



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N	A2	(11) International Publication Number: WO 99/61579 (43) International Publication Date: 2 December 1999 (02.12.99)
(21) International Application Number: PCT/US99/11268 (22) International Filing Date: 21 May 1999 (21.05.99) (30) Priority Data: 60/086,497 22 May 1998 (22.05.98) US (71) Applicant: TUFTS UNIVERSITY [US/US]; 4th floor, Ballou Hall, Medford, MA 02155 (US). (72) Inventor: LEVY, Stuart, B.; 144 Warren Avenue, Boston, MA 02116 (US). (74) Agents: HANLEY, Elizabeth, A. et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, 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, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, 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), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: MarA FAMILY HELIX-TURN-HELIX DOMAINS AND THEIR METHODS OF USE		
(57) Abstract An important advance in the battle against drug resistance by elucidating the domains of MarA which are critical in mediating its function. Accordingly, MarA family protein helix-turn-helix domains, mutant MarA family protein helix-turn-helix domains and methods of their use, for example, in screening assays to identify compounds which are useful as anti-infective agents and in screening assays to identify loci which are involved in mediating antibiotic resistance are described.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

MarA FAMILY HELIX-TURN-HELIX DOMAINS AND THEIR METHODS OF USE

Background of the Invention

5 Multidrug resistance in bacteria is generally attributed to the acquisition of multiple transposons and plasmids bearing genetic determinants for different 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
10 (*mar*) locus in *Escherichia coli* (George and Levy. 1983. *J. Bacteriol.* 155:531; George and Levy 1983. *J. Bacteriol.* 155:541). Mar mutants of *E. coli* arose at a frequency of 10^{-6} to 10^{-7} and were selected by growth on subinhibitory levels of tetracycline or chloramphenicol (George and Levy, supra). These mutants exhibited resistance to tetracyclines, chloramphenicol, penicillins, cephalosporins, puromycin,
15 nalidixic acid, and rifampin (George and Levy, supra). Later, the resistance phenotype was extended to include fluoroquinolones (Cohen et al. 1989. *Antimicrob. Agents Chemother.* 33:1318), oxidative stress agents (Ariza et al. 1994. *J. Bacteriol.* 176:143; Greenberg et al. 1991. *J. Bacteriol.* 173:4433), and more recently, organic solvents (White et al. 1997. *J. of Bacteriology* 179:6122; Asako, et al. 1997. *J. Bacteriol.*
20 176:143) and household disinfectants, e.g., pine oil and/or triclosan® (McMurry et al. 1998. *FEMS Microbiology Letters* 166:305; Moken et al. 1997. *Antimicrobial Agents and Chemotherapy* 41:2770).

The expression of the Mar phenotype is greater at 30°C than at 37°C (Seoane and Levy. 1995. *J. Bacteriol.* 177:3414). Continued growth in the same or higher
25 antibiotic concentrations led to increased levels of resistance, thus demonstrating a multiple antibiotic resistance phenotype which could be amplified (George and Levy, supra). Both high- and low-level resistance were decreased or completely reversed by a Tn5 insertion into a single locus at 34 min (1,636.7 kb) on the *E. coli* chromosome, called the *mar* locus. The genetic basis for high-level resistance is only partially
30 attributed to the *mar* locus, since transduction of the locus from high-or low-level *mar* mutants produces only a low level of multidrug resistance.

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 Levy. 1997. *Antimicrobial Agents and Chemother.* 41: 2067). One operon encodes MarC, a putative integral inner membrane protein without any yet
5 apparent function, but which appears to contribute to the Mar phenotype in some strains. The other operon comprises *marRAB*, encoding the Mar repressor (MarR), which binds *marO* and negatively regulates expression of *marRAB* (Cohen et al. 1994. *J. Bacteriol.* 175:1484; Martin and Rosner. 1995. *Proc. Natl. Acad. Sci. USA* 92:5456; Seoane and Levy. 1995. *J. Bacteriol.* 177:530), an activator (MarA), which controls expression of
10 other genes on the chromosome, e.g., the *mar* regulon (Cohen et al. 1994. *J. Bacteriol.* 175:1484; Gambino et. al. 1993. *J. Bacteriol.* 175:2888; Seoane and Levy. 1995. *J. Bacteriol.* 177:530), and a putative small protein (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). Proteins within this family activate
15 many different genes, some of which produce antibiotic and oxidative stress resistance or control microbial metabolism and virulence (Gallegos et al. supra).

Summary

The present invention represents an important advance in the battle against drug
20 resistance by elucidating the domains of MarA which are critical in mediating its function. Accordingly, the invention provides, *inter alia*, MarA family protein helix-turn-helix (HTH) domains, mutant MarA family protein helix-turn-helix domains and methods of their use. This new understanding of how MarA family proteins work to activate gene transcription will be invaluable to understanding and ultimately controlling
25 multidrug resistance.

In one aspect, the invention pertains to a method for identifying an antiinfective compound which affects the activity of a MarA family helix-turn-helix domain, by contacting a polypeptide comprising a MarA family helix-turn-helix domain derived from a MarA family protein with a compound under conditions which allow interaction
30 of the compound with the polypeptide such that a complex is formed; and

measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain as an indication of whether the compound is an antiinfective compound.

In another aspect, the invention pertains to a method for identifying an antiinfective compound which affects the activity of a MarA family helix-turn-helix domain, by contacting a cell expressing a Mar A family helix-turn-helix domain polypeptide derived from a MarA family protein with a compound under conditions which allow interaction of the compound with the polypeptide; and measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain polypeptide as an indication of whether the compound is an antiinfective compound.

10 In one embodiment, the step of measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain comprises detecting the ability of the complex to activate transcription from a MarA family member responsive promoter. In a preferred embodiment, the Mar A responsive promoter is selected from the group consisting of *marO*, *micF*, and *fumC*.

15 In one embodiment, the Mar A responsive promoter is linked to a reporter gene. In a preferred embodiment, the reporter gene is selected from the group consisting of *lacZ*, *phoA*, or green fluorescence protein.

In one embodiment, the step of measuring comprises measuring the amount of reporter gene product. In another embodiment, the step of measuring comprises measuring the amount of RNA produced by the cell. In yet another embodiment, the step of measuring comprises measuring the amount of a protein produced by the cell. In still another embodiment, the step of measuring comprises using an antibody against a protein produced by the cell.

In another aspect, the invention pertains to a method for identifying an antiinfective compound which affects the activity of a MarA family helix-turn-helix domain, by contacting a polypeptide comprising a Mar A family helix-turn-helix domain derived from a MarA family protein with a compound in a cell-free system under conditions which allow interaction of the compound with the polypeptide such that a complex is formed; and measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain as an indication of whether the compound is an antiinfective compound.

In one embodiment, the MarA family helix-turn-helix domain is an isolated polypeptide and the step of measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain comprises measuring the ability of the complex to bind to DNA.

5 In another embodiment of the invention, the method comprises screening a library of bacteriophage displaying on their surface a MarA family helix-turn-helix domain polypeptide, said polypeptide sequence being encoded by a nucleic acid contained within the bacteriophage, for ability to bind a compound to obtain those compounds having affinity for the helix-turn-helix domain, said method by contacting
10 the phage which display the helix-turn-helix domain with a sample of a library of compounds so that the helix-turn-helix domain can interact with and form a complex with any compound having an affinity for the helix-turn-helix domain; contacting the complex of the helix-turn-helix domain and bound compound with an agent that dissociates the bacteriophage from the compound; and identifying the compounds that
15 bound to the helix-turn-helix domain.

In another aspect, the invention pertains to a method for screening a library of bacteriophage displaying on their surface a plurality of polypeptide sequences, each polypeptide sequence being encoded by a nucleic acid contained within the bacteriophage, for ability to bind an immobilized MarA family helix-turn-helix domain,
20 to obtain those polypeptides having affinity for the helix-turn-helix domain, said method by contacting the immobilized helix-turn-helix domain with a sample of the library of bacteriophage so that the helix-turn-helix domain can interact with the different polypeptide sequences and bind those having affinity for the helix-turn-helix domain to form a set of complexes consisting of immobilized helix-turn-helix domain and bound
25 bacteriophage; separating the complexes from bacteriophage which have not formed the complex; contacting the complexes of the helix-turn-helix domain and bound bacteriophage with an agent that dissociates the bound bacteriophage from the complexes; and isolating the dissociated bacteriophage and obtaining the sequence of the nucleic acid encoding the displayed polypeptide, so that amino acid sequences of
30 displayed polypeptides with affinity for helix-turn-helix domain are obtained.

In certain embodiments, the polypeptides of the invention comprise the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein from which it is derived. In other embodiments, the polypeptides of the invention comprise the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein from which it is derived. In preferred embodiments, the polypeptides consist of the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein from which it is derived. In still other preferred embodiments, the polypeptides consist of the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein from which it is derived.

10 In preferred embodiments of the invention, the MarA family helix-turn-helix domain is derived from a protein selected from the group consisting of: MarA, RamA, AarP, Rob, SoxS, and PqrA.

In certain embodiments of the invention, a compound identified using the subject methods increases antibiotic susceptibility. In other embodiments, a compound identified using the subject methods reduces infectivity or virulence of a microbe.

15 In certain embodiments of the invention, the compound is effective against Gram negative bacteria. In other embodiments, the compound is effective against Gram positive bacteria. In preferred embodiments, the Gram positive bacteria are from a genus selected from the group consisting of: *Enterococcus*, *Staphylococcus*, *Clostridium* and *Streptococcus*. In other preferred embodiments, the compound is effective against bacteria from the family *Enterobacteriaceae*. In still other preferred embodiments, the compound is effective against a bacteria of a genus selected from the group consisting of: *Escherichia*, *Proteus*, *Klebsiella*, *Providencia*, *Enterobacter*, *Burkholderia*, *Pseudomonas*, *Aeromonas*, *Acinetobacter*, and *Mycobacteria*.

25 In another aspect, the invention pertains to a cell based method of identifying genetic loci in a microbe which affect antibiotic resistance comprising introducing into said microbe a nucleotide sequence encoding a helix-turn-helix motif of a MarA family protein and assaying for changes in the antibiotic resistance profile of said microbe. In certain embodiments, the invention further comprises assaying for changes in transcription of genetic loci of said microbe. In other embodiments the invention further comprises identifying proteins which are present in different amounts in resistant and

30

susceptible microbes. In other embodiments, the invention further comprising identifying the genes which encode said proteins.

In certain embodiments of the invention the antibiotic to which sensitivity is measured is selected from the group consisting of: tetracycline, fluoroquinolones, chloramphenicol, penicillins, cephalosporins, puromycin, nalidixic acid, and rifampin. In other embodiments, the antibiotic is a disinfectant, antiseptic, or surface delivered antibacterial compound. In yet other embodiments, the antibiotic is an antifungal. In still other embodiments, the antibiotic is an antiparasitic.

In another aspect, the invention pertains to a cell-free method of identifying genetic loci in an microbe which affect resistance to antibiotics comprising contacting a nucleic acid molecule of said microbe with a MarA family protein helix-turn-helix domain and allowing complexes to form; separating the nucleic acid molecule which has formed a complex with a helix-turn-helix domain from the helix-turn-helix domain; and identifying the sequence of those nucleic acid molecules which can bind to a MarA family protein helix-turn-helix domain.

In certain embodiments of the invention, the compound to be tested is derived from a library of small molecules. In other embodiments, the compound is a nucleic acid molecule. In still other embodiments, the compound is an antisense or sense oligonucleotide. In yet other embodiments, the compound is a naturally occurring small organic molecule.

In another aspect, the invention pertains to a kit for identifying genetic loci in an microbe which affect resistance to compounds comprising a nucleotide sequence encoding a naturally occurring helix-turn-helix domain of a MarA family protein and mutant, inactive form of a MarA family protein helix-turn-helix domain.

25

Brief Description of the Drawings

Figure 1 is a schematic showing how the MarA mutants of the present invention were constructed. Panel A shows the nucleotide sequence of the mutagenic oligonucleotide. Panel B shows the wild-type MarA amino acid residues 27 to 44, of MarA (based on the sequence provided in Cohen et al. 1993. *J. Bacteriol.* 175:1484). Mutants contained the amino acids shown at the indicated sites of insertion in helix A

30

Preferred antiinfective compounds increase the susceptibility of microbes to antibiotics or decrease the infectivity or virulence of a microbe. The term "microbe" includes any unicellular microbe, e.g., bacteria, fungi, or protozoa. Therefore, agents which inhibit the proliferation and/or viability of fungi or protozoa are also included in this term. In
5 preferred embodiments, microbes are pathogenic for humans, animals, or plants, however in other embodiments, microbes are involved, e.g., in fouling or spoilage.

As used herein, the term "antibiotic" includes antimicrobial agents isolated from natural sources or chemically synthesized. The term "antibiotic" includes the antimicrobial agents to which the Mar phenotype has been shown to mediate resistance
10 and, as such, includes disinfectants, antiseptics, and surface delivered compounds. For example, any antibiotic, biocide, or other type of antibacterial compound, including agents which induce oxidative stress agents, and organic solvents are included in this term. Preferred antibiotics include: tetracycline, fluoroquinolones, chloramphenicol, penicillins, cephalosporins, puromycin, nalidixic acid, and rifampin.

As used herein, the language "MarA family protein" includes the many naturally occurring transcription regulation proteins which have sequence similarities to MarA and which contain the MarA family signature pattern, which can also be referred to as an XylS/AraC signature pattern. An exemplary signature pattern which defines MarA family proteins is shown, e.g., on PROSITE and is represented by the sequence: [KRQ]-
20 [LIVMA]-X(2)-[GSTALIV]-{FYWPGDN}X(2)-[LIVMSA]-X(4,9)-[LIVMF]-X(2)-[LIVMSTA]-X(2)-[GSTACIL]-X(3)-[GANQRF]-[LIVMFY]-X(4,5)-[LFY]-X(3)-[FYIVA]-{FYWHCM}-X(3)-[GSADENQKR]-X-[NSTAPKL]-[PARL], where X is any amino acid. MarA family proteins have two "helix-turn-helix" domains. This signature pattern was derived from the region that follows the first, most amino terminal, helix-
25 turn-helix domain (HTH1) and includes the totality of the second, most carboxy terminal helix-turn-helix domain (HTH2). (See PROSITE PS00041).

MarA family polypeptide sequences are "structurally related" to one or more known MarA family members, preferably to MarA. This structural relatedness can be shown by sequence similarity between two MarA family polypeptide sequences or
30 between two MarA family nucleotide sequences. Sequence similarity can be shown, e.g., by optimally aligning MarA family member sequences using an alignment program

for purposes of comparison and comparing corresponding positions. To determine the degree of similarity between sequences, they will be aligned for optimal comparison purposes (*e.g.*, gaps may be introduced in the sequence of one protein for nucleic acid molecule for optimal alignment with the other protein or nucleic acid molecules). The amino acid residues or bases and corresponding amino acid positions or bases are then 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 molecules 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 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.

MarA family polypeptides share some amino acid sequence similarity with MarA. The nucleic acid and amino acid sequences of MarA as well as other MarA family polypeptides are available in the art. For example, the nucleic acid and amino acid sequence of MarA can be found, *e.g.*, on GeneBank (accession number M96235 or in Cohen et al. 1993. *J. Bacteriol.* 175:1484, or in SEQ ID NO:1 and SEQ ID NO:2).

The nucleic acid and/or amino acid sequences of MarA can be used as "query sequences" to perform a search against databases (*e.g.*, either public or private) to, for example, identify other MarA family members having related sequences. Such searches can be performed, *e.g.*, 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 NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to MarA family nucleic acid molecules. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 5 3 to obtain amino acid sequences homologous to MarA protein 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
10 <http://www.ncbi.nlm.nih.gov>.

MarA family members can also be identified as being structurally similar based on their ability to specifically hybridize to nucleic acid sequences specifying MarA. 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.
15 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
20 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
25 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).

Preferably, the nucleic acid sequence of a MarA family member identified in this way is at least about 10%, 20%, more preferably at least about 30%, more preferably at
30 least about 40% identical and most preferably at least about 50%, or 60% identical or

more with a MarA nucleotide sequence. Preferably, MarA family members have an amino acid sequence at least about 20%, more preferably at least about 30%, more preferably at least about 40% identical and most preferably at least about 50%, or 60% or more identical with a MarA amino acid sequence. However, it will be understood that the level of sequence similarity among microbial regulators of gene transcription, 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 MarA family members can also be determined based on “three-dimensional correspondence” of amino acid residues. As used herein, the language “three-dimensional correspondence” is meant to include residues which spatially correspond, e.g., are in the same functional position of a MarA family protein 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 DNA or bind the same cofactor, as determined, e.g., by mutational analysis.

Exemplary MarA family proteins are shown in Table 1, in Figures 2 and 3, and at Prosite (PS00041) and include: AarP, Ada, AdaA, AdiY, AfrR, AggR, AppY, AraC, CafR, CelD, CfaD, CsvR, D90812, EnvY, ExsA, FapR, HrpB, InF, InvF, LcrF, LumQ, MarA, MelR, MixE, MmsR, MsmR, OrfR, Orf_f375, PchR, PerA, PocR, PqrA, RafR, RamA, RhaR, RhaS, Rns, Rob, SoxS, S52856, TetD, TcpN, ThcR, TmbS, U73857, U34257, U21191, UreR, VirF, XylR, XylS, Xys1, 2, 3, 4, Ya52, YbbB, YfiF, YisR, YzbC, and YijO.

In preferred embodiments, a MarA family protein excludes one or more of XylS, AraC, and MelR. In other preferred embodiments, a MarA family protein is involved in antibiotic resistance. In particularly preferred embodiments, a MarA family protein is selected from the group consisting of: MarA, RamA, AarP, Rob, SoxS, and PqrA.

Preferred MarA family polypeptides are “naturally occurring.” As used herein, a “naturally-occurring” molecule refers to an MarA family molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural MarA family protein). 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 bind to DNA and regulate transcription. Such variants can be made, e.g., by mutation using techniques which are known in the art. Alternatively, variants can be chemically synthesized. For example, it will be understood that the MarA family polypeptides described herein, are also meant to include equivalents thereof. Such variants can be made, e.g., by mutation using techniques which are known in the art. Alternatively, variants can be chemically synthesized. For instance, mutant forms of MarA family 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 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 protein, corresponding either to defined structural or functional determinants (e.g., the first or second helix of a helix-turn-helix 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, Gene 96:125-128).

In certain embodiments, such variants have at least 60% amino acid identity with a naturally occurring MarA family member protein. In preferred embodiments, such variants have at least about 70% amino acid identity with a naturally occurring MarA family member protein. In more preferred embodiments, such variants have at least about 80% amino acid identity with a naturally occurring MarA family member protein. 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 MarA family member protein. In yet other embodiments, a nucleic acid molecule encoding a variant of a MarA family protein is capable of hybridizing under stringent conditions to a nucleic acid molecule encoding a naturally occurring MarA family protein.

The language "mutant form of a MarA family helix-turn-helix domain" includes mutant forms of such MarA family helix-turn-helix domains which do not retain the same biological activity as the naturally occurring form. For example, such mutants may not bind to a MarA family member promoter or may not initiate transcription from
5 MarA family member responsive promoter or may initiate transcription at a lower level than the naturally occurring MarA family member.

As used herein the language "activity of a MarA family helix-turn-helix domain" includes the ability of the helix-turn-helix domain to interact with DNA, e.g., to bind to a MarA family protein responsive promoter, or to initiate transcription from such a
10 promoter.

As used herein, the language "*marA* family protein responsive promoter" includes promoters which initiate transcription of an operon in a microbe and is structurally or functionally related to the *marA* promoter, e.g., is bound by MarA or a protein related to MarA. Preferably, the *marA* family protein responsive promoter is a
15 *marRAB* promoter. For example, in the *mar* operon, several promoters are *marA* family protein responsive promoters as defined herein, e.g., the 405-bp *Thal* fragment from the *marO* region is a *marA* family responsive promoter (Cohen et al. 1993. *J. Bact.* 175:7856). In addition, MarA has been shown to bind to a 16 bp MarA binding site (referred to as the "marbox" within *marO* (Martin et al. 1996. *J. Bacteriol.* 178:2216).
20 MarA also initiates transcription from the *acrAB*; *micF*; *mlr 1,2,3*; *slp*; *nfo*; *inaA*; *fpr*; *sodA*; *soi-17,19*; *zwf*; *fumC*; or *rpsF* promoters (Alekshun and Levy. 1997. *Antimicrobial Agents and Chemother.* 41:2067). Other *marA* family responsive promoters are known in the art and include: *araBAD*, *araE*, *araFGH* and *araC*, which are activated by AraC; *Pm*, which is activated by XylS; *melAB* which is activated by
25 MelR; and *oriC* which is bound by Rob.

The language "MarA family protein responsive promoter" also includes portions of the above promoters which are sufficient to activate transcription upon interaction with a MarA family member protein. The portions of any of the MarA family protein-responsive promoters which are minimally required for their activity can be easily
30 determined by one of ordinary skill in the art, e.g, using mutagenesis. Exemplary techniques are described by Gallegos et al. (1996. *J. Bacteriol.* 178:6427). A "MarA

family protein responsive promoter" also includes non-naturally occurring homologs of MarA family protein responsive promoters which have the same function as naturally occurring MarA family promoters. Preferably such variants have at least 60% nucleotide sequence identity with a naturally occurring MarA family protein responsive promoter. In preferred embodiments, such variants have at least about 70% nucleotide sequence identity with a naturally occurring MarA family protein responsive promoter. In more preferred embodiments, such variants have at least about 80% nucleotide sequence identity with a naturally occurring MarA family protein responsive promoter. In particularly preferred embodiments, such variants have at least about 90% nucleotide sequence identity and preferably at least about 95% nucleotide sequence identity with a naturally occurring MarA family protein responsive promoter. In yet other embodiments nucleic acid molecules encoding variants of MarA family protein responsive promoters are capable of hybridizing under stringent conditions to nucleic acid molecules encoding naturally occurring MarA family protein responsive promoters.

15 The term "interact" includes close contact between molecules that results in a measurable effect, e.g., the binding of one molecule with another. For example, a MarA family polypeptide can interact with a MarA family protein responsive promoter and alter the level of transcription of DNA. Likewise, compounds can interact with a MarA family polypeptide and alter the activity of a MarA family polypeptide.

20 As used herein, the term "multiple drug resistance (MDR)" includes resistance to both antibiotic and non-antibiotic compounds. MDR results from the increased transcription of a chromosomal or plasmid encoded genetic locus in an organism, e.g., a *marRAB* locus, that results in the ability of the organism to minimize the toxic effects of a compound to which it has been exposed, as well as to other non-related compounds, e.g., by stimulating an efflux pump(s) or microbiological catabolic or metabolic processes.

As used herein the term "reporter gene" includes any gene which encodes an easily detectable product which is operably linked to a regulatory sequence, e.g., to a MarA family protein responsive 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 such that the reporter gene is

30

transcribed. In preferred embodiments, a reporter gene consists of the MarA family protein responsive promoter linked in frame to 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 effected 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.

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 protein (U.S. patent 5,491,084; WO96/23898).

As used herein the term "compound" includes any reagent or test agent which is employed in the assays of the invention and assayed for its utility as an antiinfective compound based on its ability to influence the activity of a MarA family helix-turn-helix domain, e.g., by binding to that domain. More than one compound, e.g., a plurality of compounds, can be tested at the same time for their ability to modulate the activity of a MarA family HTH domain activity in a screening assay.

Compounds that can be tested in the subject assays include antibiotic and non-antibiotic compounds. In one embodiment, compounds include candidate 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 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 L-amino acid residues linked by natural peptide bonds. "Non-peptidic compounds" also are intended to include natural products.

10 As used herein, the term "genetic loci" includes an oligonucleotide sequence encoding a peptide or a transcriptional regulatory element (e.g., a promoter, operator, or other regulatory element). A locus may consist of a start codon, a stop codon, and at least one codon encoding an amino acid residue. Typically, a locus is transcribed to produce an mRNA transcript and that transcript is translated to produce a polypeptide.

15

II. MarA Family Protein Helix-Turn-Helix Domains

Helix-turn-helix domains are known in the art and have been implicated in DNA binding (*Ann Rev. of Biochem.* 1984. 53:293). An example of the consensus sequence for a helix-turn domain can be found in Brunelle and Schleif (1989. *J. Mol. Biol.* 209:607). The domain has been illustrated by the sequence XXXPhoAlaXXPhoGlyPhoXXXXPhoXXPhoXX, where X is any amino acid and Pho is a hydrophobic amino acid.

The crystal structure of MarA has been determined and the first (most amino terminal) HTH domain of MarA has been identified as comprising from about amino acid 31 to about amino acid 52 and the second HTH domain of MarA has been identified as comprising from about amino acid 79 to about amino acid 102 (Rhee et al. 1998. *Proc. Natl. Acad. Sci. USA.* 95:10413).

Locations of the helix-turn-helix domains in other MarA family members can easily be found by one of skill in the art. For example using the MarA protein sequence and an alignment program, e.g., the ProDom program, a portion of the MarA amino acid sequence e.g., comprising one or both HTH domains of MarA (such as from about

30

amino acid 30 to about amino acid 107 of MarA as was done to generate Figure 2) to produce an alignment. Exemplary alignments are shown in Figures 2 and 3. Using such an alignment, the amino acid sequences corresponding to the HTH domains of MarA can be identified in other MarA family member proteins. An exemplary consensus
5 sequence for the first helix-turn-helix domain of a MarA family protein can be illustrated as XXXXAXXXXSXXXLXXXFX, where X is any amino acid. An exemplary consensus sequence for the second helix-turn-helix domain of a MarA family protein is illustrated as XXIXXIAXXXGFXSXXXFXXX[F/Y], where X is any amino acid. Preferably, a MarA family protein first helix-turn-helix domain comprises the consensus
10 sequence E/D-X-V/L-A-D/E-X-A/S-G-X-S-X3-L-Q-X2-F-K/R/E-X2-T/I. Preferably, a MarA family protein second helix-turn-helix domain comprises the consensus sequence I-X-D-I-A-X3-G-F-X-S-X2-F-X3-F-X4.

Preferably, a MarA family member HTH domain is a MarA HTH domain. The first and second helix-turn-helix domains of MarA are, respectively,
15 EKVSESGYSKWHLQRMFKKET and ILYLAERYGFESQQTLTRTFKNYF. Other exemplary MarA family helix-turn-helix domains include: about amino acid 210 to about amino acid 229 and about amino acid 259 to about amino acid 278 of MelR; about amino acid 196 to about amino acid 215 and about amino acid 245 to about amino acid 264 of AraC; and about amino acid 230 to about amino acid 249 (or 233-253) and about
20 amino acid 281 to about amino acid 301 (or 282-302) of XylS (see e.g., Brunelle et al. 1989. J. Mol. Biol. 209:607; Niland et al. 1996. J. Mol. Biol. 264:667; Gallegos et al. 1997. Microbiology and Molecular Biology Reviews. 61:393).

“MarA family protein helix-turn-helix domains” are derived from or are homologous to the helix-turn-helix domains found in the MarA family proteins as
25 described supra. In preferred embodiments, a MarA family protein excludes one or more of XylS, AraC, and MelR. In particularly preferred embodiments, a MarA family protein is selected from the group consisting of: MarA, RamA, AarP, Rob, SoxS, and PqrA.

Both of the helix-turn-helix domains present in MarA family proteins are in the
30 carboxy terminal end of the protein. Proteins or portions thereof comprising either or both of these domains can be used in the instant methods. In certain embodiments, a

polypeptide which is used in screening for compounds comprises the helix-turn-helix domain most proximal to the carboxy terminus (HTH2) of the MarA family protein from which it is derived. In other embodiments, such a polypeptide comprises the helix-turn-helix domain most proximal to the amino terminus (HTH1) of the MarA family protein from which it is derived. In one embodiment, other polypeptide sequences may also be present, e.g., sequences that might facilitate immobilizing the domain on a support, or, alternatively, might facilitate the purification of the domain.

In preferred embodiments, such a polypeptide consists essentially of the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein from which it is derived. In other preferred embodiments, such a polypeptide consists essentially of the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein from which it is derived.

In preferred embodiments, such a polypeptide consists of the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein from which it is derived. In other preferred embodiments, such a polypeptide consists of the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein from which it is derived.

MarA family protein helix-turn-helix domains can be made using techniques which are known in the art. The nucleic acid and amino acid sequences of MarA family proteins are available, for example, from GenBank. Using this information, the helix-turn-helix consensus motif and mutational analysis provided herein, one of ordinary skill in the art can identify MarA family helix-turn-helix domains.

In certain embodiments of the invention it will be desirable to obtain "isolated or recombinant" nucleic acid molecules encoding MarA family helix-turn-helix domains or mutant forms thereof. By "isolated or recombinant" is meant a nucleic acid molecule which has been (1) amplified in vitro by, for example, polymerase chain reaction (PCR); (2) recombinantly produced by cloning, or (3) purified, as by cleavage and gel separation; or (4) synthesized by, for example, chemical synthesis. Such a nucleic acid molecule is isolated from the sequences which naturally flank it in the genome and from cellular components.

The isolated or recombinant nucleic acid molecules encoding MarA family helix-turn-helix protein domains can then, for example, be utilized in binding assays, can be expressed in a cell, or can be expressed on the surface of phage, as discussed further below.

5 In yet other embodiments of the invention, it will be desirable to obtain a substantially purified or recombinant MarA family helix-turn-helix polypeptide. Such polypeptides, for example, can be purified from cells which have been engineered to express an isolated or recombinant nucleic acid molecule which encodes a MarA family helix-turn-helix domain. For example, as described in more detail below, a bacterial cell
10 can be transformed with a plasmid which encodes a MarA family helix-turn-helix domain. The MarA family helix-turn-helix protein can then be purified from the bacterial cells and used, for example, in the cell-free assays described herein.

Purification of a MarA family helix-turn-helix domain can be accomplished using techniques known in the art. For example, column chromatography could be used,
15 or antibodies specific for the domain or for a polypeptide fused to the domain can be employed, for example on a column or in a panning assay.

In preferred embodiments, cells used to express MarA family helix-turn-helix domains for purification, e.g., host cells, comprise a mutation which renders any endogenous MarA family member protein nonfunctional or causes the endogenous
20 protein to not be expressed. In other embodiments, mutations may also be made in MarR or related genes of the host cell, such that repressor proteins which bind to the same promoter as a MarA family protein are not expressed by the host cell.

III. Mutant MarA Family Helix-Turn-Helix Domains

25 In certain embodiments of the invention, it will be desirable to use a mutant form of a MarA family protein helix-turn-helix domain, e.g., a non-naturally occurring form of a MarA family helix-turn-helix domain which has altered activity, e.g., does not retain wild type MarA family protein helix-turn-helix domain activity, or which has reduced activity or which is more active when compared to a wild-type MarA family protein
30 helix-turn-helix domain.

Such mutant forms can be made using techniques which are well known in the art. For example, random mutagenesis can be used. Using random mutagenesis one can mutagenize an entire molecule or one can proceed by cassette mutagenesis. In the former instance, 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 second approach, discrete regions of a protein, corresponding either to defined structural or functional determinants (e.g., the first or second alpha helix of a helix-turn-helix 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 a preferred embodiment, PCR mutagenesis is used. For example, Example 2 describes the use of Megaprimer PCR (O.H. Landt, *Gene* 96:125-128) used to introduce an *NheI* restriction site into the centers of both the helix A (position 1989) and helix B (position 2016) regions of the *marA* gene.

In one embodiment, such mutant helix-turn-helix domains comprise one or more mutations in the helix-turn-helix domain most proximal to the carboxy terminus (HTH2) of the MarA family protein molecule. In a preferred embodiment, the mutation comprises an insertion into helix A and helix B of the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein. In one embodiment, such mutant helix-turn-helix domains comprise one or more mutations in the helix-turn-helix domain most proximal to the amino terminus (HTH1) of the MarA family protein molecule. In a preferred embodiment, the mutation comprises an insertion into helix A and helix B of the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein. In particularly preferred embodiments, the mutation is selected from the group consisting of: an insertion at an amino acid corresponding to about position 33 of MarA and an insertion at an amino acid position corresponding to about position 42 of MarA. "Corresponding" amino acids can be determined, e.g, using an alignment of the helix-turn-helix domains, such as that shown in Figure 2.

Such mutant forms of MarA family helix-turn-helix motifs are useful as controls to verify the specificity of antiinfective compounds for a MarA family helix-turn-helix domain or as controls for the identification of genetic loci which affect resistance to

antiinfectives. For example, the mutant MarA family helix-turn-helix domains described in the appended Examples demonstrate that insertional inactivation of MarA at either helix A or helix B in the first HTH domain abolished the multidrug resistance phenotype in both *E. coli* and *M. smegmatis*. By the use of an assay system such as that
5 described in Example 2, which demonstrates the ability of MarA family protein helix-turn-helix domains to increase antibiotic resistance and that mutant forms of these domains do not have the same effect, one can clearly show that the response of any genetic loci identified is specific to a MarA family helix-turn-helix domain.

10 **IV. Expression of MarA Family Helix-Turn-Helix Domains**

Nucleic acids encoding MarA family protein helix-turn-helix domains can be expressed in cells using vectors. Almost any conventional delivery vector can be used. Such 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
15 use with a given microbial cell. The sequences encoding these domains 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.

These nucleic acids can be introduced into microbial cells using standard
20 techniques, for example, by transformation using calcium chloride or electroporation. Such techniques for the introduction of DNA into microbes are well known in the art.

V. Methods of Identifying Antimicrobial/Antiinfective Compounds Which Interact With MarA Family Helix-Turn-Helix Domains

25 In one embodiment, the invention provides for methods of identifying an antiinfective compound which affects the activity of a MarA family helix-turn-helix domain, by contacting a polypeptide comprising a Mar A family helix-turn-helix domain derived from a MarA family protein with a compound under conditions which allow interaction of the compound with the polypeptide. The ability of the compound to
30 reduce an activity of a MarA family protein helix-turn-helix domain is used as an indication of whether the compound is an antimicrobial compound which interferes with

the ability of a microbe to grow or an antiinfective compound which interferes with the ability of a microbe to cause infection in a host.

A variety of different techniques can be used to determine whether a compound reduces the activity of a helix-turn-helix domain. For example, the ability of a
5 compound to decrease binding of a MarA family protein to DNA, e.g., to a MarA family protein responsive promoter, or the ability of the compound to reduce MarA family protein initiated transcription from such a promoter can be measured. As described in more detail below, either whole cell or cell free assay systems can be employed.

10 **A. Whole Cell Assays**

In certain embodiments of the invention, the step of determining whether a compound affects the activity of a MarA family helix-turn-helix domain comprises measuring the ability of the compound to reduce the ability of a MarA family helix-turn-helix domain to activate transcription from a MarA responsive promoter. In such an
15 assay, since the MarA family member helix-turn-helix domain would normally bind to the MarA responsive promoter to induce transcription, a compound would be identified based on its ability to reduce this control level of transcription as compared to a cell which had been transfected with the MarA family helix-turn-helix domain but which had not been treated with the compound.

20 In preferred embodiments, to provide a convenient readout of the transcription from a MarA family protein responsive promoter, such a promoter is linked to a reporter gene. For example, a bacterial cell, e.g., an *E. coli* cell, can be transfected with plasmids comprising a *pm-lacZ* reporter gene construct and a plasmid comprising XylS. XylS activates transcription from the *pm* promoter under control conditions leading to
25 transcription and the production of reporter gene product. The ability of a compound to interfere with this interaction is indicated by a reduction in this control level of transcription and, thus, a reduction in the amount of reporter gene product. The amount of reporter gene product can be measured grossly in intact cells, e.g., by looking at color changes in cells, for example, by using the *lacZ* reporter gene and plating the cells on
30 media supplemented with X-Gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside

or, for example, by lysing cells and measuring the amount of product produced, e.g., by absorbance or enzyme activity.

In yet other embodiments, the step of detecting the ability of a compound to induce a change in transcription comprises measuring the amount of RNA produced by the cell. In such embodiments, the cells may or may not comprise a reporter gene construct. For example, the RNA can be isolated from cells which express a MarA family helix-turn-helix domain which have been incubated in the presence and absence of compound. Northern Blots using probes specific for the sequences to be detected can then be performed using techniques known in the art. Sequences which can be detected include any sequences which are linked to a MarA family responsive promoter, including, for example, both endogenous sequences and reporter gene sequences. Exemplary endogenous sequences which can be detected include: *acrAB*; *micF*; *mlr 1,2,3*; *slp*; *nfo*; *inaA*; *fpr*; *sodA*; *soi-17,19*; *zwf*; *fumC*; or *rpsF*; *araBAD*, *araE*, *araFGH* and *araC*, which are activated by AraC; *pm*, which is activated by XylS; *melAB* which is activated by MelR; and *oriC* which is activated by Rob., as well as sequences from genetic loci that are identified using the assays described infra.

In yet other embodiments, the ability of a compound to induce a change in transcription from a MarA responsive promoter can be accomplished by measuring the amount of a protein produced by the cell. Proteins which can be detected include any proteins which are produced upon the activation of a MarA family responsive promoter, including, for example, both endogenous sequences and reporter gene sequences. Exemplary endogenous proteins which can be detected include: AcrAB; Mlr 1,2,3; Slp; Nfo; InaA; Fpr; SodA; Soi-17,19; Zwf; FumC; or RpsF promoters (Alekhshun and Levy. 1997. *Antimicrobial Agents and Chemother.* 41:2067). Others are known in the art and include: AraBAD, AraE, AraFGH and AraC, which are activated by AraC; *pm*, which is activated by XylS; MelAB which is activated by MelR; and *oriC* which is bound by Rob, as well as proteins translated from genetic loci that are identified using the assays described infra. In one embodiment, the amount of protein made by a cell can be detected using an antibody against that protein. In other embodiments, the activity of such a protein can be measured.

B. Cell-Free Assays

In other embodiments, the ability of a compound to affect the activity of a MarA family helix-turn-helix domain is accomplished using isolated MarA family helix-turn-helix polypeptides in a cell-free system. In such an assay the step of measuring the ability of a compound to affect the activity of the MarA family helix-turn-helix polypeptide is accomplished by measuring the effect of the compound on the ability of helix-turn-helix domain to bind to DNA.

For example, the ability of a helix-turn-helix domain to bind to DNA can be measured by end labeling a nucleic acid molecule which encodes for a MarA family responsive promoter with ^{32}P using techniques which are known in the art (see e.g., Martin and Rosner. 1995. *Proc. Natl. Acad. Sci. USA* 92:5456). The helix-turn-helix domain can then be incubated with the compound to be tested to form a complex. The complex can then be incubated with the labeled MarA family protein responsive promoter. The sample can then be electrophoresed to look for changes in the mobility of the sample as compared to the mobility of the helix-turn-helix domain-promoter complex in the absence of the compound (Martin and Rosner, supra).

In yet another method of detecting the ability of a compound to bind a MarA family helix-turn-helix domain, the helix-turn-helix domain polypeptide sequence can be expressed by a bacteriophage. In this embodiment the phages which display the helix-turn-helix domain would then be contacted with a compound so that the helix-turn-helix domain can interact with and potentially form a complex with the compound. Phage which have formed complexes with compounds can then be separated from those which have not. The complex of the helix-turn-helix domain and compound can then be contacted with an agent that dissociates the bacteriophage from the compound. Any compounds that bound to the helix-turn-helix domain can then be isolated and identified.

In a variation of this method that allows for screening of compounds which are polypeptides and which bind to helix-turn-helix domains, a library of bacteriophage which display on their surface a plurality of polypeptide sequences can be tested for their ability to bind a MarA family helix-turn-helix domain to obtain those polypeptides having affinity for the helix-turn-helix domain. The complexes of bound bacteriophage and helix-turn-helix domain can be separated, and then treated with an agent that

dissociates the bound bacteriophage from the complexes and the sequence of the nucleic acid encoding the displayed polypeptide can be obtained.

VII. Microbes Suitable For Testing

5 Numerous different microbes are suitable for testing in the instant assays. As such, they may be used as intact cells or as sources of DNA as described herein.

In preferred embodiments, microbes for use in the claimed methods are bacteria, either Gram negative or Gram positive bacteria. More specifically, any bacteria that are shown to become resistant to antibiotics, e.g., to display a Mar phenotype are
10 appropriate for use in the claimed methods.

In preferred embodiments, microbes suitable for testing are bacteria from the family *Enterobacteriaceae*. In more preferred embodiments, the antiinfective is effective against a bacteria of a genus selected from the group consisting of:
Escherichia, Proteus, Salmonella, Klebsiella, Providencia, Enterobacter, Burkholderia,
15 *Pseudomonas, Aeromonas, Haemophilus, Yersinia, Neisseria, and Mycobacteria.*

In yet other embodiments, the microbes to be tested are Gram positive bacteria and are from a genus selected from the group consisting of: *Lactobacillus, Azorhizobium, Streptomyces, Pediococcus, Photobacterium, Bacillus, Enterococcus, Staphylococcus, Clostridium, and Streptococcus.*

20 In other embodiments, the microbes to be tested are fungi. In a preferred embodiment the fungus is from the genus *Mucor* or *Candida*, e.g., *Mucor racmeosus* or *Candida albicans*.

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

25

VIII. Test Compounds

Compounds for testing in the instant methods can be derived from a variety of different sources and can be known or can be novel. In one embodiment, libraries of compounds are tested in the instant methods to identify MarA family protein blocking
30 agents. In another embodiment, known compounds are tested in the instant methods to identify MarA family protein blocking agents. In a preferred embodiment, compounds

among the list of compounds generally regarded as safe (GRAS) by the Environmental Protection Agency are tested in the instant methods.

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
5 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
10 described an approach for the synthesis of molecular libraries of small organic molecules with a diversity of 10⁴-10⁵ (Carell et al. 1994. *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell et al. *Angew. Chem. Int. Ed. Engl.* 1994. 33:2061).

The compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including:
15 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 libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries
20 of compounds (Lam, K.S. *Anticancer Drug Des.* 1997. 12:145).

Exemplary compounds which can be screened for activity include, but are not limited to, peptides, nucleic acids, carbohydrates, small organic molecules, 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
25 non-peptidic compounds.

Other exemplary methods for the synthesis of molecular libraries can be found in 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.

30

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 '409), plasmids (Cull et al. (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage
5 (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, combinatorial polypeptides can be produced from a cDNA library.

10 In other embodiments, the compounds can be nucleic acid molecules. In preferred embodiments, nucleic acid molecules for testing are small oligonucleotides. Such oligonucleotides can be randomly generated libraries of oligonucleotides or can be specifically designed to reduce the activity of a MarA protein family member helix-turn-helix domain. For example, in one embodiment, these oligonucleotides are sense or
15 antisense oligonucleotides. In preferred embodiments, oligonucleotides for testing are sense to the binding site of a MarA protein family member helix-turn-helix domain. Methods of designing such oligonucleotides given the sequences of the MarA family member protein helix-turn-helix domains is within the skill of the art.

In preferred embodiments, controls should be included to ensure that any
20 compounds which are identified using the subject assays do not merely appear to decrease the activity of a MarA family helix-turn-helix domain because they inhibit protein synthesis. For example, if a compound appears to inhibit the synthesis of a protein being translated from RNA which is transcribed upon activation of a MarA family responsive promoter, it may be desirable to show that the synthesis of a control,
25 e.g., a protein which is being translated from RNA which is not transcribed upon activation of a MarA family responsive promoter, is not affected by the addition of the same compound. For example, the amount of the MarA family helix-turn-helix polypeptide being made or the amount of an endogenous protein could be tested. In another embodiment the microbe could be transformed with another plasmid comprising

a promoter which is not a MarA family responsive promoter and a protein operably linked to that promoter. The expression of this protein could be used to normalize the amount of protein produced in the presence and absence of compound.

5 **X. Methods of Identifying Genetic Loci in an Microbe Which Affect Resistance**

A variety of different techniques can be used to identify new genetic loci which are involved in mediating antibiotic resistance. For example, either whole cell or cell free assay systems can be employed utilizing at least one MarA family helix-turn-helix domain.

10

A. Cell-Based Assays

In one embodiment the invention provides cell based method of identifying genetic loci in an microbe which affect resistance to antibiotics. In such an assay a nucleotide sequence encoding a helix-turn-helix motif of a MarA family protein is
15 introduced into a microbe and the microbe is tested for changes in its antibiotic resistance profile, for example, by monitoring changes in growth in media containing antibiotics or by detecting a reduction in the zone of inhibition around an antibiotic disc. The ability of a MarA family protein helix-turn-helix domain to decrease antibiotic sensitivity is an indication that the microbe comprises a MarA family protein responsive
20 endogenous genetic loci which are involved in mediating antibiotic resistance.

In another embodiment, the above method can further involve assaying for changes in transcription of the genetic loci identified in the microbe. For example, a test microbe can be transformed with a vector bearing a MarA family helix-turn-helix domain. Suitable control microbes include those which lack any such heterologous
25 DNA or are transformed with a vector bearing a mutant form of a MarA family helix-turn-helix domain. The total RNA from the test and control microbes can be isolated. This can be done, for example, by making a cDNA library from both of the strains. The cDNAs from the test and control strains can then be incubated together under conditions which are favorable to hybridization. cDNAs which do not hybridize and remain single
30 stranded may be involved in mediating antibiotic resistance and can be isolated and sequenced using standard techniques.

In another example of a method by which the total mRNA from cells bearing a MarA family helix-turn-helix domain can be compared to cells with lack such a domain or bear a mutant form of such a domain, a cDNA library can be made from the total RNA of these cells. This cDNA library can be used to generate labeled probes which can
5 be used in a standard Northern blot screen. Any cDNA probes that hybridize to the mRNA of the cells comprising a MarA family helix-turn-helix domain, but not to the mRNA from control cells will be involved in mediating antibiotic resistance. Once these cDNA probes which specifically hybridize to cells comprising a MarA family helix-turn-helix domain are identified, these probes can be used to identify genes
10 involved in mediating antibiotic resistance using standard techniques.

A. Cell-Free Assays

In other embodiments of the invention, MarA family protein responsive genetic loci involved in mediating antibiotic resistance are identified using cell-free assays. In
15 one embodiment, a cell-free method of identifying such genetic loci involves contacting a nucleic acid molecule of the microbe with a MarA family protein helix-turn-helix domain and allowing complexes to form. The helix-turn-helix domain-nucleic acid molecule complexes are separated from the uncomplexed helix-turn-helix domains and the sequence of those nucleic acid molecules which can bind to a MarA family protein
20 helix-turn-helix domain can then be sequenced to identify loci involved in mediating antibiotic resistance.

For example, substantially purified MarA family protein helix-turn-helix domain polypeptide is mixed with the fragmented genomic DNA of an microbe under conditions which permit the polypeptide to bind to appropriate DNA sequences. DNA fragments to
25 which the helix-turn-helix domain has bound can be isolated using a column, filters, polyacrylamide gels, or any other methods well known to those of ordinary skill in the art. The DNA which has bound to the helix-turn-helix domain can then be released from the domain and cloned into vectors or used as probes to locate and isolate the genes to which they correspond. Any such genes can then be sequenced.

In another aspect the invention also provides for kits for identifying genetic loci in an microbe which affect resistance to compounds. Such kits comprise a nucleotide sequence encoding a MarA family protein helix-turn-helix domain and/or substantially purified MarA family protein helix-turn-helix domains and nucleotide sequence
5 encoding a mutant form of a MarA family protein helix-turn-helix domain and/or substantially purified mutant forms of a MarA family protein helix-turn-helix domain. By providing both functional MarA family protein helix-turn-helix domains and mutant forms of such domains, the subject kits provide both the test and control reagents which facilitate both optimal performance of the claimed methods and optimal interpretation of
10 results.

XI. Formulations Comprising Compounds Identified in the Instant Assays

The invention provides pharmaceutically acceptable compositions which include a therapeutically-effective amount or dose of a compound identified in any of the instant
15 assays and one or more pharmaceutically acceptable carriers (additives) and/or diluents. A composition can also include a second antimicrobial agent, e.g., an antibiotic.

As described in detail below, the pharmaceutical compositions can be formulated for administration in solid or liquid form, including those adapted for the following: (1) oral administration, for example, aqueous or non-aqueous solutions or suspensions,
20 tablets, boluses, powders, granules, pastes; (2) parenteral administration, for example, by subcutaneous, intramuscular or intravenous injection as, for example, a sterile solution or suspension; (3) topical application, for example, as a cream, ointment or spray applied to the skin; (4) intravaginally or intrarectally, for example, as a pessary, cream, foam, or suppository; or (5) aerosol, for example, as an aqueous aerosol, liposomal preparation or
25 solid particles containing the compound.

The phrase "pharmaceutically-acceptable carrier" as used herein means a 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 antiinfective agents or compounds of the invention from one organ, or
30 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

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: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) phosphate buffer solutions; and (21) other non-toxic compatible substances employed in pharmaceutical compositions. Proper fluidity can be maintained, for example, by the use of coating materials, such as 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 adjuvants such as preservatives, wetting agents, emulsifying agents and dispersing agents. Prevention of the action of microbes 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.

Pharmaceutical 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
5 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 mean modes of administration other than enteral and topical administration, usually by injection, and includes, without limitation, intravenous, intramuscular,
10 intraarterial, intrathecal, intracapsular, intraorbital, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intraarticular, subcapsular, subarachnoid, intraspinal and intrasternal injection and infusion.

The phrases "systemic administration," "administered systemically," "peripheral administration" and "administered peripherally" as used herein mean the administration
15 of a sucrose octasulfate and/or an antibacterial, drug or other material other than directly into the central nervous system, such that it enters the 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
20 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 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
25 surfaces.

Compositions can be formulated in a variety of conventional forms employed for 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,
30 ointments, salves, balms, douches, drops, troches, chewing gums, lozenges, mouthwashes, rinses.

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, hectorite, ispaghula husk, polyvinylpyrrolidone, silica and derivatives thereof, xanthan gum, kaolin, talc, starch and derivatives thereof, paraffin, 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, bicarbonate, citrate, gluconate, lactate, acetate, gluceptate or tartrate).

Such compositions can be particularly useful, for example, for treatment or prevention of an unwanted cell, e.g., vaginal *Neisseria gonorrhoeae*, or infections 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 antiinfective compounds or a pharmaceutically acceptable salt 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, 5 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.

10 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 such as benzylalkonium chloride. Alternatively, for ophthalmic uses, the compositions can be formulated in an ointment such as petrolatum. Exemplary ophthalmic 15 compositions include eye ointments, powders, solutions and the like.

Powders and sprays can contain, in addition to sucrose octasulfate and/or antibiotic or contraceptive agent(s), carriers such as lactose, talc, aluminum hydroxide, calcium silicates and polyamide powder, or mixtures of these substances. Sprays can additionally contain customary propellants, such as chlorofluorocarbons and 20 volatile unsubstituted hydrocarbons, such as butane and propane.

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, Pluronics, or 25 polyethylene glycol), proteins like serum albumin, sorbitan esters, oleic acid, 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, 30 lozenges (using a flavored basis, usually sucrose and acacia or tragacanth), powders, 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
5 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
10 agents. If desired, certain sweetening, flavoring or coloring agents may also be added.

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
15 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
20 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
25 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,
30 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 wetting agents, emulsifying and suspending agents, sweetening, flavoring, coloring,
5 perfuming and preservative agents.

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

10 Sterile injectable forms of the compositions of this invention can be aqueous or oleaginous suspension. 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
15 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
20 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,
25 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, such as Ph. Helv or similar alcohol.

The antiinfective agent or a pharmaceutically acceptable salt thereof will represent some percentage of the total dose in other dosage forms in a material forming a
30 combination product, including liquid solutions or suspensions, suppositories, douches, enemas, gels, creams, emulsions, lotions slurries, soaps, shampoos, detergents, powders,

sprays, lipsticks, foams, pastes, toothpastes, ointments, salves, balms, douches, drops, troches, lozenges, mouthwashes, rinses and others. Creams and gels for example, are typically limited by the physical chemical properties of the delivery medium to concentrations less than 20% (e.g., 200 mg/gm). For special uses, far less concentrated preparations can be prepared, (e.g., lower percent formulations for pediatric applications). For example, the pharmaceutical composition of the invention can comprise sucrose octasulfate in an amount of 0.001-99%, typically 0.01-75%, more typically 0.1-20%, especially 1-10% by weight of the total preparation. In particular, a preferred concentration thereof in the preparation is 0.5-50%, especially 0.5-25%, such as 1-10%. It can be suitably applied 1-10 times a day, depending on the type and severity of the condition to be treated or prevented.

Given the low toxicity of an antiinfective agent or a pharmaceutically acceptable salt thereof over many decades of clinical use as an anti-ulcerant [W.R. Garnett, *Clin. Pharm.* 1:307-314 (1982); R.N. Brogden et al., *Drugs* 27:194-209 (1984); D.M. McCarthy, *New Eng J Med.*, 325:1017-1025 (1991), an upper limit for the therapeutically effective dose is not a critical issue.

For prophylactic applications, the pharmaceutical composition of the invention can be applied prior to potential infection. The timing of application prior to potential infection can be optimized to maximize the prophylactic effectiveness of the compound. The timing of application will vary depending on the mode of 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. One skilled in the art will be able to determine the most appropriate time interval required to maximize prophylactic effectiveness of the compound.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. 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*
5 *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
10 published patents, cited throughout this application are hereby expressly incorporated by reference. In addition, the contents of U.S. patent 5,650,321 are also expressly incorporated by this reference.

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

15

Examples

Example 1. The identification of genetic loci in *Mycobacterium smegmatis* which are involved in antibiotic resistance.

Multidrug resistance in *Mycobacterium* is presumed to occur via the
20 accumulation of independent chromosomal mutations which affect susceptibility to individual drugs or a single pleiotropic mutation, e.g., in *mar*. In *Escherichia coli* and other *Enterobacteriaceae* multidrug resistance is generally attributed to plasmids and transposons. Still, multidrug resistance can arise via derepression of the *E. coli mar* (multiple antibiotic resistance) operon, either by mutation or exposure to inducing
25 compounds (S.P. Cohen et al, *J. Bacteriol.*, 1993, 175:1484-1492). In *Mycobacterium*, the observed relatively high frequency of multidrug resistance and the suggested relationship of inadequate treatment to the emergence of resistance (B.R. Bloom et al, *Science*, 1992, 257:10544-10642) fit with the selection of *E. coli* *Mar* mutants. The possible existence of a *mar*-like regulatory drug resistance response in *Mycobacterium*

smegmatis antimicrobial susceptibility in cells expressing the cloned *E. coli marA* gene was examined.

PCR oligonucleotide primers were used to prepare a wild-type *marA* amplicon from *E. coli* AG100 (A.M. George, et al, *J. Bacteriol.*, 1983, 155:531-540)
5 chromosomal DNA, based on the annotated sequence (S.P. Cohen, et al. *J. Bacteriol.*, 1993, 175:1484-1492). The oligonucleotide primers corresponded with nucleotide positions 1893 to 1910 and 2265 to 2282 and contained terminal *EcoRI* restriction enzyme sites to allow insertion of *marA* in frame with the *hsp60* mycobacterial heat shock promoter resident on the *E. coli-Mycobacterium* shuttle plasmid pMV261 (W.R. Jacobs, et al. *Microbiol. Immunol.*, 1990, 155:153-160). The "megaprimer" PCR method (O.H. Landt, et al., *Gene.*, 1990, 96:125-128) was used to create insertional mutants of *marA* in the center of each alpha-helical region of the putative helix-turn-helix (HTH) domain of MarA (Figure 1). These mutant *marA* genes were ligated to pMV261 and pET13a for testing in *M. smegmatis* and *E. coli*, respectively. Plasmids
10 were introduced into *E. coli* and *M. smegmatis* mc²155 by electroporation with a Gene Pulser transfection apparatus (Bio-Rad. Richmond, Calif.) and selected on kanamycin (10 or 25 µg/ml).

Cultures of *M. smegmatis* mc²155 with and without plasmids were grown at 30 or 37°C by using 7H9 or 7H10 Middlebrook medium (Difco) enriched with
20 Middlebrook Dubos albumic complex supplement (OADC), respectively, supplemented with 0.05% Tween 80 and with kanamycin (10 µg/ml) where appropriate to maintain the Kan^r plasmids. Antimicrobial susceptibilities were tested without kanamycin in 7H10-OADC antibiotic gradient plates (M.S. Curial, et al. *J. Bacteriol.*, 1982, 151:209-215) at 30 and 37°C. Tetracycline, chloramphenicol, norfloxacin, and phenazine methosulfate
25 were purchased from Sigma Chemical Co. (St. Louis, Mo.). Isoniazid, rifampin, streptomycin sulfate, and ethambutol were kindly provided by J. Crawford (Centers for Disease Control and Prevention, Atlanta. Ga.), and sparfloxacin was received from Rhone-Poulenc (Paris, France).

M. smegmatis mc²155 bearing pMV261:*marA* showed increased resistance to multiple antimicrobial agents, including rifampin, isoniazid, ethambutol, chloramphenicol, and tetracycline, compared to the microbe with vector alone (Table 3) when grown at 37°C but not at 30°C. Increased resistance to rifampin, however, was also noted at 30°C. Rifampin resistance also increased in the presence of vector alone at both temperatures, although this finding was variable. When it did occur, this was the only drug to which the vector appeared to affect parental susceptibility levels. Resistance of *M. smegmatis* to chloramphenicol increased two-fold and resistance to tetracycline increased nearly five-fold in the presence of *marA*. Ethambutol and isoniazid resistance increased 1.5- and 2.5-fold at 37°C. Little if any change in susceptibility to nalidixic acid, phenazine methosulfate, or sparfloxacin occurred; some increased susceptibility was observed for norfloxacin and streptomycin.

These changes in drug susceptibility were not seen with the *marA* gene cloned in the reverse orientation relative to the mycobacterial *hsp60* promoter. Also, introduction of *marR* cloned with the same vector by PCR methods (primer nucleotide positions 1446 to 1462 and 1864 to 1879) caused no changes in susceptibility of *M. smegmatis* to any of the compounds tested (Table 3). Strains selected for spontaneous loss of plasmids by growth in the absence of kanamycin showed a return of the wild-type susceptibility phenotype. While multidrug resistance was clearly temperature dependent, and correlated with the presence of *marA* behind the heat shock promoter, it could reflect a resistance mechanism(s) per se which functions better at 37°C than at 30°C regardless of MarA expression. Of note, however, no temperature-dependent differences in susceptibility of wild-type cells were observed with any of the agents tested (Table 3).

Example 2. Demonstration that the resistance phenotype in *M. smegmatis* was a direct result of MarA activity in the cell as demonstrated by insertional mutants targeted to the predicted helix-turn-helix domain

To obtain support for the notion that the resistance phenotype was a direct result of MarA activity in the cell, insertional mutants targeted to the first helix-turn-helix domain (HTH1) (M.N. Alekshun, et al. *Chemother*, 41:2067-2075) of the MarA protein

were constructed. Megaprimer PCR (O.H. Landt, *Gene* 96:125-128) was used to introduce an *NheI* restriction site into the centers of both the helix A (position 1989) and helix B (position 2016) regions of *marA*. A double-stranded synthetic oligonucleotide with compatible ends was ligated to the *NheI* sites to produce two distinct insertional mutants interrupting each of the two putative alpha-helical regions (Fig. 1). These mutated genes were cloned into pMV261, as described above, for susceptibility testing in *M. smegmatis* at 37°C. They were also expressed from the isopropyl-β-D-thiogalactopyranoside (IPTG)-regulated T7 promoter resident on plasmid pET13a for testing in *E. coli* BL21 (Studier et al. 1990. *Methods Enzymol.* 185:60). Insertional inactivation of MarA at either helix A or helix B of the first HTH abolished the multidrug resistance phenotype in both *E. coli* and *M. smegmatis* (Table 3 and 4).

To confirm MarA expression, Northern blotting (S.K. Goda, *Nucleic Acids Res.* 23:3357-3358) was performed with total cellular RNA isolated by the TRIzol method (Gibco BRL, Gaithersburg, Md) from mid-log-phase cells grown at 30 and 37°C in Middlebrook 7H9-OADC medium following 1 h of pretreatment with lysozyme (4 mg of Tris-EDTA, pH 8.0, per ml) at 30°C. Equal amounts of RNA separated electrophoretically in 20 mM guanidine isothiocyanate, were probed with a radiolabeled *marA* PCR product. Hybridization signals were visualized with a PhosphorImager (Molecular Dynamics, Sunnyvale, Calif.).

MarA expression was observed in cells carrying the wild-type *E. coli marA* gene but not in the host carrying vector. Northern analysis was performed with PCR-amplified *marA* probe (nucleotide primers 1910 to 1893 and 2265 to 2282). The intensity of the *marA* hybridization signal was approximately fivefold higher in cells grown at 37 than at 30°C. As expected, a hybridization signal was detected in vector carrying *marA* in the reverse orientation, since a double-stranded *marA* probe was used.

Anti-MarA antiserum was prepared with MarA purified from BL21(DE3)pLysS cells (F.W. Studier, et al., *Methods Enzymol.* 1990, 185:60-89) bearing *marA* (M.N. Alekshun, *Antimicrob. Agents Chemother.* 1997, 41:2067-2075) cloned under the control of the T7 RNA polymerase initiation signals of pET13a. After induction with IPTG for 30 minutes, rifampin was added to maximize MarA synthesis.

MarA was purified by a combination of the procedures of Li and Demple (Z. Li, et al, 1994, *Biol. Chem.*, 1994, 269:18371-18377) and Langley et al. (K.E. Langley, et al., 1987, *Euro. J. Biochem.*, 163:313-321). Anti-MarA rabbit antiserum was generated with purified MarA by Biodesign International (Kennebunk, Maine).

5 For Western analysis, cell lysates were prepared from mid-log-phase *M. smegmatis* or *E. coli* cultures by sonication in buffer (10 mM Tris-HCl, pH 8.0; 3C7O sodium dodecyl sulfate) on ice. Prior to electrophoresis, samples were treated by boiling for 5 min in sample buffer (125 mM Tris-HCl, pH 6.8; 20% glycerol; 6 mM β -mercaptoethanol; 0.05% bromphenol blue), and equivalent amounts of total protein were
10 resolved by electrophoresis in a sodium dodecyl sulfate-17.5% polyacrylamide gel electrophoresis gel. Each lane contains 15 μ g of mycobacterial protein from supernatant fractions. Proteins from *E. coli* AG100 and AG102 were used as negative and positive controls. Proteins were transferred to Immobilon-P membranes (Amersham) and analyzed by using rabbit anti-MarA antiserum and chemiluminescent detection (with a
15 kit from New England Biolabs, Beverly Mass.).

A protein band migrating to the same place as purified MarA and having the expected molecular mass (14.3 kDa) was detected in MarA-containing cells grown at both temperatures (Fig. 2B); however, considerably more MarA was produced in cells incubated at 37°C. Since small amounts of MarA were detected at 30°C (Fig. 2B), the
20 variable resistance to rifampin and the increased susceptibility phenotypes at 30°C may have been produced by relatively low cytoplasmic levels of MarA protein. By the same Western analysis, MarA was easily detected in *E. coli* and *M. smegmatis* lysates containing the mutant *marA* genes.

The mechanism of MarA-mediated multidrug resistance in *Mycobacterium* is
25 unknown. The lack of a resistance phenotype mediated by the two different expressed mutant MarA proteins suggests that the multidrug resistance observed resulted from direct transcriptional activation of cognate promoters by MarA in *M. smegmatis*. Alternatively, MarA may have acted indirectly through induction of, or interaction with, endogenous proteins that mediate the mycobacterial Mar phenotype. In both instances,
30 the multidrug resistance phenotype would have resulted from a *mar*-like regulatory system operating on other genes in this microbe. The effect, as with MarA in *E. coli*,

may be linked to activation of a yet-to-be-discovered multidrug efflux system. The presence of efflux-like proteins (J.L. Doran, et al., 1997, *Clin. Diagn. Lab. Immunol.*, 4:23-32 and J.H. Lui, et al., *J. Bacteriol.*, 1996, 178:3791-3795), along with the earlier report of the existence of mycobacterial porin proteins (S.D. Mukhopadhyay, et al., *J. Bacteriol.*, 1997, 179:6205-6207 and J.V. Trias, et al., *Science*, 1992, 258:1479-1481), indicate that, like in *E. coli*, effector proteins for *mar*-like multidrug resistance are present in *Mycobacterium*. The recently completed *Mycobacterium tuberculosis* genome sequencing project (W.J. Philipp, et al., *Proc. Natl. Acad. Sci.*, 1996, 93:3132-3137) identified at least two proteins similar to MarA. Determination of whether these elements represent an endogenous *mar*-like system in this species awaits further study.

In addition to defining a function for MarA in a heterologous genus, our results are the first direct evidence of structurally important regions of MarA. The helix-turn-helix region targeted in site-directed mutagenesis corresponds to regions in the homologous proteins AraC and XylS (M.T. Gallegos, et al., *Microbiol. Mol. Biol. Rev.*, 1997, 61:393-410), which are involved in DNA binding and transcriptional activation (A. Brunelle, et al. *J. Mol. Biol.*, 1989, 209:607-622 and M.T. Gallegos, et al., *J. Bacteriol.*, 1996, 178:6427-6434). Although the insertional mutations reported here involve significant changes to the wild-type protein, they point to the predicted helix-turn-helix domain as critical for protein function.

20

Example 3. Development of a reporter gene screening assay for identifying compounds that reduce the activity of a MarA family protein helix-turn-helix domain.

The MarA family helix-turn-helix domain described in the previous examples is expressed in *E. coli* along with an *inaA1::phoA* reporter construct, made as previously described for *inaA1::lacZ* (Martin et al. 1995. *J. Bacteriol.* 177:4176), or a *micf::lacZ* reporter gene construct. The cells containing the *phoA* reporter construct are plated on media supplemented with 5-bromo-4-chloro-3-indolyl phosphate and containing kanamycin, while the cells containing the *lacZ* marker are grown on 5-bromo-4-chloro-3-indolyl β -D-galactoside with kanamycin. Colonies which turn blue, indicating transcription of the reporter gene construct, are isolated and placed in suspension. These

30

cells are divided into two populations for treatment with each compound to be assayed: one test population to be treated with test compound and a control population to remain untreated. The appropriate population of cells are contacted with each compound to be tested. The cells are plated onto the same selective medium as indicated above.

- 5 Colonies which turn blue indicate that the compound has no effect on the ability of the MarA helix-turn-helix domain to activate transcription of the reporter gene construct, while colonies which remain white after treatment with a compound indicate that the compound reduces the activity of the MarA helix-turn-helix domain.

10 **Example 4. Development of a screening assay for identifying compounds that increase the antibiotic sensitivity of an organism bearing a MarA family protein helix-turn-helix domain.**

M. smegmatis mc²155 bearing pMV261::*marA*, which was shown above to have increased resistance to the antimicrobial agents: rifampin, isoniazid, ethambutol, chloramphenicol, and tetracycline is used to test for compounds which decrease this resistance. These cells are divided into two populations for treatment with each compound to be assayed: one test population to be treated with test compound and a control population to remain untreated. The appropriate population of cells is contacted with each compound to be tested. The treated and untreated cells are plated onto plates containing medium. Antibiotic sensitivity discs for the antibiotics listed above are placed on the plated cells and the plates are incubated. Compounds which reduce the zone of antibiotic sensitivity, i.e., show a larger zone of growth inhibition around the antibiotic discs from that seen in the cells which are not treated with compound are selected for identification.

25

Example 5. Analysis of MarA expressed transcripts using DNA "chip" (GeneChip®) technology.

To identify all MarA induced/regulated transcripts in *Escherichia coli in vivo*, DNA "chip" (GeneChip®) technology is employed. DNA computer chips containing the entire *E. coli* chromosome are available through Affymetrix (Santa Clara, CA). In brief, *E. coli* containing a *marA* expression vector is induced in order to overexpress

30

MarA *in vivo*. This treatment results in the activation of MarA regulated promoters. Total cDNA is prepared from these cells and used to probe the DNA chips. As a control, cDNA prepared from *E. coli* containing the expression vector lacking *marA* is used. This approach is also used to identify all genes expressed following exposure to
5 any compound that induces *mar* expression.

Example 6. Negative antisense control of MarA regulated loci.

Negative regulation of MarA responsive transcripts is achieved using a method similar to a previously described protocol (White et al., *Antimicrob. Agents Chemother.*
10 41:2699). *E. coli* is transformed with a vector encoding antisense-oligonucleotides complementary to 5' portions of the *marA*, *rob*, *soxS*, or other MarA family member protein transcripts following expression of these antisense-oligonucleotides *in vivo*. These structures interfere with translation of the *marA*, *rob*, *soxS*, etc. transcripts are targeted and degraded by endogenous RNaseH, an enzyme that degrades RNA-DNA
15 hybrids. A vector which encodes antisense-oligonucleotides targeted toward all of the MarA homologs in *E. coli* is designed. Transfection of this vector into *E. coli* diminishes or eliminates the host's MarA family member regulated adaptational response to many antibiotics, disinfectants, organic solvents, and/or other environmental stimuli.

20 **Example 7. Resistance Pattern of *E. coli* microbes with plasmids bearing different *E. coli mar* genes.**

Antibiotic MICs were determined for *E. coli* strains comprising different *E. coli mar* genes having mutations in the first helix-turn-helix domain. The antibiotic MICs are shown in Table 5. The A10 strain has an (Ala Ser Arg4 Ala Ser) insertion at amino
25 acid 31. The A12 strain has a substitution of Val to Ala at position 33. The B7 strain has an AlsSerAla5Ser insertion at amino acid position 40. The B9 strain has a Trp to Ala substitution at position 42 and a His to Ser substitution at position 43. The Lys14 strain has a Lys to Gln substitution at position 41. The Trp15 strain has a Trp to Ala substitution at position 42. The Phe21 strain has a Phe to Leu substitution at position
30 48. The Ser2Leu3 strain has a Ser to Ala substitution at position 29 and a Leu to Ala

substitution at position 30. The Ser2Ala strain has a Ser to Ala substitution at position 29.

Example 8. Mutations in the second alpha helix of the first helix-turn-helix

5 **domain of MarA.**

To determine the roles of the MarA HTH domains in more detail, additional mutants in these regions were devised. These mutants were tested for their ability to promote a Mar phenotype in *E. coli* and for changes in their affinity for the *mar* promoter. The mutants were synthesized in amounts comparable to the wild type MarA as determined by Western blot analysis. These mutants were defective in their ability to promote drug resistance, and exhibited lowered DNA binding affinities. Changes within the first alpha helix of the first HTH domain were less detrimental to MarA function than mutations in the second alpha helix of the first HTH domain.

15 **Example 9. Design of oligomers to create mutations in the second helix-turn-helix domain of MarA.**

As was done for the first helix-turn-helix domain of MarA, mutagenic oligomers can be designed to make mutations in the second helix-turn-helix domain of MarA. Exemplary oligomers for mutating this region of MarA are illustrated in Figure 4.

20

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.

25

TABLE 1. Some Bacterial MarA homologs^a

	Gram-negative bacteria		Gram-positive bacteria
5	<i>Escherichia coli</i>		<i>Lactobacillus helveticus</i>
	MarA (1)	<i>Kiebsiella pneumoniae</i>	U34257 (39)
	OrfR (2, 3)	RamA (27)	
	SoxS (4, 5)		<i>Azorhizobium caulinodans</i>
	AfrR (6)	<i>Haemophilus influenzae</i>	S52856 (40)
10	AraC (7)	Ya52 (28)	
	CelD (8)		<i>Streptomyces spp.</i>
	D90812 (9)	<i>Yersinia spp.</i>	U21191 (41)
	FapR (10, 11)	CafR (29)	AraL (42)
	MelR (12)	LcrF (30) or VirF (30)	
15	ORF _f375 (13, 14)		<i>Streptococcus mutans</i>
	RhaR (15, 16, 17)	<i>Providencia stuartii</i>	MsmR (43)
		AarP (31)	
	RhaS (18)		<i>Pediococcus pentosaceus</i>
	Rob (19)	<i>Pseudomonas spp.</i>	RafR (44)
20	U73857 (20)	MmsR (32)	
	XylR (21)	TmbS (33)	<i>Photobacterium leiognathi</i>
	YijO (22)	XylS (34)	LumQ (45)
		Xys1,2,3,4 (35, 36)	
	<i>Proteus vulgaris</i>		<i>Bacillus subtilis</i>
25	PqrA (23)		AdaA (46)
			YbbB (47)
	<i>Salmonella typhi-</i>		YfiF (48)
	<i>murium</i>	Cyanobacteria	YisR (49)
	MarA (24)	<i>Synechocystis spp.</i>	YzbC (50)
30	InvF (25)	LumQ (38)	
	PocR (26)	PchR (38)	

^a The smaller MarA homologs, ranging in size from 87 (U34257) to 138 (OrfR) amino acid residues, are represented in boldface. References are given in parentheses.

References for Table 1:

- (1) S.P. Cohen, et al. 1993. *J. Bacteriol.* 175:1484-1492
- 5 (2) G.M. Braus, et al. 1984. *J. Bacteriol.* 160:504-509
- (3) K. Schollmeier, et al., 1984. *J. Bacteriol.* 160:499-503
- (4) C.F. Amabile-Cuevas, et al., 1991. *Nucleic Acids Res.* 19:4479-4484
- (5) J. Wu, et al., 1991. *J. Bacteriol.* 173:2864-2871
- (6) M.K. Wolf, et al., 1990. *Infect. Immun.* 58:1124-1128
- 10 (7) C.M. Stoner, et al. 1982. *J. Mol. Biol.* 153:649-652
- (8) L.L. Parker, et al., 1990. *Genetics* 123:455-471
- (9) H. Mori, 1996. Unpublished data taken from the NCBI databases
- (10) P. Klaasen, et al., 1990. *Mol. Microbiol.* 4:1779-1783
- (11) M. Ahmed, et al., 1994. *J. Biol. Chem* 269:28506-28513
- 15 (12) C. Webster, et al., 1989. *Gene* 83:207-213
- (13) G. Plunkett, III. 1995. Unpublished
- (14) C Garcia-Martin, et al., 1992. *J. Gen. Microbiol.* 138:1109-1116
- (15) G. Plunkett, III., et al. 1993. *Nucleic Acids Res.* 21:3391-3398
- (16) C. G. Tate, et al. 1992. *J. Biol. Chem.* 267:6923-6932
- 20 (17) J.F. Tobin et al., 1987. *J. Mol. Biol.* 196:789-799
- (18) J. Nishitani, 1991. *Gene* 105:37-42
- (19) R.E. Benz, et al., 1993. *Zentralbl. Bakteriologie. Parasitenkd. Infektionskr. Hyg. Abt.*
- (20) 1 Orig. 278:187-196
- (21) M. Duncan, et al., 1996. Unpublished data
- 25 (22) H.J. Sofia, et al., 1994. *Nucleic Acids Res.* 22:2576-2586
- (23) F.R. Blattner, et al., 1993. *Nucleic Acids Res.* 21:5408-5417
- (24) H. Ishida, et al., 1995. *Antimicrob. Agents Chemother.* 39:453-457
- (25) M.C. Sulavik, et al., 1997. *J. Bacteriol.* 179:1857-1866
- (26) K. Kaniga, et al., 1994. *Mol. Microbiol.* 13:555-568
- 30 (27) J.R. Roth, et al. 1993. *J. Bacteriol.* 175:3303-3316
- (28) A.M. George, et al., 1983. *J. Bacteriol.* 155:541-548

- (29) R.D. Fleischmann, et al., 1995. *Science* 269:469-512
- (30) E.E. Galyov, et al., 1991. *FEBS Lett.* 286:79-82
- (31) N.P. Hoe, et al., 1992. *J. Bacteriol.* 174:4275-4286
- (32) G. Cornelis, et al., 1989. *J. Bacteriol.* 171:254-262
- 5 (33) D.R. Macinga, et al., 1995. *J. Bacteriol.* 177:3407-3413
- (34) M.I. Steele, et al., 1992. *J. Biol. Chem.* 267:13585-13592
- (35) G. Deho, et al., 1995. Unpublished data
- (36) N. Mermoud, et al., 1984. *EMBO J.* 3:2461-2466
- (37) S.J. Assinder, et al., 1992. *Nucleic Acids Res.* 20:5476
- 10 (38) S.J. Assinder, et al., 1993. *J. Gen. Microbiol.* 139:557-568
- (39) E.G. Dudley, et al., 1996. *J. Bacteriol.* 178:701-704
- (40) D. Geelen, et al., 1995. Unpublished data
- (41) J. Kormanec, et al., 1995. *Gene* 165:77-80
- (42) C.W. Chen, et al., 1992. *J. Bacteriol.* 174:7762-7769
- 15 (43) R.R. Russell, et al., 1992. *J. Biol. Chem.* 267:4631-4637
- (44) K.K. Leenhouts, et al., 1995. Unpublished data
- (45) J.W. Lin, et al., 1995. *Biochem. Biophys. Res. Commun.* 217:684-695
- (46) F. Morohoshi, et al. 1990. *Nucleic Acids Res.* 18:5473-5480
- (47) M. Rosenberg, et al., 1979. *Annu. Rev. Genet.* 13:319-353
- 20 (48) H. Yamamoto, et al., 1996. *Microbiology* 142:1417-1421
- (49) L.B. Bussey, et al., 1993. *J. Bacteriol.* 175:6348-6353
- (50) P.G. Quirk, et al., 1994. *Biochim. Biophys. Acta* 1186:27-34

Table 2. Bacterial strains and plasmids used in the Examples

5	<u>Strain or plasmid</u>	<u>Description</u>	<u>Reference or Source</u>
	<u>Strain</u>		
	<i>E. coli</i> AG100	Wild type	J. Bacteriol., 1983 155:531-540
	<i>E. coli</i> AG102	Mar mutant of AG100	J. Bacteriol., 1983 155:531-540
10	<i>E. coli</i> BL21	Expression strain for pET vectors	Novagen
	<i>M. smegmatis</i> MC ² 155	Electroporation competent	Mol. Microbiol, 1990, 4:1911-1919
15	<u>Plasmids</u>		
	pMV261	<i>Mycobacterium-E. coli</i> shuttle vector	Nature, 1991, 351:456-460
	pPM10	pMV261:: <i>marA</i>	This study
20	pPM10R	pMV261:: <i>marA</i> in antisense orientation	This study
	pPM11	pMV261:: <i>marR</i>	This study
	pPM1989R	pPM10 insertional mutant of helix A	This study (Fig. 1)
25	pPM2016A	pPM10 insertional mutant of helix B	This study (Fig. 1)
	pET13a	T7 expression vector	Methods Enzymol, 1990, 185:60-89
	pEC10	pET13:: <i>marA</i>	This study
	pEC1989R	pEC10 insertional mutant of helix A	This study (Fig. 1)
30	pEC2016A	pEC10 insertional mutant of helix B	This study (Fig. 1)

Table 3. Antibiotic susceptibilities of *M. smegmatis* mc²155 microbes with and without plasmids bearing different *E. coli* *mar* genes

		% Growth in antibiotic gradient ^a									
10		RIF		INH		CML		TET		ETM	
	<i>M. smegmatis</i> microbe	30°C	37°C	30°C	37°C	30°C	37°C	30°C	37°C	30°C	37°C
15	Wild type	11±1.2	11±1.1	21±1.2	20±1.0	20±0.8	22±0.9	20±1.9	19±12	40±2.2	44±2.1
	Transformants bearing plasmids:										
20	pMV261	22±1.1	45±2.6	22±1.3	20±1.0	22±0.9	23±1.1	20±1.8	20±1.3	42±2.4	44±1.5
	pPM10	68±1.6	90±2.1	15±1.5	45±2.2	13±1.1	46±1.6	15±2.1	95±3.1	38±1.8	63±2.7
	pPM10R	24±1.0	47±2.2	20±0.9	24±1.7	22±0.9	27±1.1	21±2.0	20±1.0	38±2.0	45±2.4
	pPM11	24±1.0	45±2.1	22±1.0	23±1.2	22±1.0	27±1.0	22±2.2	20±1.3	40±2.0	46±2.4
	pPM1989R	ND	10±1.0	ND	18±2.1	ND	26±1.4	ND	20±2.0	ND	48±2.2
25	pPM2016A	ND	13±0.8	ND	22±1.0	ND	27±2.2	ND	22±1.7	ND	50±2.4

^aThe antibiotics were tested in gradient plates at various concentrations (in micrograms per milliliter) as follows: rifampin (RIF), 150; isoniazid (INH), 3.5; chloramphenicol (CML), 30; tetracycline (TET), 0.3; and ethambutol (ETM), 2.0. Values are means ± standard deviations of experiments performed in triplicate. ND, not determined.

Table 4. Antibiotic susceptibilities of *E. coli* microbes with plasmids bearing different *E. coli mar* genes

5

% Growth in antibiotic gradient^a

Plasmid borne by	CML	TET	AMP	NAL
10 <i>E. coli</i> transformant				
pET13a	20±1.1	33±1.0	2±1.1	17±0.8
pEC10	100	100	11±2.0	100
pEC1989R	20±1.6	33±1.4	2±1.0	17±1.2
15 style="padding-left: 20px;">pEC2016A	20±1.3	33±1.4	2±1.7	17±1.1

^a The antibiotics were tested in gradient plates at various concentrations (in micrograms per milliliter) as follows: CML, 2.0; TET, 1.0; ampicillin (AMP), 3.0; and nalidixic acid (NAL), 1.0 Values are means ± standard deviations of experiments performed in

20 triplicate.

Table 5. Resistance Pattern of MarA Mutants in 1st alpha Helix of First Helix-Turn-Helix Domain

5

Strain	Tet	Chlor	Nal Ac	Rif	Cipro	Nor	Amp	Genta	Ceph
A ₁₀ Ala ₁₄	0.75±.25	1.0	0.38±.12	10.7±2.3	<0.002	<0.016	0.74±.02	1.0	1.5
Ser ₂ Ala	1.25±.43	3.3±.57	1.1±.38	>32	<0.002	<0.016	2.4±.27	1.0	3.0
pET13A	0.71±.31	1.0	0.23±.03	13.3±4.6	<0.002	<0.016	0.76±.11	0.75	1.5
A10	0.75±.25	1.08±.38	0.38±.13	12±4	<0.002	<0.016	0.8±.07	0.75	1.5
A12	0.75±.25	1.5±0.5	0.54±.18	16	<0.002	<0.016	1.3±0.18	0.75	0.094
B7	0.58±.14	1.3±.28	0.33±.14	17.3±6.1	<0.002	<0.016	0.9±.24	0.75	3
B9	0.92±.52	1.2±.28	0.42±.06	15±2.3	<0.002	<0.016	0.84±.26	4	1.5
Lys14	0.67±.14	1.2±.28	0.38±.12	13.3±2.3	<0.002	<0.016	0.9±.23	3	2
Ser2Leu3	0.75	1.0	0.33±.07	14.6±2.3	<0.002	<0.016	0.75±.03	0.75	1.5
Trp15	0.58±.14	0.83±.28	0.25±.12	10.6±2.3	<0.002	<0.016	0.84±.05	0.75	3
A10Ala 6	0.92±.52	0.83±.14	0.37±.12	12±4	<0.002	<0.016	0.75±.03	0.75	1.5
Phe21	0.66±.14	0.92±.14	0.33±.14	10.6±2.3	<0.002	<0.016	0.78±.04	1.0	2
MarA	1.2±.28	2.5±.86	0.96±.56	>32	<0.002	<0.016	3.0±.38	1.0	3
BL21	0.75	0.66±.14	0.23±.03	6.7±1.1	<0.002	<0.016	<0.125	1.0	1.5

Data shown are the Avg ± SD (3 replicates).

What is Claimed is:

1. A method for identifying an antiinfective compound which affects the activity of a MarA family helix-turn-helix domain, comprising:
 - 5 contacting a polypeptide comprising a Mar A family helix-turn-helix domain derived from a MarA family protein with a compound under conditions which allow interaction of the compound with the polypeptide such that a complex is formed; and
 - measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain as an indication of whether the compound is an antiinfective
 - 10 compound.

2. A method for identifying an antiinfective compound which affects the activity of a MarA family helix-turn-helix domain, comprising:
 - contacting a cell expressing a Mar A family helix-turn-helix domain polypeptide
 - 15 derived from a MarA family protein with a compound under conditions which allow interaction of the compound with the polypeptide; and
 - measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain polypeptide as an indication of whether the compound is an antiinfective compound.
 - 20

3. The method of claim 2, wherein the step of measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain comprises detecting the ability of the complex to activate transcription from a MarA family member responsive promoter.
- 25

4. The method of claim 3, wherein the Mar A responsive promoter is selected from the group consisting of *marO*, *micF*, and *fumC*

5. The method of claim 3, wherein the Mar A responsive promoter is linked to a
- 30 reporter gene.

6. The method of claim 5, wherein the reporter gene is selected from the group consisting of *lacZ*, *phoA*, or green fluorescence protein.
7. The method of claim 5, wherein the step of measuring comprises measuring the amount of reporter gene product.
8. The method of claim 3, wherein the step of measuring comprises measuring the amount of RNA produced by the cell.
9. The method of claim 3, wherein the step of measuring comprises measuring the amount of a protein produced by the cell.
10. The method of claim 9, wherein the step of measuring comprises using an antibody against a protein produced by the cell.
11. A method for identifying an antiinfective compound which affects the activity of a MarA family helix-turn-helix domain, comprising:
contacting a polypeptide comprising a Mar A family helix-turn-helix domain derived from a MarA family protein with a compound in a cell-free system under conditions which allow interaction of the compound with the polypeptide such that a complex is formed; and
measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain as an indication of whether the compound is an antiinfective compound.
12. The method of claim 11, wherein the MarA family helix-turn-helix domain is an isolated polypeptide and the step of measuring the ability of the compound to affect the activity of a MarA family helix-turn-helix domain comprises measuring the ability of the complex to bind to DNA.

13. The method of claim 11, wherein said method comprises screening a library of bacteriophage displaying on their surface a MarA family helix-turn-helix domain polypeptide, said polypeptide sequence being encoded by a nucleic acid contained within the bacteriophage, for ability to bind a compound to obtain those compounds
- 5 having affinity for the helix-turn-helix domain, said method comprising:
- contacting the phage which display the helix-turn-helix domain with a sample of a library of compounds so that the helix-turn-helix domain can interact with and form a complex with any compound having an affinity for the helix-turn-helix domain;
 - 10 contacting the complex of the helix-turn-helix domain and bound compound with an agent that dissociates the bacteriophage from the compound; and
 - identifying the compounds that bound to the helix-turn-helix domain.
14. A method for screening a library of bacteriophage displaying on their surface a plurality of polypeptide sequences, each polypeptide sequence being encoded by a
- 15 nucleic acid contained within the bacteriophage, for ability to bind an immobilized MarA family helix-turn-helix domain, to obtain those polypeptides having affinity for the helix-turn-helix domain, said method comprising
- 20 contacting the immobilized helix-turn-helix domain with a sample of the library of bacteriophage so that the helix-turn-helix domain can interact with the different polypeptide sequences and bind those having affinity for the helix-turn-helix domain to form a set of complexes consisting of immobilized helix-turn-helix domain and bound bacteriophage;
 - separating the complexes from bacteriophage which have not formed the complex;
 - 25 contacting the complexes of the helix-turn-helix domain and bound bacteriophage with an agent that dissociates the bound bacteriophage from the complexes; and
 - isolating the dissociated bacteriophage and obtaining the sequence of the nucleic acid encoding the displayed polypeptide, so that amino acid sequences of displayed
 - 30 polypeptides with affinity for helix-turn-helix domain are obtained.

15. The method of claim 1, wherein said polypeptide comprises the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein from which it is derived.
- 5 16. The method of claim 1, wherein said polypeptide comprises the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein from which it is derived.
- 10 17. The method of claim 1, wherein said polypeptide consists of the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein from which it is derived.
- 15 18. The method of claim 1, wherein said polypeptide consists of the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein from which it is derived.
19. The method of claim 1, wherein the MarA family helix-turn-helix domain is derived from a protein selected from the group consisting of: MarA, RamA, AarP, Rob, SoxS, and PqrA.
- 20 20. The method of claim 1, wherein the compound increases antibiotic susceptibility.
21. The method of claim 1, wherein the compound reduces infectivity or virulence.
- 25 22. The method of claim 1, wherein the compound is effective against Gram negative bacteria.
23. The method of claim 1, wherein the compound is effective against Gram positive bacteria.

24. The method of claim 23, wherein the Gram positive bacteria are from a genus selected from the group consisting of: *Enterococcus*, *Staphylococcus*, *Clostridium* and *Streptococcus*.
- 5 25. The method of claim 1, wherein the compound is effective against bacteria from the family *Enterobacteriaceae*.
26. The method of claim 1, wherein the compound is effective against a bacteria of a genus selected from the group consisting of: *Escherichia*, *Proteus*, *Klebsiella*,
10 *Providencia*, *Enterobacter*, *Burkholderia*, *Pseudomonas*, *Aeromonas*, *Acinetobacter* and *Mycobacteria*.
27. A cell based method of identifying genetic loci in an microbe which affect antibiotic resistance comprising introducing into said microbe a nucleotide sequence
15 encoding a helix-turn-helix motif of a MarA family protein and assaying for changes in the antibiotic resistance profile of said microbe.
28. The method of claim 27, further comprising assaying for changes in transcription of genetic loci of said microbe.
20
29. The method of claim 28, further comprising identifying proteins which are present in different amounts in resistant and susceptible microbes.
30. The method of claim 29, further comprising identifying the genes which encode
25 said proteins.
31. The method of claim 27, wherein said polypeptide comprises the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein from which it is derived.
30

32. The method of claim 27, wherein said polypeptide comprises the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein from which it is derived.
- 5 33. The method of claim 27, wherein said polypeptide consists of the helix-turn-helix domain most proximal to the carboxy terminal end of the MarA family protein from which it is derived.
34. The method of claim 27, wherein said polypeptide consists of the helix-turn-
10 helix domain most proximal to the amino terminal end of the MarA family protein from which it is derived.
35. The method of claim 27, wherein the MarA family helix-turn-helix domain is derived from a protein selected from the group consisting of: MarA, RamA, AarP, Rob,
15 SoxS, and PqrA.
36. The method of claim 27, wherein said microbe is a bacterium.
37. The method of claim 20, wherein the antibiotic is selected from the group
20 consisting of: tetracycline, fluoroquinolones, chloramphenicol, penicillins, cephalosporins, puromycin, nalidixic acid, and rifampin.
38. The method of claim 20, wherein the antibiotic is a disinfectant, antiseptic, or surface delivered antibacterial compound.
25
39. The method of claim 20, wherein the antibiotic is an antifungal.
40. The method of claim 20, wherein the antibiotic is an antiparasitic.

41. A cell-free method of identifying genetic loci in an microbe which affect resistance to antibiotics comprising contacting a nucleic acid molecule of said microbe with a MarA family protein helix-turn-helix domain and allowing complexes to form; separating the nucleic acid molecule which has formed a complex with a helix-turn-helix domain from the helix-turn-helix domain; and identifying the sequence of those nucleic acid molecules which can bind to a MarA family protein helix-turn-helix domain.
42. The method of claim 40, wherein said polypeptide comprises the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein from which it is derived.
43. The method of claim 40, wherein said polypeptide comprises the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein from which it is derived.
44. The method of claim 40, wherein said polypeptide consists of the helix-turn-helix domain most proximal to the carboxy terminal end of the MarA family protein from which it is derived.
45. The method of claim 40, wherein said polypeptide consists of the helix-turn-helix domain most proximal to the amino terminal end of the MarA family protein from which it is derived.
46. The method of claim 40, wherein the MarA family helix-turn-helix domain is derived from a protein selected from the group consisting of: MarA, RamA, AarP, Rob, SoxS, and PqrA.
47. The method of claim 40, wherein said microbe is a bacterium.
48. The method of claim 1, wherein the compound is derived from a library of small molecules.

49. The method of claim 1, wherein the compound is a nucleic acid molecule.
50. The method of claim 49, wherein the compound is an antisense or sense
5 oligonucleotide.
51. The method of claim 1, wherein the compound is a naturally occurring small organic molecule.
- 10 52. An isolated or recombinant mutant MarA family helix-turn-helix domain comprising a mutation in the helix-turn-helix domain most proximal to the carboxy terminus of the MarA family protein molecule.
- 15 53. An isolated or recombinant mutant MarA family helix-turn-helix domain comprising a mutation in the helix-turn-helix domain most proximal to the amino terminus of the MarA family protein molecule.
- 20 54. A kit for identifying genetic loci in an microbe which affect resistance to compounds comprising a nucleotide sequence encoding a naturally occurring helix-turn-helix domain of a MarA family protein and mutant, inactive form of a MarA family protein helix-turn-helix domain.

A. NheI
 CTAGCGCGGCGGCGGCGG
Mutagenic | | | | | | | | | | | | | | | |
 GCGCCGCCGCCGCCGATC
 NheI

B. Helix ASRRRRAS
 Y
Wild type HTH PLSLEKVSERSGYSKWHL
 ^
Helix B ASAAAAAS

Figure 1

1	marA (30-107)	LEKVERS	SKWHLQRMFK	KETGSHLQGY	IRSRKMTEIA	QKL...KESN
	RHAR_SALTY (19-96)	LDAFCQEQC	SERVLRAQFR	AQTGMTINQY	LRQVRICHAQ	YLL...QHSP
	RHAR_ECOLI (225-302)	LDKFCDEASC	SERVLRQQFR	QQTGMTINQY	LRQVRICHAQ	YLL...QHSP
	RHAS_ECOLI (190-267)	WDAVADQFSL	SLRTLHRQLK	QQTGLTPQRY	LNRLRLMKAR	HLL...RHSE
	RHAS_SALTY (189-266)	WEAVAEQFSL	SLRTLHRQLK	QQTGLTPQRY	LNRLRLIKAR	HLL...RHSD
	EXSA_PSEAE (207-284)	LSDFSREFGM	GLTTFKELFG	SVYGVSPRAW	ISERRILYAH	QLL...LNSD
	LCRF_YERPE (183-260)	LSKFAREFGM	GLTTFKELFG	TVYGISPRAW	ISERRILYAH	QLL...LNGK
	VIRF_YEREN (183-260)	LSKFAREFGM	GLTTFKELFG	TVYGISPRAW	ISERRILYAH	QLL...LNGK
	ARAC_CITFR (197-274)	IASVAQHVCV	SPSRLSHLFR	QQLGISVLSW	REDQRISQAK	LLL...STTR
	ARAC_ECOLI (197-274)	IASVAQHVCV	SPSRLSHLFR	QQLGISVLSW	REDQRISQAK	LLL...STTR
	ARAC_SALTY (197-274)	IASVAQHVCV	SPSRLSHLFR	QQLGISVLSW	REDQRISQAK	LLL...STTR
	ARAC_ERWCH (202-279)	IDEVARHVCL	SPSRLAHLFR	EQVGINIORW	REDQRVIRAK	LLL...QTTQ
	POCR_SALTY (2 1-288)	LEDVASHVYL	SPYFYSKLFK	KYQGIGFNAW	VNRORMVSAR	ELL...CHSD
	YIJO_ECOLI (188-256)	RESVAQAFYI	SPNYLSHLFQ	KTGAIGFNEY	LNHTRLEHAK	TLL...KGYD
	YBBB_BACSU (182-259)	LAQLSQMAGI	SAKHYSSEFK	KWTGQSVTEF	ITKTRITKAK	RIM...AKSN
	YFIF_BACSU (208-284)	LTDVASHFHI	SGRHLSRLEA	AELGVSYSEF	VQNEKINKAA	ELL...KSTN
	LUMQ_PHOLE (164-240)	VAELESSVAF	ACOSFYALFK	SOMGITPHOY	VLRKRDLAK	QLI...AERQ
	ARAL_STRAT (218-295)	VAELASAAA	SRSILAARFK	ATVQGPLEY	LTRWRIELTA	RQL...REGS
	ARAL_STRLI (218-295)	VAELASAAA	SRSILAARFK	ATVQGPLEY	LTRWRIELTA	RQL...REGN
	YKGD_ECOLI (196-273)	VESLASIAHM	SRASFAQLFR	DVSGTTPLAY	LTKLRLQIAA	QMF...SRET
	YA52_HAEIN (213-290)	IEQLAELATM	SRANFIRIFQ	QHIGMSPGRF	LTKVRLQSA	FLL...KQSQ
	MMSR_PSEAE (217-294)	LERLA AFCNL	SKFHVSRYK	AITGRTPIQH	FLHLKIEYAC	QLL...DSSD
	MMSR_STRMU (192-269)	VNDIAKKLNL	SRSYLYKIFR	KSTNLSIKEY	ILQVRMKRSQ	YLL...ENPK
	RAFR_PEDPE (192-269)	IMDLCHYLNL	SRSYLYTLFK	THANTSPQKL	LTCLRLEDK	QRL...STSN
	YIDL_ECOLI (213-290)	VADMAATIPC	SEAWLRRLLFL	RYTGKTPKEY	YLDARLDLAL	SLL...KQCG
	YCGK_ALTCA (83-158)	IDTVAFSVGV	SRSYLVKQFK	LATNKTINNR	IEEVRIEQAK	KVL...LKK.
	KYLR_ECOLI (304-381)	VDQVLDVAVGI	SRSNLEKRFK	EEVGETIHAM	IHAEKLEKAR	SLL...ISTT
	XYLR_HAEIN (304-381)	VGQVLDHLET	SRSNLEQRFK	NEMNKTIHQV	IHEEKISRAK	NLL...QQTID
	ADAA_BACSU (118-195)	LESLADICHG	SPYHMRTFK	KIKGITLVEY	IQQVRVHAAK	KYL...IQTN
	CELD_ECOLI (192-269)	LENMVALSAK	SQEYLTRATQ	RYYGKTPMQI	INEIRINFAK	KQL...EMTN

Figure 2A

MEIR_ECOLI (210-287)	INDVAEHVKL	NANYAMGIFQ	RVMQLTMKQY	ITAMRINHVR	ALL...SDTD
PCHR_PSEAE (217-293)	LDTLASRVGM	NPRKLTAGFR	KVFGASVFGY	LQEYRLREAH	RML...CDEE
MARA_SALTY (30-107)	LEKVSERSGY	SKWHLQRMFK	KETGSHLGQY	IRSRKMTEIA	QKL...KESN
MARA_ECOLI (30-107)	LEKVSERSGY	SKWHLQRMFK	KETGSHLGQY	IRSRKMTEIA	QKL...KESN
PQRA_PROVU (25-102)	IDTIANKSGY	SKWHLQRIEK	DFKGCITLGEY	VRKRRLLAEA	KSL...QEKD
AARP_PROST (38-115)	LDDIAQHSYG	TKWHLQRVFR	KIVGMPLGEY	IRRRRICEAA	KEL...QTTN
ROB_ECOLI (24-101)	LDNVAAKAGY	SKWHLQRMFK	DVTGHAIGAY	IRARRLSKSA	VAL...RLTA
TETD_ECOLI (47-124)	LDDVANKAGY	TKWYFQRLFK	KVTGVTLASY	IRARRLTKAA	VEL...RLTK
SOXS_SALTY (23-100)	IDVVAKKSGY	SKWYLQRMFR	TVTHQTLGEY	IRQRRLLAA	VEL...RTTE
SOXS_ECOLI (23-100)	IDVVAKKSGY	SKWYLQRMFR	TVTHQTLGDI	IRQRRLLAA	VEL...RTTE
YKGA_ECOLI (35-111)	IEDIAQKSGY	SRNIQLLFR	NFMHVPLGEY	IRKRRLCRAA	ILV...RLTA
CAFR_YERPE (25-102)	IDCLVLYSGF	SRRYLQISFK	EYVGMPIGTY	IRVRRASRAA	ALL...RLTR
RAMA_ENTCL (25-101)	IEDIARHAGY	SKWHLQRLFL	QYKGESLGRY	IRERKLLAA	RDL...RESD
RAMA_KLEPN (25-86)	IEDIARHAGY	SKWHLQRLFL	QYKGESLGRY	IRERKLLAA	RDL...RDTD
Y4FK_RHISN (334-412)	IEDLAAAARC	TPRALQRMFR	TYRGGSPMSV	LCNYRLAAAH	GAI...KAGRA
HRPB_BURSO (392-472)	TREVAAHINV	TERALQLAFK	SAVGMSPSSV	IRRMRLLEGIR	SDLLDSEPNP
THCR_RHOSO (243-323)	VAQVARNVGV	SVRSLVQVGFQ	NSLGTTPMRQ	LKIRIMQKAR	KDLLLRADPAS
XYS1_PSEPU (230-310)	LERLAELAMM	SPRSLYNLFE	KHAGTTPKNY	IRNRKLESIR	ACLNDPSPANV
XYS1_PSEPU (230-310)	LERLAELAMM	SPRSLYNLFE	KHAGTTPKNY	IRNRKLESIR	ACLNDPSPANV
XYS3_PSEPU (230-310)	LERLAELAMM	SPRSLYNLFE	KHAGTTPKNY	IRNRKLESIR	ARLSDPNANV
XYS4_PSEPU (230-310)	LERLAELAMM	SPRSLYNLFE	KHAGTTPKNY	IRNRKLECI	ARLSDPNANV
XYS2_PSEPU (55-135)	LEQLAELALM	SPRSLYTMFE	KHTGTPMNY	IRNRKLECI	ARLSDPNANV
YFEG_ECOLI (259-339)	VLDLCNQLHV	SRRTLQNRFH	AILGIGNAW	IRNRKLECI	ARLSDPNANV
V38K_MYCTU (255-328)	...IADQLDM	HPRTLQRRLA	AE.GLRCHDL	IERERRAQAA	RYL...AQPG
YBCM_ECOLI (184-256)	...VARYLYI	SVSTLHRRLA	SE.GVVSQFI	LDDVRLNNAL	SAI...QTTV
ADIV_ECOLI (165-241)	LSMVASCLCL	SPSLLKKKLK	SE.NTSYSQI	ITTCRMRVAV	NEL...MMDG
ENVY_ECOLI (165-241)	LRIVASSLCL	SPSLLKKKLK	NE.NTSYSQI	VTECRMRYAV	QML...LMDN
YHIX_ECOLI (161-237)	LARIASELLM	SPSLLKKKLR	EE.ETSYSQI	LTECRMORAL	QLI...VIHG
APPY_ECOLI (149-225)	LKDIAELIYT	SESLIKKKLR	DE.GTSFTEI	LRDTRMRYAK	KLI...TSNS

Figure 2A Continued

YHIW_ECOLI (155-231)	LRDIAERMYT	SESLIKKKLQ	DE.NTCFSKI	LLASRSMAR	RLL...ELRQ
URER_ECOLI (187-263)	LDDVAKALFT	TPSTLRRHLN	RE.GVSFRQL	LLDVRMGMAL	NYL...TFSN
URER_PROMI (187-263)	LDDVAKALYT	TPSTLRRHLN	KE.GVSFCQL	LLDVRIPIAL	NYL...TFSN
AGGR_ECOLI (180-256)	LAIIADEFNV	SEITIRKRLE	SE.YITFNQI	LMQSRMSKAA	LLL...LDNS
CFAD_ECOLI (180-256)	LGIIAADDANA	SEITIRKRLE	SE.YITFNQI	LMQSRMSKAA	LLL...LDNS
RNS_ECOLI (180-256)	LGIIAADDVVA	SEITIRKRLE	SE.YITFNQI	LMQSRMSKAA	LLL...LDNS
CSVR_ECOLI (182-258)	LAIIADVFNV	SEITIRKRLE	SE.DTNFNQI	LMQSRMSKAA	LLL...LENS
FAPR_ECOLI (170-246)	LSDIAEEMHI	SEISVRKRLE	QE.CLNFNQL	ILDVRMNQAA	KFI...IRSD
VIRF_SHIDY (177-253)	LSDISNNLNL	SEIAVRKRLE	SE.KLTFQOI	LLDIRMHAA	KLL...INSQ
PERA_ECOLI (184-260)	LGDVSSSMFM	SDSCLRKQLN	KE.NLTFKKI	MLDIKMKHAS	LFL...RTD
TCPN_VIBCH (188-264)	WADICGELRT	NRMILKKELE	SR.GVKFREL	INSIRISYSI	SIM...KTGE
FEAR_ECOLI (219-293)	...IAGETGM	SVRSLYRMFA	DK.GLVVAQY	IRNRRLDFCA	DAI.RHAADD
YISR_BACSU (113-183)ALHY	HQDYVSRMCQ	QVLGVTPAQY	TNRVRMTEAK	RLO...SSTN
Consensus	a.....s.....l.....f.....g.....r.....a.....l.....			

Figure 2B

1	marA (30-107)	EPILYLAERY	GFESQOTLTR	TFKNYFDVPP	H
	RHAR_SALTY(19-96)	LMISEISMQC	GFEDSNYFSV	VFTRETGMT	S
	RHAR_ECOLI(225-302)	LLISDISTEC	GFEDSNYFSB	VFTRETGMT	S
	RHAS_ECOLI(190-267)	ASVTDIAYRC	GFSDSNHFST	LFRREFNWS	R
	RHAS_SALTY(189-266)	HSVTEIAYRC	GFSDSNHFST	LFRREFNWS	R
	EXSA_PSEAE(207-284)	MSIVDIAMEA	GFSSQSYFTQ	SYRRRFGCT	S
	LCRF_YERPE(183-260)	MSIVDIAMEA	GFSSQSYFTQ	SYRRRFGCT	S
	VIRF_YEREN(183-260)	MSIVDIAMEA	GFSSQSYFTQ	SYRRRFGCT	S
	ARAC_CITFR(197-274)	MPIATVGRNV	GFDDQLYFSR	VFKKCTGASP	S
	ARAC_ECOLI(197-274)	MPIATVGRNV	GFDDQLYFSR	VFKKCTGASP	S
	ARAC_SALTY(197-274)	MPIATVGRNV	GFDDQLYFSR	VFKKCTGASP	S
	ARAC_ERWCH(202-279)	ESIANIGRVV	GYDDQLYFSR	VFRKRVGVSP	S
	POCR_SALTY(211-288)	WSIASIARNL	GFSDSNYFCR	LFRKNTERSP	S
	YIJO_ECOLI(188-265)	LKVKEVAHAC	GFVDSNYFCR	LFRKNTERSP	S
	YBBB_BACSU(182-259)	CKLKEIAHQ	GYQDEFYFSR	IFKKYTGCS	T
	YFIF_BACSU(208-284)	LSIKEIAEEI	GFS.VHYFTR	VFSAKIGSS	G
	LUMQ_PHOLE(164-240)	KPLSQVAQLC	GFSSQSSFSQ	AFRRLYGMSP	G
	ARAL_STRAT(218-295)	APLAAIAHSV	GYGSESALS	AFKRVLGMNP	G
	ARAL_STRLI(218-295)	ATLASIAHSV	GYGSESALS	AFKRVLGMPP	G
	YKGD_ECOLI(196-273)	LPVVVIAESV	GYASESFHK	AFVREFGCT	G
	YA52_HAEIN(213-290)	QSVLAIALEV	GYQSEAHFCK	VFKNYQLSP	S
	MMSR_PSEAE(217-294)	QSVARVGOAV	GYDDSYFYSR	LFSKVMGLSP	S
	MMSR_STRMU(192-269)	LSIAEISNSV	GFSDSLAFSK	AFKNYFGKSP	S
	RAFR_PEDPE(192-269)	NSVQSIANMY	GYKDSFTFSK	AFKRYSGASP	S
	YIDL_ECOLI(213-290)	NSVGEVADTL	NFFDSFHFSK	AFKHKFGYAP	S
	YCGK_ALTCA(83-158)	.SVTETAYEV	GFNNSNYFAT	VFKKRTNYTP	K
	XYLR_ECOLI(304-381)	LSINEISQMC	GYPSLQYFYS	VFKKAYDTTP	K
	XYLR_HAEIN(304-381)	ISIKEITEIC	GYPSIQYFYS	VFKKEFEMTP	K
	ADAA_BACSU(118-195)	KAIGDIAICV	GIANAPYFIT	LFKKTGTQTP	A

Figure 2C

CELD_ECOLI (192-269)	YSVTDIAFEA	GYSSPSLFIK	TFKKLTSFTP	K
MELR_ECOLI (210-287)	KSILDIALTA	GFRSSSRFYS	TFGKYVGMSP	Q
PCHR_PSEAE (217-293)	ANVSTVAYRV	GY.SPAHFSI	AFRKRYGISP	S
MARA_SALTY (30-107)	EPIOYAERY	GFESQQLTR	TFKNYFDVPP	H
MARA_ECOLI (30-107)	EPIOYAERY	GFESQQLTR	TFKNYFDVPP	H
PQRA_PROVU (25-102)	MSILDIALMY	GFSSQATFTR	IFKKHFNTP	A
AARP_PROST (38-115)	LQVIDIALKY	QFDSQQSEFAK	RFKAYLGISP	S
ROB_ECOLI (24-101)	RPILDIALQY	RFDSQQTFTR	AFKKQFAQTP	A
TETD_ECOLI (47-124)	KTILEIALKY	QFDSQQSFTR	RFKYIFKVT	S
SOXS_SALTY (23-100)	RPIFDIAMDL	GYVSQQTFSR	VRRREFDRT	S
SOXS_ECOLI (23-100)	RPIFDIAMDL	GYVSQQTFSR	VRRREFDRT	S
YKGA_ECOLI (35-111)	KSMLDIALSL	HFDSQQSEFSR	EFKKLFGCSP	.
CAFR_YERPE (25-102)	LTIIIEISAKL	FYDSQQTFTR	EFKKIFGYTP	R
RAMA_ENTCL (25-101)	ERVYEICLRY	GFESQQTFTR	IFTRTFHQPP	.
RAMA_KLEPN (25-86)	QRVYDICKLY	GFDSQ.....
Y4FK_RHISN (334-412)	GSITELALNL	QFSNPGRFVS	LYKSAYGLSP	S
HRPB_BURSO (392-472)	SNIIDTASRW	GIRRSALVK	GyrKQFNEAP	S
THCR_RHOSO (243-323)	EGVTEIAQRW	GFLHVGRFAG	EYKQTFGVSP	S
XYS1_PSEPU (230-310)	RSITEIALDY	GFLHLGRFAE	NYRSAFGELP	S
XYS1_PSEPU (230-310)	RSITEIALDY	GFLHLGRFAE	NYRSAFGELP	S
XYS3_PSEPU (230-310)	RSVTEMALDY	GGFHTGRFAE	NYRSTFGELP	S
XYS4_PSEPU (230-310)	RSVTEMALDY	GGFHTGRFAE	NYRSTFGELP	S
XYS2_PSEPU (55-135)	RSITEVALDY	GFLHLGRFAE	KYRSTFGELP	S
YFEG_ECOLI (259-339)	MTVKDAAMQW	GFWHLGQFAT	DYQQLFSEKP	S

Figure 2C Continued

```

V38K_MYCTU(255-328)
YBCM_ECOLI(184-256)
ADIY_ECOLI(165-241)
ENVY_ECOLI(165-241)
YHIX_ECOLI(161-237)
APPY_ECOLI(149-225)
YHIW_ECOLI(155-231)
URER_ECOLI(187-263)
URER_PROMI(187-263)
AGGR_ECOLI(180-256)
CFAD_ECOLI(180-256)
RNS_ECOLI(180-256)
CSVR_ECOLI(182-258)
FAPR_ECOLI(170-246)
VIRF_SHIDY(177-253)
PERA_ECOLI(184-260)
TCPN_VIBCH(188-264)
FEAR_ECOLI(219-293)
YISR_BACSU(113-183)
..... Consensus ..
1  L Y L S Q I A V L L G Y S E Q S A R N R S C R R W F G M T P R
31 K P I S E I A R E N G Y K C P S R F T E R E H N R F N I T P .
K N I S Q V S Q S C G Y N S T S Y F I S V F K D F Y G M T P L
K N I T Q V A Q L C G Y S T S Y F I S V F K A F Y G L T P L
F S I K R V A V S C G Y H S V S Y F I Y V E R N Y Y G M T P T
Y S I N V V A Q K C G Y N S T S Y F I C A F K D Y Y G V T P S
I P L H T I A E K C G Y S T S Y F I N T F R Q Y Y G V T P H
Y S V F Q I S H R C G F G S N A Y F C D V F K R K Y N M T P S
Y S V F Q I S H R C G F G S N A Y F C D A F K R K Y G M T P S
Y Q I S Q I S N M I G F S T S Y F I R L F V K H F G I T P K
Y Q I S Q I S N M I G I S S A S Y F I R V E N K H Y G V T P K
Y Q I S Q I S N M I G I S S A S Y F I R I F N K H Y G V T P K
Y Q I S Q I S N M I G I S S A S Y F I R I F N K H F G V T R S
H Q I G M I A S L V G Y T S V S Y F I K T F K E Y Y G V T P K
S Y I N D V S R L I G I S S P S Y F I R K F N E Y Y G I T P K
K N I D E I S C L V G F N S T S Y F I K V F K E Y Y N T T P K
F K I K Q I A Y Q S G F A S V S N F S T V F K S T M N V A P S
E K L A G I G F H W G F S D Q S H F S T V F K Q R F G M T P .
D K M G V I A E T V G M E D P T Y F S K L F K Q I E G I S P .
.. i...ia.....gf.s....f... .k....g...P

```

FIGURE 2D

MarA (HTH) Homology Alignment

MarA	28	LSLEKVSERSGYSKWHLQRMFKKETGHSLGQYIRSRKMTETIAQKL	72
AarP	36	LSLDDIAQHSGYTKWHLQRVFRKIVGMPLGEYIRRRRICEAAKEL	80
AfrR	23	IDINALVDYSGYSRRYLQLLFKENIGVTLGKYIQLRRITRAAILL	67
AraC	195	FDIASVAQHVCLSPSRSLSHLFRQQLGISVLSWREDQRISQAKLLL	239
CafR	23	INIDCLVLYSGFSRRYLQISFKEYVGMPIGTYIRVRRASRAAALL	67
LumQ	162	ISVAELSSVAFLAQSQFYALFKSQMGITPHQYVLRKRDLAKQLI	306
MelR	208	LTINDVAEHVKNLANYAMGIFQRMQLTMKQYITAMRINHVRALL	252
MmsR	215	LELERLAAFCNLSKFHFVSRYKAITGRTPIQHFLHLKIEYACQLL	259
MsmR	190	LRVNDIAKKNLSRSYLYKIFRKSTNLSIKEYILQVRMKRSQYLL	234
ORF_f375	284	LSVEQLAAEANMSVSFAFHNFKSVTSTSPLOQYLKNYRLHKARMMI	328
PocR	135	LRLEDVASHVYLSPPYFYSKLFKQYQIGFNWVNRQRMVSARELL	179
PqrA	23	IKIDTIANKSGYSKWHLQRIKDFKCTLGEYVRKRRLLEAAKSL	67
RamA	23	LRIDDIARHAGYSKWHLQRLFLQYKGESLGRYIRERKLLLAARDL	67
RhaR	223	FALDKFCDEASCSESVLRQQFRQQTGMTINQYLRQVRVCHAQYLL	267
RhaS	188	VNWDVAVADQFSLSLRTLHRQLKQQTGLTPQRYLNRLRLMKARHLL	232
Rob	22	LSLDNVAAGAGYSKWHLQRMFKDVTGHAIGAYIRARRLSKSAVAL	66
SoxS	22	LNIDVVAKKSGYSKWYLQRMFRTVTHQTLGDYIRQRLLLAAVEL	66
TetD	45	LLLDDVANKAGYTKWYFQRLFKKVTGVTLAS YIRARRLTKAAVEL	89
XylR	302	IKVDQVLDVAGISRSNLEKRFKEEVGETIHAMIHAEKLEKARSL	346
XylS	228	ISLERLAELAMMSPRSLYNLFKXAGTTPKNYIRNRKLESIRACL	272
Xys3	228	ISLERLAELALMSPRSLYTLFEKHAGTTPKNYIRNRKLECIARL	272
Ya52	211	WHIEQLAELATMSRANFIRIFQQHIGMSPGRFLTQVRLQSAAFLL	255
YfiF	206	LKLTDVASHFHISGRHLRSLFAAELGVS YSEFVQNEKINKAAELL	250
YzbC	180	ITLAQLSQMAGISAKHYSESFKKWTGQSVTEFITKTRITKAKRLM	244
Consensus		E/D-X-V/L-A-D/E-X-A/S-G-X-S-X3-L-Q-X2-F-K/R/E-X2-T/I	

FIGURE 3A

MarA Homology Alignment

MarA	79	ILYLAERYGFESQQTLLTRTFKKNYFDVPPHKYR	110
AarP	87	VIDIALKYQFDSQQSFAKRFKAYLGISPSLYR	118
AfrR	74	IVDISERLFYDSQQTFTRTFKKNVSGYTPLOQR	105
AraC	246	IATVGRNVGFDDQLYFSRVFKKCTGASPSEFR	277
CafR	74	IIIEISAKLFYDSQQTFTRTFKKNVSGYTPLOQR	105
LumQ	213	LSQVAQLCGFSSQSSFSQAFRRLYGMSPTRYQ	244
MelR	259	ILDIALTAGFRSSRFYSTFGKYVGMSPQOYR	290
MmsR	266	VARVGQAVGYDDSYFSLRFSKVMGLSPSAYR	297
MsmR	241	IAEISNSVGFSDSLAFSKAFKNYFGKSPSKFR	272
ORF_f375	335	ASAAAMRVGYESASQFSREFKRYFGVTPGEDA	366
PocR	186	IASIARNLGFSSQTSYFCKVFRQTYQVTPQAYR	217
PqrA	74	ILDIALMYGFSSQATFTRIFKKNHFNTTPAKFR	105
RamA	74	VYDICALKYGFDSQQTFTRVFTTRTFNPPGAYR	105
RhaR	274	ISDISTECGFEDSNYFSVVFTRETGMTSPQWR	305
RhaS	239	VTDIAYRCGFSDSNHFSTLFRREFNWSPRDIR	270
Rob	73	ILDIALQYRFDSQQTFTRAFKKQFAQTPALYR	104
SoxS	73	IFDIAMDLYVSSQTFSRVFRQFDRTPSDYR	104
TetD	96	ILEIALKYQFDSQQSFTTRRFKYIFKVTPSYR	127
XylR	353	INEISQMCGYPSLQYFYFVFKKAYDTPKEYR	384
XylS	282	ITEIALDYGFLHLGRFAENYRSTFGELPSDTL	313
Xys3	282	VTEMALDYGFFHTGRFAENYRSTFGELPSDTL	313
Ya52	262	VLAIALEVGYQSEAHFCKVFNYYQLSPSQYR	293
YfiF	256	SIKEIAEEIGFSVHYFTRVFSKIGSSPGLFR	287
Yzbc	231	LKEIAHQTYQDEFYFSRIFKKYTGCSPTSYP	262

Concensus IXDIAXXGFXSXXYFXXXFXXXGXTPSXXR

FIGURE 3B

SEQUENCE LISTING

SEQ ID NO:1

```

1  gttaactgtg  gtggttgtca  ccgcccatta  cacggcatac  agctatateg  agccttttgt
61  acaaaacatt  gcgggattca  gcgccaactt  tgccacggca  ttaactggtat  tactcgggtg
121  tgcgggcatt  attggcagcg  tgattttcgg  taaactgggt  aatcagtatg  cgtctgcgtt
181  ggtgagtacg  gcgattgcmc  tgttgctggt  gtgcctggca  ttgctgttac  ctgcgggcaa
241  cagtgaataa  cacctcgggg  tgctgagtat  tttctggggg  atcgcgatga  tgatcatcgg
301  gcttggtatg  caggttaaag  tgctggcgtc  ggaccagatg  gctaccgacg  tcgcatggc
361  gctattctcc  ggcataatta  atattggaat  cggggcgggt  gcgttggtag  gtaatcaggt
421  gagtttgcac  tggccaatgt  cgatgattgg  ttatgtgggc  gcggtgcctg  cttttgcgc
481  gttaatttgg  tcaatcatta  tatttcgccc  ctggccagtg  aactcgaag  aacagacgca
541  atagttgaaa  ggcccattcg  ggcctttttt  aatggtacgt  tttaatgatt  tccaggatgc
601  cgtaataaat  aaactgcaca  cccatacata  ccagcaggaa  tcccacaga  cgggagatcg
661  cttcaatgcc  acccttgccc  accagccgca  taattgcgcc  ggagctgcgt  aggcttcccc
721  acaaaataac  cgccaccagg  aaaaagatca  gcggcggcgc  aaccatcagt  acccaatcag
781  cgaaggttga  actctgacgc  actgtggacg  ccgagctaat  aatcatcgtc  atggttcccg
841  gaccggcagt  acttggcatt  gccagcggca  caaaggcaat  attggcactg  ggttcatctt
901  ccagctcttc  cgacttgctt  ttcgctccg  gtgaatcaat  cgctttctgt  tgcggaaga
961  gcatccgaaa  accgataaac  gcgacgatta  agccgcctgc  aattgcgaga  ccgggaatcg
1021  aaatgccaaa  tgtatccatc  accagttgcc  cggcgttaata  cgccaccatc  atgatggcaa
1081  atacgtacac  cgaggccatc  aacgactgac  gattacgttc  ggcactgttc  atggtgctc
1141  ccaggccaag  aaataacgcg  acagtgtta  atgggttagc  taacggcagc  aacaccacca
1201  gccccaggcc  aattgcttta  acaaaatcta  acattggtgg  ttgttatcct  gtgtatctgg
1261  gttatcagcg  aaaagtataa  ggggtaaaac  aggataaagt  gtcactcttt  agctagcctt
1321  gcatcgcatt  gaacaaaact  tgaaccgatt  tagcaaaacg  tggcatcggg  caattcattc
1381  atttgactta  tacttgctcg  ggcaatatta  tcccctgcaa  ctaattactt  gccagggcaa
1441  ctaatgtgaa  aagtaccagc  gatctgttca  atgaaattat  tccattgggt  cgcttaatcc
1501  atatggttaa  tcagaagaaa  gatcgcctgc  ttaacgagta  tctgtctccg  ctggatatta
1561  ccgcggcaca  gtttaaggtg  ctctgctcta  tccgctgcmc  ggcgtgtatt  actccggtt
1621  aactgaaaaa  ggtattgtcg  gtcgacctgg  gagcactgac  ccgatgctg  gatcgcctgg
1681  tctgtaaagg  ctgggtggaa  aggttgccga  acccgaatga  caagcgcggc  gtactggtaa
1741  aacttaaccac  cggcggcgcg  gcaatatgtg  aacaatgcca  tcaattagtt  ggccaggacc
1801  tgcaccaaga  attaacaaaa  aacctgacgg  cggacgaagt  ggcaacactt  gagtatttgc
1861  ttaagaaagt  cctgccgtaa  acaaaaaaga  ggtatgacga  tgccagacg  caatactgac
1921  gctattacca  ttcataagcat  tttggactgg  atcggagaca  acctggaatc  gccactgtca
1981  ctggagaaag  tgtcagagcg  ttcgggttac  tccaaatggc  acctgcaacg  gatgtttaa
2041  aaagaaccg  gtcattcatt  aggcaatac  atccgcagcc  gtaagatgac  ggaaatcgcg
2101  caaaagctga  aggaaagtaa  cgagccgata  ctctatctgg  cagaacgata  tggcttcgag
2161  tcgcaacaaa  ctctgacctg  aaccttcaaa  aattactttg  atgttccgcc  gcataatac
2221  cggatgacca  atatgcaggg  cgaatcgcgc  tttttacatc  cattaatca  ttacaacagc
2281  tagttgaaaa  cgtgacaacg  tcaactgagg  aatcatgaaa  ccactttcat  ccgcaatagc
2341  agctgcgctt  attctctttt  ccgcgaggg  cgttgcgcaa  caaccacgc  agccagttgt
2401  tacttcttgt  gccaatgtcg  tggttgttcc  cccatcgcag  gaacaccac  cgtttgattt
2461  aaatcacatg  ggtactggca  gtgataagtc  ggatgcgctc  ggctgcccct  attataatca
2521  acacgctatg  tagtttgctt  tggccccgac  atctcggggc  ttattaactt  cccaccttta
2581  ccgctttacg  ccaccgcaag  ccaatacat  tgatatacag  cccggtcata  atgagcaccg
2641  cacctaaaaa  ttgcagacct  gttaaagcgt  catccaaca  tagtgcgca  cttgccagtc
2701  ctactacggg  caccagtaac  gataacgggt  caaccgcca  ggttcatag  cgtcccagta
2761  acgtccccc  gatccataa  ccaacaattg  tcgccacaaa  cgccagatac  atcagagaca
2821  agatggtggt  catatcgata  gtaaccagac  tgtgaatcat  ggttgcgaa  ccatcgagaa
2881  tcagcgagg  acaagaag  ggaatgattg  ggattaaagc  gctccagatt  accagcgaca
2941  tcaccgccg  acgcttgag  tgcgacatga  tctttttatt  gaagatgtt  ccacacgcc
3001  aactaaatgc  tgccgccagg  gtcaacataa  agccgagcat  cccacatgc  tgaccgttca
3061  gactatcttc  gattaacacc  agtacgcaa  aaatcgctaa  ggcgatcccc  gccaatgtt
3121  tgccatgcag  tcgctccccg  aaagtaaacc  cgccaagcat  gatagtaaaa  aacgctgtg
3181  cctgtaaac  cagcgaagcc  agtccagcag  gcataaccgaa  gttaatggca  caaaaaagaa
3241  aagcaactg  cgcaaaactg  atggttaatc  cataccocag  cagcaaatc  agtggtaact
3301  tcggtcgtgc  gacaaaaaag  atagccggaa  aagcgaccag  cataaagcgc  aaaccggcca

```

3361 gcatcagcgt ggcattgttat gaagccccac tttgatgacc acaaaattta gccccatac
3421 gaccactacc agtagcgcca acacccccatc ttttcgcgac attctaccgc ctctgaattt
3481 catcttttgt aagcaatcaa cttagctgaa tttacttttc tttaacagtt gattcgttag
3541 tcgccgggta cgacggcatt aatgcgcaaa taagtgcgta tacttcggat ttttgccatg
3601 ctattttctt acatctctaa aacaaaacat aacgaaacgc actgccggac agacaaatga
3661 acttatccct acgacgctct accagcgccc tttctgcctc gtcgttgta ttaaccatcg
3721 gacgcgggcg taccgtgcca tttatgacca tttacttgag tcgccagtac agcctgagtg
3781 tcgatctaata cggttatgcg atgacaattg cgctcactat tggcgtcggt tttagcctcg
3841 gttttggtat cctggcggat aagtctcgaca agaaacgcta tatgttactg gcaattaccg
3901 ccttcgcccag cggttttatt gccattactt tagtgaataa cgtgacgctg gttgtgctct
3961 tttttgccc cattaactgc gcctattctg tttttgctac cgtgctgaaa gcctggtttg
4021 ccgacaatct ttcgtccacc agcaaaacga aaatcttctc aatcaactac accatgctaa
4081 acattggctg accatcggtc cgccgctcgg cacgctggtg gtaatgcaga gcatcaatct
4141 gcccttctgg ctggcagcta tctgttccgc gtttcccatg cttttcattc aaatttgggt
4201 aaagcgcagc gagaaaatca tcgccacgga aacaggcagt gtctggtcgc cgaaagtttt
4261 attacaagat aaagcactgt tgtggtttac ctgctctggt tttctggctt cttttgtaag
4321 cggcgcattt gcttcatgca tttcacaata tgtgatggtg attgctgatg gggattttgc
4381 cgaaaagggtg gtcgcggttg ttcttccggt gaatgctgcc atggtggtta cgttgcaata
4441 ttccgtgggc cgccgactta acccggtctaa catccgcgcg ctgatgacag caggcacctc
4501 ctgtttcgtc atcggctctgg tcggttttat ttttccggc aacagcctg agtataaggt
4561 tatgtcagct gcggtattta ctgtcggta aatcatttat gcgcccggcg agtataaggt
4621 gattgaccat attgcgccc cagaaatgaa agccagctat ttttccgcc agtctttagg
4681 ctggcttggg gcgcgatta acccattagt gagtggcgta gtgctaacca gcctgccgcc
4741 ttccctcgctg tttgtcatct tagcgttggg gatcattgct gcgtgggtgc tgatgttaaa
4801 agggattcga gcaagaccgt gggggcagcc cgcgctttgt tgatttaagt cgaacacaat
4861 aaagatttaa ttcagccttc gtttaggtta cctctgctaa tatctttctc attgagatga
4921 aaattaaggt aagcgaggaa acacaccaca ccataaacgg aggcaataa tgctgggtaa
4981 tatgaatggt tttatggccg tactgggaat aattttattt tctggttttc tggccgcgta
5041 tttcagccac aaatgggatg actaatgaac ggagataatc cctcacctaa cgggccctt
5101 gttacagttg tgtacaaggg gcctgatttt tatgacggcg aaaaaaac gccagtaaac
5161 cggcggtgaa tgcttgcag gatagatttg tgttttgctt ttacgctaac aggcattttc
5221 ctgcaactgat aacgaatcgt tgacacagta gcatcagttt tctcaatgaa tgtaaacgg
5281 agcttaaact cggttaatca cttttgttc gtcaataaac atgcagcgtat ttctccggt
5341 ttgcttacc tcatacattg cccggctcgc ctctccaatg accacatcca gaggctcttc
5401 aggaaatgcg cgactcacac ctgctgtcac ggtaatggtg atatgccctt cagaatgtgt
5461 gatggcagtg ttatcgacta actggcaaat tctgacacct gcacgacatg cttcttcac
5521 attagccgct ttgacaataa tgataaatc ttcgccccg tagcgataaa ccgtttcgta
5581 atcacgcgct caactggcta agtaagttgc cagggtgctg aatactacat cgcgattaa
5641 atgcccgtag tatcattaac caatttaaat cggctcaatat ccaacaacat taaataaaga
5701 ttcagaggct cagcgttgcg taactgatga tcaaaggatt catcaagaac ccgacgacc
5761 ggcaatcccg tcaaacatc catattgcta cggatcgtca gcaataaat tttgtaatcg
5821 gttaatgccc cagtaaaaga aagcaacccc tccgaaagg cgtcgaaatg cgcgtcctgc
5881 cagtgatttt caacaatagc cagcattaat tcccgaccac agttatgcat atggtgatgg
5941 gcagaatcca ttagccgaac gtaaggtaat tcatcgttat cgagtggccc cagatgatca
6001 atccaccgac caaactggca cagtccataa gaatggttat ccgttatttc tggcttactg
6061 gcatctctcg cgaccacgct gtgaaacata ctaccagcc actggtatg ggcacgata
6121 gccttattga gatttaacaa gatggcatca atttccggtg tcttcttgat cattgccact
6181 cctttttcac agttccttgt gcgctgctatt ctaacgagag aaaagcaaaa ttacgtcaat
6241 attttcatag aaatccgaag ttatgagtca tctctgagat aacatttgta tttaaaaca
6301 aatcagcggg taaaaaagtg ttaattctg taaattacct ctgcattatc gtaataaaaa
6361 ggatgacaaa tagcataacc caatacccta atggcccagt agttcaggcc atcaggctaa
6421 tttattttta tttctgcaaa tgagtgacc gaacgacggc cggcgcgctt ttcttatcca
6481 gactgccact aatggtgatc atctggtccg gctgaacttc tcgtccatca aagacggccg
6541 caggaataac gacattaatt tcaccgctct tatcgcgaaa aacgtaacgg tcctctcctt
6601 tgtgagaaat caaattaccg cgtagtgaaa ccgaagcgc atcgtgcatg gtttttgcca
6661 aatcaacggg cattttttt gcatcatcgg ttcgcgata gccatcttct attgcatgag
6721 gcggcgggtg cgctgcatcc tgttttaaac cgccctggte atctgccaac gcataaggca
6781 tgacaagaaa acttgctaat acaatggcct gaaatttcat actaactcct taattgcgtt
6841 tggtttgact tattaagtct ggttgctatt tttataattg ccaataaaga atattgccaa

6901 ttgttataag gcatttaaaa tcagccaact agctgtcaaa tatacagaga atttaactca
6961 ctaaagttaa gaagattgaa aagtcttaaa catattttca gaataatcgg atttatatgt
7021 ttgaaaatta ttatattgga cgagcataca gaaaaagcaa atcaccttta catataaaag
7081 cgtggacaaa aaacagtgaa cattaataga gataaaattg tacaacttgt agataccgat
7141 actattgaaa acctgacatc cgcgttgagt caaagactta tcgcggatca attacgctta
7201 actaccgccg aatcatgcac cggcggtaag ttggctagcg cctgtgtgac agctgaagat
7261 acacccaaat ttacgggtgc aggtcttgtt actttcaccg atcaggcaaa gatgaaaatc
7321 ctcagcgtaa gccagcaatc tcttgaacga tattctgagg tgagtggaga agtggcagca
7381 gaaatggcaa ccggtgccat agagcgtgag gatgctgatg tcagtattgc cattaccggc
7441 tacggcggac cggaggggcg tgaagatggt acgccagcgg gtaccgtctg gtttgcgtgg
7501 catattaaag gccagaacta cactgcggtt atgcattttg ctggcgactg cgaaacggta
7561 ttagctttag cggtgagggt tgccctcgcc cagctgctgc aattactgct ataaccaggc
7621 tggcctggcg atatctcagg ccagccattg gtggtgttta tatgttcaag ccacgatggt
7681 gcagcatcgg cataatctta ggtgccttac cgcgccattg tcgatcacagg cgttccagat
7741 cttcgctggt acctctggaa aggatcgct cgcgaaaacg cagcccattt tcacgcgta
7801 atccgccctg ctcaacaaac cactgataac catcatcggc caacatttgc gtccacagat
7861 aagcgtaata acctgcag

SEQ ID NO:2

MTMSRRNTDAITIHSLDWIEDNLESPLSLEKVSERSGYSKWHLQRMFKKETGHSILGQYIRSRKMTEIAQKLKESNEP
ILYLAERYGFESQOTLTRTFKKNYFDVPPHKYRMTNMQGESRFLHPLNHYNS