	ctg	att	aaa	aaa	cgc	gag	cag	gaa	806
Ala	Glu	Gln	Ile	Tyr	Ile	Pro	Val	Leu	Ile	Lys	Lys	Arg	Glu	Gln	Glu	
	110				115					120					125	
aaa	ggc	ctg	gat	cgc	agc	aaa	atg	gtg	gcg	ttc	tct	aac	tat	ttc	ttt	854
Lys	Gly	Leu	Asp	Arg	Ser	Lys	Met	Val	Ala	Phe	Ser	Asn	Tyr	Phe	Phe	
				130					135					140		
gat	acc	acg	cag	ggc	cat	agc	cag	atc	aac	ggc	tgt	acc	gtg	cgt	aac	902
Asp	Thr	Thr	Gln	Gly	His	Ser	Gln	Ile	Asn	Gly	Cys	Thr	Val	Arg	Asn	
			145					150					155			
gtc	tat	atc	aaa	gaa	gcc	ttc	gat	acg	ggc	gtg	cgt	tac	gac	ttt	aaa	950
Val	Tyr	Ile	Lys	Glu	Ala	Phe	Asp	Thr	Gly	Val	Arg	Tyr	Asp	Phe	Lys	
		160					165					170				
ggc	aac	ttt	gac	ctt	gag	gga	tta	gaa	cgc	ggt	att	gaa	gaa	gtt	ggt	998
Gly	Asn	Phe	Asp	Leu	Glu	Gly	Leu	Glu	Arg	Gly	Ile	Glu	Glu	Val	Gly	
	175					180					185					

- 178 -

ccg	aat	aac	gtg	ccg	tat	atc	gtt	gca	acc	atc	acc	agt	aac	tct	gca	1046
Pro	Asn	Asn	Val	Pro	Tyr	Ile	Val	Ala	Thr	Ile	Thr	Ser	Asn	Ser	Ala	
190					195					200					205	
ggt	ggt	cag	ccg	gtt	tca	ctg	gca	aac	tta	aaa	gcg	atg	tac	agc	atc	1094
Gly	Gly	Gln	Pro	Val	Ser	Leu	Ala	Asn	Leu	Lys	Ala	Met	Tyr	Ser	Ile	
				210					215					220		
gcg	aag	aaa	tac	gat	att	ccg	gtg	gta	atg	gac	tcc	gcg	cgc	ttt	gct	1142
Ala	Lys	Lys	Tyr	Asp	Ile	Pro	Val	Val	Met	Asp	Ser	Ala	Arg	Phe	Ala	
			225					230					235			
gaa	aac	gcc	tat	ttc	atc	aag	cag	cgt	gaa	gca	gaa	tac	aaa	gac	tgg	1190
Glu	Asn	Ala	Tyr	Phe	Ile	Lys	Gln	Arg	Glu	Ala	Glu	Tyr	Lys	Asp	Trp	
		240					245					250				
acc	atc	gag	cag	atc	acc	cgc	gaa	acc	tac	aaa	tat	gcc	gat	atg	ctg	1238
Thr	Ile	Glu	Gln	Ile	Thr	Arg	Glu	Thr	Tyr	Lys	Tyr	Ala	Asp	Met	Leu	
	255					260					265					
gcg	atg	tcc	gcc	aag	aaa	gat	gcg	atg	gtg	ccg	atg	ggc	ggc	ctg	ctg	1286
Ala	Met	Ser	Ala	Lys	Lys	Asp	Ala	Met	Val	Pro	Met	Gly	Gly	Leu	Leu	
270					275					280				285		
tgc	atg	aaa	gac	gac	agc	ttc	ttt	gat	gtg	tac	acc	gag	tgc	aga	acc	1334
Cys	Met	Lys	Asp	Asp	Ser	Phe	Phe	Asp	Val	Tyr	Thr	Glu	Cys	Arg	Thr	
				290					295					300		
ctt	tgc	gtg	gtg	cag	gaa	ggc	ttc	ccg	aca	tat	ggc	ggc	ctg	gaa	ggc	1382
Leu	Cys	Val	Val	Gln	Glu	Gly	Phe	Pro	Thr	Tyr	Gly	Gly	Leu	Glu	Gly	
			305					310					315			
ggc	gcg	atg	gag	cgt	ctg	gcg	gta	ggt	ctg	tat	gac	ggc	atg	aat	ctc	1430
Gly	Ala	Met	Glu	Arg	Leu	Ala	Val	Gly	Leu	Tyr	Asp	Gly	Met	Asn	Leu	
		320					325					330				
gac	tgg	ctg	gct	tat	cgt	atc	gcg	cag	gta	cag	tat	ctg	gtc	gat	ggt	1478
Asp	Trp	Leu	Ala	Tyr	Arg	Ile	Ala	Gln	Val	Gln	Tyr	Leu	Val	Asp	Gly	
	335					340				345						
ctg	gaa	gag	att	ggc	gtt	gtc	tgc	cag	cag	gcg	ggc	ggt	cac	gcg	gca	1526
Leu	Glu	Glu	Ile	Gly	Val	Val	Cys	Gln	Gln	Ala	Gly	Gly	His	Ala	Ala	
350					355					360					365	
ttc	gtt	gat	gcc	ggt	aaa	ctg	ttg	ccg	cat	atc	ccg	gca	gac	cag	ttc	1574
Phe	Val	Asp	Ala	Gly	Lys	Leu	Leu	Pro	His	Ile	Pro	Ala	Asp	Gln	Phe	
				370					375					380		
ccg	gca	cag	gcg	ctg	gcc	tgc	gag	ctg	tat	aaa	gtc	gcc	ggt	atc	cgt	1622
Pro	Ala	Gln	Ala	Leu	Ala	Cys	Glu	Leu	Tyr	Lys	Val	Ala	Gly	Ile	Arg	
			385					390					395			
gcg	gta	gaa	att	ggc	tct	ttc	ctg	tta	ggc	cgc	gat	ccg	aaa	acc	ggt	1670
Ala	Val	Glu	Ile	Gly	Ser	Phe	Leu	Leu	Gly	Arg	Asp	Pro	Lys	Thr	Gly	
		400					405					410				
aaa	caa	ctg	cca	tgc	ccg	gct	gaa	ctg	ctg	cgt	tta	acc	att	ccg	cgc	1718
Lys	Gln	Leu	Pro	Cys	Pro	Ala	Glu	Leu	Leu	Arg	Leu	Thr	Ile	Pro	Arg	

- 179 -

415	420	425	
gca aca tat act caa aca cat atg gac ttc att att gaa gcc ttt aaa			1766
Ala Thr Tyr Thr Gln Thr His Met Asp Phe Ile Ile Glu Ala Phe Lys			
430	435	440	445
cat gtg aaa gag aac gcg gcg aat att aaa gga tta acc ttt acg tac			1814
His Val Lys Glu Asn Ala Ala Asn Ile Lys Gly Leu Thr Phe Thr Tyr			
	450	455	460
gaa ccg aaa gta ttg cgt cac ttc acc gca aaa ctt aaa gaa gtt taa			1862
Glu Pro Lys Val Leu Arg His Phe Thr Ala Lys Leu Lys Glu Val			
	465	470	475

ttaatactac agagtggcta taaggatggt agccactctc ttaccctaca tcctcaataa 1922
 caaaaatagc cttcctctaa aggtggcatc atgactgac aagctgaaaa aaagcactct 1982
 gcattttggg gtgttatggt tatagcaggt acagtaattg gtggagggtat gtttgcttta 2042
 cctgttgatc ttgccggtgc ctgggtttttc tgggggtgcct ttatccttat cattgcctgg 2102
 ttttcaatgc ttcattccgg gttattgtta ttagaagcaa atttaaatta tcccgtcggc 2162
 tccagtttta acaccatcac caaagattta atcggtaca cctggaacat tatcagcgg 2222
 attaccgttg ccttcgttct ctatatcctc acttatgcct atatctctgc taatgggtgcg 2282
 atcattagt aaacgatatc aatgaatttg ggttatcacg ctaatccacg tattgtcggg 2342
 atctgcacag ccattttcgt tgccagcgta ttgtgggttaa gttcggttagc cgccagtcgt 2402
 attacctcat tgttcctcgg gctgaagatt atctcctttg tgatcgtgtt tgggttctttt 2462
 ttcttccagg tcgattactc cattctgcgc gacgccacca gctccactgc gggaacgtct 2522
 tacttcccg atatatcttat ggctttgccg gtgtgtcttg cgtcatttg tttccacggc 2582
 aatattccca gcctgattat ttgctatgga aaacgcaaa ataagttaat caaaagcgtg 2642
 gtatttggtt cgctgctggc gctgggtgatt tatctcttct ggctctattg caccatgggg 2702
 aatattccgc gagaaagctt taaggcgatt atctcctcag gcggcaacgt tgattcgctg 2762
 gtgaaatcgt tcctcggcac caaacagcac ggcatatcg agttttgcct gctgggtgttc 2822
 tctaacttag ctgttgccag ttcgttcttt ggtgtcacgc tgggggttgtt cgattatctg 2882
 gcggacctgt ttaagattga taactccac ggcgggcggt tcaaaaccgt gctgttaacc 2942
 ttctgccac ctgcgttggt gtatctgac tttccgaacg gctttattta cgggatcggc 3002
 ggtgccgggc tgtgcgccac catctgggcg gtcattatc ccgcagtgt tgcaatcaaa 3062
 gctcgcaaga agtttcccaa tcagatgttc acggtctggg gcggcaatct tattccggcg 3122
 attgtcattc tctttggtat aaccgtgatt ttgtgctggt tcggcaacgt ctttaacgtg 3182
 ttacctaaat ttggctaaat ccttcaagaa gccagccatt cgctggcttc ttgcctctca 3242

- 180 -

ggaaatcact tatgtccaaa tggcaactcg cctgatactc cttcaccacg tatgctttgc 3302
gtcaccttac tatcaggacg ctttagccca tgtcccgctt tttgatttgt agttttgccc 3362
tggttttact ttatcccgcc gggattgata tgtacctcgt tggtttaccg cgcatacgccg 3422
ccgatctcaa tgccagcgaa gcgcagttgc atattgcgtt ctccgtatat ctggcgggga 3482
tggcagctgc gatgttatTTT gccggtaaag tggccgatcg ttcagggaga aagccggtcg 3542
ccatacccgg cgcggcgcta tttattattg cctcgggtgtt ctgttctactg gctgaaacca 3602
gcacgttatt tcttgcaggc cgatttctac aggggttggg cgcaggctgt tgttacgtag 3662
tggcgttcgc tattttgcgc gacacgctgg atgatcgacg tcgggctaaa gtgctgtcat 3722
tactcaacgg tattacctgc atcattccgg tgttagcgcc agtgctcgga catctgatta 3782
tgcttaaatt cccgtggcag agtctgttct gggcgatggc aatgatgggc atcgcggtac 3842
tgatgttgtc tttgtttatt ttaaaagaaa cgcgcccgac ggcccccgca gcttcggata 3902
aaccacgaga aatagcgag tcgctgctta accgtttttt cctcagccgt gttgttatca 3962
ccaccctcag cgtttcggtg atcctcactt tcgtcaacac gtcaccggta ttgctgatgg 4022
aaatcatggg gtttgagcgc ggtgaatacg ccaccattat ggcgctgacc gctggcgtca 4082
gcatgaccgt ttcattctcc acgccatttg cgctgggaat ttttaagcca cgtacgttga 4142
tgatcacctc gcagggtgta ttcctggcgg cggggatcac tcttgccgtt tcaccttccc 4202
atgcgggtttc tctgtttggg atcacgctga tttgcgccgg tttctcggtg ggttttgggtg 4262
tggcgatgag tcaggcggtta gggccgtttt cattacgcgc gggcgtagcc agctcgacct 4322
taggtattgc gcaggtttgc ggttcgtcac tgtggatttg gctggcagcg gtggttggtgta 4382
tcggcgcatg gaatatgctg atcgggatc tcgattgcctg tagcatagtg agcctgttgc 4442
tgattatgtt cgtcgcgcct ggacgccccg ttgccgctca tgaagaaatc catcaccacg 4502
cttgatctca atctgctgct ttgtctgcaa ctgctgatgc aggagcgag cgtaacccaaa 4562
gcggcgaagc ggataaacgt gacaccttcg gcggtgagta agtcgctggc aaagttaaga 4622
gcgtgggtttg acgaccgct ctttgtgaac tcaccgctgg gtctgtcgcc cacaccgctg 4682
atggtcagca tggagcaaaa tctggcggag tggatgcaaa tgagcaacct gctgctggat 4742
aaaccgcacc accagacacc gcgcggcctg aagtttgagc tggcggcgga atcaccgctg 4802
atgatgatca tgcttaatgc gctgtcgaaa cagatctacc aacgttacct gcaggcgacc 4862
atcaaattac gtaactggga ttacgattcc ttagatgcc tttactcgtgg tgaagtggat 4922
atcggttttt ccggtcgcga aagccatcct cgctcgcggg agctgttaag ctcgctaccg 4982

- 181 -

ttagccattg attatgaagt gctgttttagt gatgtgccct gcgtctgggtt acgcaaagat 5042
catccggcac tgcataaac gtggaatctg gacaccttct tacgttatcc gcatatcagc 5102
at ttgctggg aacagagcga tacctgggcg ctggacaatg tgttacagga gctgggacgc 5162
gaacgcacga ttgctatgag cctgccggaa ttcgagcagt cactgtttat ggcagcgcaa 5222
cccgacaatc tgctactggc gaccgcgccg cgctactgtc agtactacaa tcaactccat 5282
caactgccgt tggttgctct tcctctcccg tttgacgaaa gccagcaaaa aaagctggaa 5342
gttcctttta ccctgctgtg gcataaacgg aacagccata atccgaagat cgtctgggta 5402
cgggaaacca ttaaaaacct ttacgcgtcg atggcataac cgaatcgtat gaaacgcgac 5462
ccaatttcac aataaaatgt aaaaaagttg taataagctt gtctgaatcg aacttttagc 5522
cgcttttagtc tgtccatcat tccagtaa at gattactctt gtattcataa tggaccatta 5582
agcatggagc gaaaaatggc gactcacttt gcccgaggga ttttaacgga aggacatctg 5642
at ttctgttc gtctcccctc ccagtgtcat caagaagccc gaaacattcc ccctcatcgt 5702
caaagccgtt ttctggcgctc cagaggttta ctgcgagaac tgatgttcat gctgtatggc 5762
attggcgaat tgccggaaat cgtcaccctg ccgaaaggta aaccggtttt cagtataaa 5822
aat ttgcctt cgttttccat ttctatgcc gggaatatgg ttggcggtggc gttaacaacc 5882
gaaggtgaat gtggcctcga tatggaacta cagcgtgcga cgcgcgggtt tcatagccca 5942
cacgcgcccg ataaccacac cttttccagc aatgaatcgc tatggatcag taaacaaaac 6002
gatcctaacg aagcgcgggc gcagctcatc acgctgcgcc gaagcgtgct aaaactaacc 6062
ggatgatgtt tgaatgacga tccgcgcgat ctgcagctgc tgccattgc cggacgcctg 6122
aaatgtgctc atgtaaatca tgtagaagcg ttatgcgacg cggaagacgt gctgggtgtg 6182
tccgtggcgg tcacgcccac gattgaaaag ctcatgtctt gggagttaga tggcaaacac 6242
ggctggaaaa gcctgccgga tattcacagc cgcgccaaca atcctaccag ccgatgatg 6302
cgttttgccc aactctctac cgtgaaggct ttttcgcaa attgatagac aaccacagga 6362
gtcatcatgt ctgaaaaatt gcagggtggtt acgttactgg ggagcctgcg caaaggctca 6422
tttaatggca tggttgcacg taccctgccg aaaattgctc cggcgagcat ggaagtcaat 6482
gcgttaccat ccattgccga cattcccttg tatgacgctg acgtacagca ggaagaagg 6542
tttccagcaa cggttgaagc tctggcgga cagatccgtc aggctgacgg tgtggtgatc 6602
gtcacgccgg aatataacta ctcggtaccg ggtgggctga aaaatgccat cgactggctt 6662
tcccgcctgc cggatcaacc gctggccggt aaaccggtat tgattcagac cagctcaatg 6722
ggcgtgattg gcggcgcgcg ctgtcagtat cacctgcgcc agattctgggt tttcctcgat 6782

- 182 -

gcaatggtga tgaacaagcc ggaatttatg ggcggcgtga ttcagaacaa agttgatccg 6842
caaaccggag aagtgattga tcagggtacg ctggaccacc tgaccgggca attgaccgca 6902
tttggtgagt ttattcagcg agttaagatc taaataaaaa acccgccagc aatcatgcat 6962
ggcggggttt taacgcgcta tcgatttttag tgagcgtcga taaagacaat cttcaggata 7022
aacagcagcg caacgatgat tacgcacggg cttagatcac gcagacgtcc ggtaccgatt 7082
ttcatcacgc agtaggagat aaagcccagc gcgatacctt cggtaatcga gaagctgaac 7142
ggcatcatca cggcggtaat aaacgccgga acagattcag taagatcctg ccagttcacg 7202
cgtgccagac tggaagtcac cagcacgcca acgtaaatca gcgcaccagc tgcagcgtag 7262
cctggcacca tccccgccag cggcgacaga aagataacca gcaggaacag cagaccaaca 7322
accactgccg tcagaccggt acgaccgcca accgatacgc cggaagagga ctcaatataa 7382
gccgtaacgg aagaagtacc gataaacgaa ccggtcacgg aagagatact gtcgacatac 7442
agcgcctgct tcatgcgcgg gaatttcccc ttctcatccg ccagacctgc tttatcggtc 7502
acgccaatca gcgtaccgga ggagtcaaac aagttgacca acatgaaaga gaaaatcacc 7562
cctgccagcc cgaggttaaa cgaccgggt aaatctacat gaccacaaac tgtcattacg 7622
ctcggcggcg cagaaacgat gccattgtag tgcacatcac ccagcatcca gccagcagc 7682
gtcgtcacca cgatagaaac cagcaccgct gcgtgaatgt tgcgcgaggc cagaatagca 7742
atgatgaaga agccgaggat acccagaagt acgctgtgag aagtcagatt accgatgctc 7802
accagcgttt ccgggttagc gacaatcaca cctgcgtttt tcagcccat catgccaatg 7862
aacagaccga taccgctggt aatacccaca cgcagactca ccggaatggt ggctatcatc 7922
cagtagcgaa cgcggaatat cgtcagtaac agcagacctc tcgcgccccca gaagattgcg 7982
cccatcccga cctgccacgg caagcccacg gcctgtacaa cgacaaaagc gaagaacgca 8042
ttcaggccca tagcgggtgc cagtgcact ggcagggttag caaacagtcc catcataata 8102
ctgccgaatg cagcgatcag acaggtagtg acgaagacgg cgctgggtatc catgccagca 8162
acgccaagaa ttgcggggtt aacaaaaacg atgtaaacca tcgtcaggaa ggtggtaaaa 8222
ccggcgatca cttcgggtccg tgccgtcgtg ccatgttcgc gcagtttaaa cacgcgttcc 8282
agcatcccct gaccagaagt ctgggtggta tggtgatgac tcattatcta tttccgaaca 8342
aggagggaaa atccgtcgct atcgtatacc aaaatgcgac aataggcgcg tttgtgagag 8402
acttttttat tggatttact tatacggcaa cgattgcgtt gcgcaaatcg gctctacgaa 8462
aacgttaaac tgattaaaaa ggaaaggcat gtcccggata gaagcggtat ttttcgactg 8522

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D’UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 290

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 290

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

- 109 -

What is claimed is:

1. A method for identifying compounds that modulate an NIMR polypeptide activity comprising:
 - 5 contacting an NIMR polypeptide with a test compound under conditions which allow interaction of the compound with the polypeptide;
 - determining the ability of the test compound to modulate the activity of an NIMR polypeptide; and
 - selecting those compounds that modulate the activity of the NIMR polypeptide to
- 10 thereby identify compounds that modulate NIMR polypeptide activity.
2. The method of claim 1, wherein the NIMR polypeptide is selected from the group consisting of: b0357, b0447, b0853, b1448, b2530, b2889, b2948, b3469, *mdaB*, *yadG*, *yadH*, *ybjC*, *yfaE*, *yggJ*, and *yhbW*.
- 15 3. The method of claim 1, wherein the NIMR polypeptide activity comprises promoting the ability of a microbe to resist an environmental challenge.
4. The method of claim 3, wherein the NIMR polypeptide is selected from the
- 20 group consisting of: *aceG*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*, *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*, *rimK*, *rplE*, *srlA_2*, *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.
5. The method of claim 1, wherein the NIMR polypeptide activity comprises
- 25 promotion of microbial virulence.
6. The method of claim 5, wherein the NIMR polypeptide is selected from the group consisting of: *aceG*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*, *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*, *rimK*, *rplE*,
- 30 *srlA_2*, *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.

- 110 -

7. The method of any of claims 1, 3, or 5 wherein said step of determining comprises measuring the efflux of the test compound or a marker compound from the cell.
- 5 8. The method of any of claims 1, 3, or 5 wherein said step of determining comprises measuring the ability of the microbe to grow or remain viable in the presence of the environmental challenge.
9. The method of any of claims 1, 3, or 5 wherein the NIMR polypeptide is present
10 in a microbial cell.
10. The method of claim 9, wherein the NIMR polypeptide is heterologous to the cell in which it is present.
- 15 11. A method for identifying compounds that modulate an NIMR polypeptide activity comprising:
contacting an NIMR polypeptide with a test compound under conditions which allow interaction of the compound with the polypeptide;
determining the ability of the test compound to modulate the expression of an
20 NIMR polypeptide; and
selecting those compounds that modulate the expression of the NIMR polypeptide to thereby identify compounds that modulate NIMR polypeptide activity.
12. The method of claim 11, wherein the NIMR polypeptide is selected from the
25 group consisting of: b0357, b0447, b0853, b1448, b2530, b2889, b2948, b3469, *mdaB*, *yadG*, *yadH*, *ybjC*, *yfaE*, *yggJ*, and *yhbW*.
13. The method of claim 11, wherein the NIMR polypeptide is selected from the group consisting of: *aceG*, *accB*, *aceF*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*,
30 *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*, *rimK*, *rplE*, *srlA_2*, *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.

- 111 -

14. The method of any one of claims 12 or 13, wherein the step of measuring comprises measuring the amount of RNA produced by the cell.

15. The method of any one of claims 12 or 13, wherein the step of measuring
5 comprises measuring the amount or activity of a reporter gene product produced by the cell.

16. The method of claim 15 wherein the step of measuring comprises detecting the ability of an antibody to bind to the reporter gene product.

10

17. The method of any of claims 1, 3, or 5 wherein the NIMR polypeptide is present in a cell free system.

18. The method of claim 17, wherein the step of determining comprises measuring
15 the ability of the compound to bind to the NIMR polypeptide.

19. A method for decreasing the virulence of a microbe comprising exposing the microbe to an environmental challenge and to an agent that modulates the activity of an NIMR polypeptide.

20

20. A method for reducing the marA mediated transcription of an NIMR gene comprising exposing the microbe to an environmental challenge and to an agent that modulates the activity of an NIMR polypeptide.

- 112 -

21. A method for identifying compounds that modulate activity of an NIMR polypeptide in a microbe comprising:

contacting an isolated NIMR nucleic acid molecule with a test compound under conditions which allow interaction of the compound with the nucleic acid molecule;

5 determining the ability of the test compound to bind to the isolated NIMR nucleic acid molecule; and

selecting those compounds that bind to the NIMR nucleic acid molecule to thereby identify compounds that modulate activity of an NIMR polypeptide.

10 22. The method of claim 21, wherein the NIMR polypeptide is selected from the group consisting of: b0357, b0447, b0853, b1448, b2530, b2889, b2948, b3469, *mdaB*, *yadG*, *yadH*, *ybjC*, *yfaE*, *yggJ*, and *yhbW*.

23. The method of claim 21, wherein the NIMR polypeptide activity comprises
15 promoting the ability of a microbe to resist an environmental challenge.

24. The method of claim 22, wherein the NIMR polypeptide is selected from the group consisting of: *accB*, *aceF*, *aceG*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*, *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*,
20 *rimK*, *rplE*, *srlA_2*, *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.

25. The method of claim 19, wherein the NIMR polypeptide activity comprises promotion of the virulence of a microbe.

25 26. The method of claim 25, wherein the NIMR polypeptide is selected from the group consisting of: *aceG*, *ackA*, *aldA*, *cobU*, *fabB*, *fecA*, *galK*, *galT*, *gatA*, *gatC*, *glpD*, *gltA*, *gshB*, *guaB*, *hemB*, *map*, *mglB*, *mtr*, *ndh*, *nfnB*, *pflB*, *pgi*, *purA*, *ribD*, *rimK*, *rplE*, *srlA_2*, *tnaA*, *tnaL*, *tpx*, *acnA*, *mdaA*, *ribA*, and *ydeA*.

30 27. The method of claim 21, wherein the environmental challenge is an antibiotic compound.

- 113 -

28. The method of claim 21, wherein the environmental challenge is non-antibiotic compound.
29. The method of claim 28, wherein the non-antibiotic compound is a candidate
5 disinfectant or antiseptic compound.
30. A vaccine comprising at least one NIMR nucleic acid molecule or an NIMR polypeptide and a pharmaceutically acceptable carrier.
- 10 31. A composition comprising at least one compound that modulates the activity of an NIMR polypeptide and at least one antibiotic.
32. A composition comprising at least one compound that modulates the activity of an NIMR polypeptide and at least one non-antibiotic compound.
15
33. A method for reducing the virulence of a microbe in a subject suffering from a microbial infection comprising administering at least one NIMR modulating agent to the subject such that the virulence of the microbe is reduced.
- 20 34. A method for treating a microbial infection in a subject comprising administering at least one NIMR modulating agent to the subject such that the infection is treated.
35. A method for reducing the infectivity of a microbe on a surface comprising contacting the microbe with at least one NIMR modulating agent such that the
25 infectivity of the microbe is reduced.
36. The method of any one of claims 33, 34, or 35, wherein the microbe is a gram positive bacteria.
- 30 37. The method of any one of claims 33, 34, or 35, wherein the microbe is a gram negative bacteria.

- 114 -

38. The method of any one of claims 33, 34, or 35, wherein the microbe is an acid fast bacteria.

1/3

Figure 1

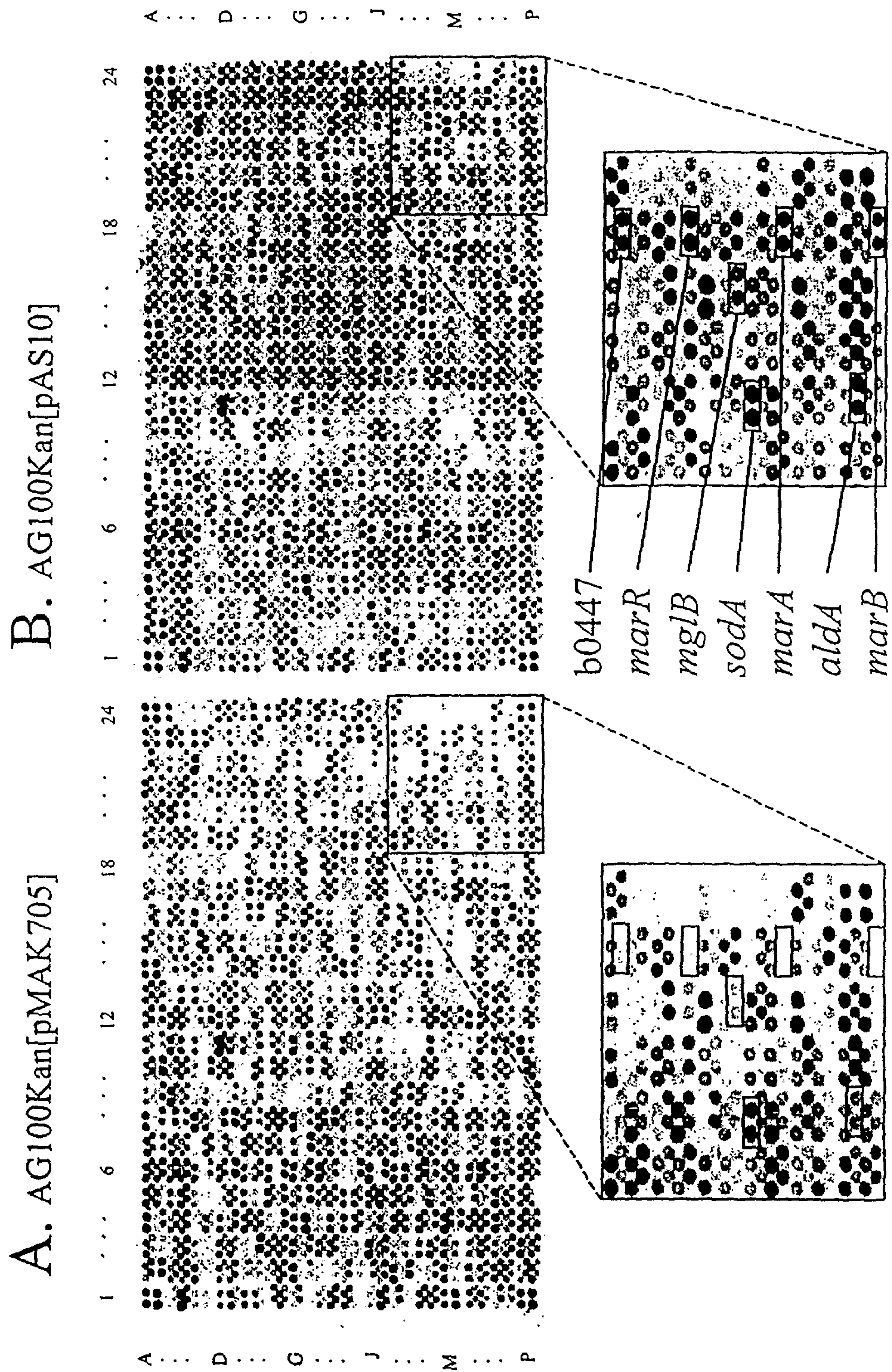
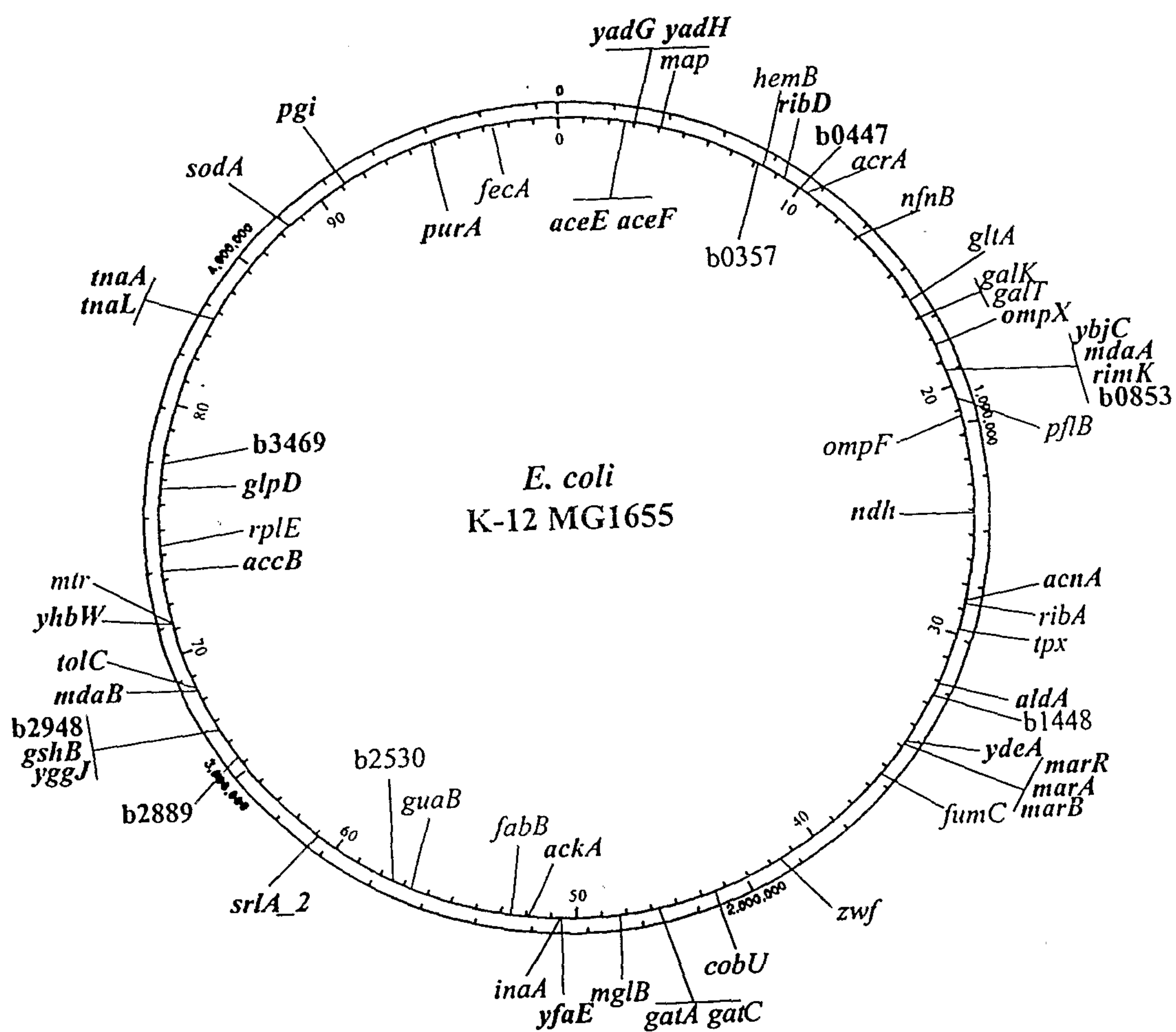


Figure 2



3/3

Figure 3

