INFECTION AND IMMUNITY, June 1996, p. 2365–2367 Vol. 64, No. 6
0019-9567/96/$04.00+0

The Predicted Amino Acid Sequence of the *Salmonella typhimurium* Virulence Gene *mviA*$^+$ Strongly Indicates that MviA Is a Regulator Protein of a Previously Unknown *S. typhimurium* Response Regulator Family

WILLIAM H. BENJAMIN, JR.,[1,2]* XINPING WU,[2] AND W. EDWARD SWORDS[2]

The Departments of Pathology[1] and Microbiology,[2] The University of Alabama at Birmingham, Birmingham, Alabama 35294

Received 20 October 1995/Returned for modification 3 January 1996/Accepted 18 March 1996

The *Salmonella typhimurium* virulence gene *mviA*$^+$ has a predicted amino acid sequence with homology to the N-terminal 112-amino-acid sequence of response regulator proteins. A previously described mutant allele (*mviA*), which restores virulence to avirulent LT2 strains, was shown to contain a point mutation which would be predicted to cause a single amino acid change, V-102→G (W. H. Benjamin, Jr., J. Yother, P. Hall, and D. E. Briles, J. Exp. Med. 174:1073–1083, 1991). A comparison of the nucleotide sequence of *mviA*$^+$ with that of the *Escherichia coli* and *Salmonella typhi* genes revealed a high degree of conservation.

We have shown that differences in the *mviA* gene in *Salmonella typhimurium* strains of the LT2 background directly affect the ability of the bacterium to grow in *Ity*^{s} (susceptible genotype) but not *Ity*^{r} mice (4–6). Thus, an understanding of the mechanism of action of *mviA*$^+$ may shed light on the resistance mechanism of *Ity* as well as on salmonella virulence. *mviA*$^+$ was initially characterized by its ability to confer avirulence upon LT2 strains of *S. typhimurium* (6). A comparison of the deduced amino acid sequence of *mviA*$^+$ with that of known proteins reveals a high degree of homology with a well-characterized family of bacterial response regulatory proteins and suggests that MviA is the second component of a previously unknown two-component response regulatory system. Two-component response regulator systems are commonly used by prokaryotes in environment-dependent control of gene expression (1, 9, 21). On the basis of sequence homology, it is estimated that approximately 40 two-component response regulator systems exist in *Escherichia coli*. However, functional data to support this assignment are available for only a portion of the response regulator-like genes that have been sequenced (1, 18). Several two-component regulatory systems (*phoP*/*phoQ*, *envZ*/*ompR*, *bvgS*/*bvgA*, and *virG*/*virA*) are known bacterial virulence genes (3, 11–13, 15, 17, 20).

The intact *mviA*$^+$ allele from an avirulent *S. typhimurium* LT2 strain was cloned and sequenced. The *mviA*$^+$ sequence was found on a 1.92-kb *Bst*XI-*Cla*I fragment of pWB4184 (6), which includes a 1.57-kb *Dra*I fragment that constitutes the minimal subclone causing avirulence (reference 6 and data not shown). DNA sequencing of double-stranded plasmids was performed with the Sequenase kit (U.S. Biochemicals). On the basis of initial sequence data obtained from the cloned *mviA*$^+$ gene in plasmid pWB4184, appropriate primers were constructed, and the entire *mviA*$^+$ gene was amplified by PCR from chromosomal DNA prepared from cultures of *S. typhimurium* by previously described methods (14). PCR products were made single stranded by exonuclease digestion with a presequencing kit (U.S. Biochemicals). The primers used include the following pUC19 universal and reverse primers: *mviAa*, 5′ ATGCAGCAAGATTTGCTCTC 3′ (nucleotides [nt] 53 to 72); *mviAb*, 5′ ACCAATGGCTGCGTCATGTT 3′ (nt 506 to 487); *mviAc*, 5′ TTTAACGGCTTATTGCAGGA 3′ (nt 1156 to 1175); *mviAd*, 5′ CATTCCGCAGACAACATCAA 3′ (nt 1503 to 1489); *mviAe*, 5′ AGCAAATTCTGACAGGC GCA 3′ (nt 422 to 431); *mviAf*, 5′ GGGGGTGAAGATGTC TTGCT 3′ (nt 789 to 809); and *mviAg*, 5′ TCAACCCGCG AGTTAAACAT 3′ (nt 895 to 875). A single open reading frame was identified which is predicted to encode a protein 337 amino acids in length with a molecular mass of 37,292 Da and with a pI of 5.5. No obvious transmembrane or leader sequences were noted. A typical promoter [−35 (TATACC) N_{17} (TTTAAT)] was identified 89 bp upstream of the predicted start codon of the open reading frame (16).

GenBank searches yielded no significant homology at the nucleotide level, except for the apparent *mviA*$^+$ homolog from *E. coli*, which was previously designated *orf37* (7, 8). Because of the high degree of homology and an almost identical chromosomal map location, we believe that *orf37* is a homolog of *mviA*$^+$; hereafter, we refer to it as *E. coli mviA*$^+$. The 110 best matches to *mviA*$^+$, obtained from a TfastA search with the 112 N-terminal amino acids predicted from the *mviA*$^+$ sequence were analyzed, and 84 unique sequences were identified. Of these, 77 were identified as having the response regulator motif, on the basis of the invariant amino acids corresponding to MviA residues D-15, D-58, T- or S-83, and K-108 (21, 22). The deduced sequence was 90% identical to the *E. coli* MviA predicted protein product (Fig. 1). In contrast to *E. coli* MviA, the additional 76 known and apparent response regulators had 20 to 37% amino acid sequence homology with MviA over at least 98 amino acids. In each case, the region of homology covered most of the N-terminal 112 amino acids of MviA (Fig. 1). The hydrophobic regions described by Stock et al. (21, 22) were also highly conserved between MviA and the response regulator protein sequences. By using the seven *mviA* oligonucleotide primers, *mviA*$^+$ homologs were amplified from *E. coli* (strain LE392) and two strains of *Salmonella typhi* (the virulent strain Q5 and the avirulent vaccine strain Ty21a). The PCR products were sequenced as described above, and the nucleotide sequences obtained were compared with that of *mviA*$^+$ from *S. typhimurium*. Q5 was found to have a single nucleotide difference, in codon 98 (TTG → TTA); this is a silent mutation, because it has no effect on the predicted product. Ty21a was

* Corresponding author. Phone: (205) 934-1923. Fax: (205) 934-0605. Electronic mail address: Benjamin@wp.path.uab.edu.

```
          1                                                  58
MviA (St) MTQPLVGKQI LIVEDEPVFR SLLDSWFSSL GATTAL..AG DGVDALELMG RFTPDLMICD
MviA (Ec) MTQPLVGKQI LIVEDEQVFR SLLDSWFSSL GATTVL...AA DGVDALELLG GFTPDLMICD
CheY (St) ..MADKELKF LVVDDFSTMR RIVRNLL.KE LGFNNVEEAE DGVDALNKLQ AGGFGFIISD
HydG (St) ..MIRGKIDI LVVDDDVSHC TILQALL.RG WGY.NVALAY SGHDALAQVR EKVFDLVLCD
NarL (Ec) .MSNQEPATI LLIDDHPMLR TGVKQLISMA PDITVVGEAS NGEQGIELAE SLDPDLILLD
OmpR (Ec) ...MQENYKN LVVDDDMRLR ALLERYLTEQ G..FQVRSVA NAEQMDRLLT RESFHLMVLD
CheB (Ec) ....MSKIRV LSVDDSALMR QIMTEIINSH SDMEMVATAP DPLVARDLIK KFNPDVLTLD
          59                                  #              114
MviA (St) IAMPRMNGLK LVENLRNRG. .DQTPILVIS ATENMADIAK ..ALRLGVED VLLKPVKDLN
MviA (Ec) IAMPRMNGLK LLEHIRNRG. .DQTPVLVIS ATENMADIAK ..ALRLGVED VLLKPVKDLN
CheY (St) WNMPNMDGLE LLKTIRADSA MSALPVLMVT AEAKKENIIA ..AQAGASG YVVKPF.TAA
HydG (St) VRMAEMDGIA TLKEIKALNP ..AIPILIMT AFSSVETAVE ..ALKAGALD YLIKPL.DFD
NarL (Ec) LNMPGMNGLE TLDKLREKSL SGRIVVFSVS NHE..EDVVT ..ALKRGADG YLLKDM.EPE
OmpR (Ec) LMLPGEDGLS ICRRLRSQS. .NPMPIIMVT AKGEEVDRIV ..GLEIGADD YIPKPF.NPR
CheB (Ec) VEMPRMDGLD FLEKLMR... LRPMPVVMVS SLTGKGSEVT LRPLELGAID FVTKPQLG..
         115                                                174
MviA (St) RLRETVFACL YPNMFNSRVE EEERLFRDWD AMVSNPTAAA QLLQELQPPV QQVISHCRIN
MviA (Ec) RLREMVFACL YPSMFNSRVE EEERLFRDWD AMVDNPAAAA KLLQELQPPV QQVISHCRVN
CheY (St) TLEEKLNKIF ..........  .......EKL GM*
HydG (St) RLQETLEKAL AHTRETG..A ELPSASAAQF GMIGSSPAMQ HLLNEI.... ..........
NarL (Ec) HLLKALHQAA AGEMVLS..E ALTPVLAASL .....RANRA TTERDV.... ..........
OmpR (Ec) ELLARIRAVL RRQANELPGA PSQEEAVIAF GKFKLNLGTR EMFREDEP.. MPLTSGEFAV
CheB (Ec) .IREGMLAYN EMIAQKVETA AKASLAAHKP LSAPTTLKAG PLLSSEKLIA IGASTGGTEA
         175                                                234
MviA (St) YRQLVSADQP GLVLDIAPLS DNELAFYCLD VTRAGDNGVL AALLLRALFN GLLQDQLGQQ
MviA (Ec) YRQLVAADKP GLVLDIAALS ENDLAFYCLD VTRAGHNGVL AALLLRALFN GLLQEQLAHQ
HydG (St) ..AMVRPSDA TVLIHGDSGT GKELVARALH ACSARSDRPL VTLNCAALNE SLLESELFGH
NarL (Ec) ..NQLTPRER DILKLIAQGL PNKMIARRLD ITESTVKVHV KHMLKKMKLK SRVEAAVWVH
OmpR (Ec) LKALVSHPRE PLSRDKLMNL ARGREYSAME RSIDVQISRL RRMVESDPAH PRYIQTVWGL
CheB (Ec) IRHVLQP... .LPLSSPALL ITQHMPPGFT RSFADRLNKL CQIGVKEAED G....ERVLP
         235                                                294
MviA (St) KHRLPELGAL LKQVNHLLRQ ANLPGQFPLF VGYYHSELKN LILVSAGLNA TLNTGAHQVQ
MviA (Ec) NQRLPELGAL LKQVNHLLRQ ANLPGQFPLL VGYYHRELKN LILVSAGLNA TLNTGEHQVQ
HydG (St) EKGAFTGADK RREGRFVEAD G...GTLFLD EIGDISPLMQ VRLLRAIQER EVQRVGSNQT
NarL (Ec) QERIF*
OmpR (Ec) GYVFVPDGSK A*
CheB (Ec) GHAYIAPGDR HMELSRSGAN YQIKIHDGPA VNRHRPSVDV LFHSVAKQAG RNAVGVILTG
         295                                    337
MviA (St) ISSGVPLGTL GNAYLNQLSQ RCDSWQCQIW GAGGRLRLML SAE*
MviA (Ec) ISNGVPLGTL GNAYLNQLSQ RCDAWQCQIW GTGGRLRLML SAE*
HydG (St) LSVDVRLIAA THRDLAEEVS AGRFRQDLYY RLNVVAIEMP SLRQRREDIP LLADHFLRRF
CheB (Ec) MGNDGAAGML AMRQAGAWTL AQNEASCVVF GMPREAINMG GVCEVVDLSQ VSQQMLAKIS

HydG (St) AERNRKVVKG FTPQAMDLLI HYDWPGNIRE LENAIERAVV LLTGEYISER ELPLAIAATP
CheB (Ec) AGQAIRI*
HydG (St) IKTEYSGEIQ PLVDVEKEVI LAALEKTGGN KTEAARQLGI TRKTLLAKLSR*
```

FIG. 1. Amino acid sequence homology of MviA and other response regulators. The GCG Pileup program was used to align representative response regulators which were found to have homology to MviA. The underlined residues are the amino acids found to be homologous to *mviA* (*S. typhimurium* [St]) in the other proteins. MviA (*E. coli* [Ec]), M64675; CheY (*S. typhimurium*), M12131; HydG (*S. typhimurium*), M64988; NarL (*E. coli*), M24910; OmpR (*E. coli*), J01656; CheB (*E. coli*), M13463. Boldface residues delineate the invariant amino acids D-15, D-58, S-83 or T-83, and K-108. The site of the mutation found in strain WB600 (V-102→G) is indicated by #.

found to have the same mutation as well as two other silent mutations (codons 46 and 231). The sequence obtained from *E. coli mviA*$^+$ matched that reported for *orf37* and was compared with the sequence obtained from *S. typhimurium mviA*$^+$. The *E. coli* sequence was >75% conserved at the nucleotide level and >90% conserved at the level of the predicted product.

Identification of the *mviA*$^+$ gene product MviA. The open reading frame of *mviA*$^+$ (*S. typhimurium*) is expected to encode a protein with a size of 37,292 Da. The *mviA*$^+$ minimal coding region was cloned into pUC19 to form plasmid pWES5. A kanamycin resistance (Km^r) cassette from plasmid pBSL14 (2) was excised with *Bam*HI and ligated into the *mviA*$^+$-internal *Bgl*II site of pWES5. The Km^r cassette was then removed with *Nco*I and *Bgl*II, and the vector was treated with T4 DNA polymerase and ligated to form plasmid pWES9. The *mviA*::STOP mutant allele in pWES9 resulted in an insertion of 16 nucleotides and thus a frameshift. Analysis of the sequence of *mviA*::STOP reveals a stop codon immediately adjacent to this frameshift, which would be predicted to halt translation early in the predicted product. In vitro transcription-translation with the *E. coli* S30 extract system for circular DNA templates (Promega, Madison, Wis.) was performed with [^{35}S]methionine label to show that *mviA*$^+$ encodes the production of a protein with an apparent size of 38.5 kDa (23). This

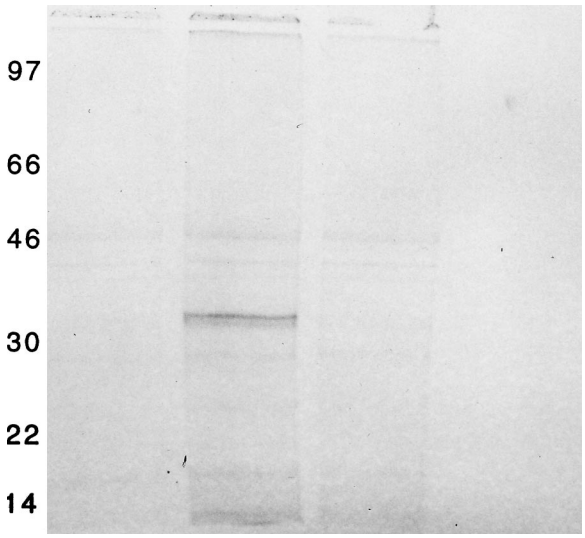

FIG. 2. Visualization of the *mviA*$^+$ gene product by in vitro transcription-translation. MviA is a 38.5-kDa protein which is clearly produced by pWES5 (lane 2) and not by the vector pUC19 (lane 1) or by an *mviA*::STOP mutant allele (lane 3). Lane 4 is a control with no exogenous DNA template. Lane numbers given above refer to positions from left to right. Molecular masses are given (kilodaltons) on the left.

protein was shown to be MviA, because it is expressed from plasmid pWES5 (*mviA*$^+$) and not from the vector (pUC19) or the *mviA*::STOP allele (pWES9). These results are shown in Fig. 2.

Sequence of a mutant *mviA* allele. Previously reported mutations which affect the activity of response regulators are mostly clustered around the invariant amino acids (D-15, D-58, T- or S-83, and K-108) (10, 19). These mutations can affect the phosphorylation and dephosphorylation of response regulators and thus affect expression of regulated genes. Some mutations, especially those in the conserved amino acids D-15 and D-58 and K-108, can abolish phosphorylation and therefore transcriptional regulation (18, 22). Other mutations near these residues result in independence of the response regulators from their normal sensor proteins. Minute changes in the rate of phosphorylation or dephosphorylation can result in altered gene regulation, as evidenced by phenotypic modifications (10). LT2 *S. typhimurium* strains with a functional *mviA*$^+$ allele are avirulent, whereas strains with interrupted *mviA*$^+$ alleles are virulent (6). We have identified a single allele, *mviA*^{V102G} (from strain WB600) which confers a virulent phenotype (3, 6). The *mviA*^{V102G} allele differed by only a single base pair from the *mviA*$^+$ sequence, resulting in a change in the predicted amino acid sequence, V-102→G, shown in Fig. 1. The close proximity of the V-102→G mutation to K-108 is consistent with its apparent effect on the activity of the response regulator. It is worth noting that the V-102→G mutation was not observed in the *mviA* genes from any source other than strain WB600. The previous designation of the *mviA*$^+$ allele was based upon the ability to confer avirulence upon LT2 strains and thus was somewhat provisional. The present sequence data support the designation of *mviA*$^+$ as the wild-type allele.

The results of this study indicate that *mviA*$^+$ encodes a response regulatory protein which is highly conserved among enteric species. The dramatic effects of mutations in *mviA*$^+$ on the virulence of *S. typhimurium* suggest that the homologous gene found in *S. typhi* may be an important virulence determinant in human typhoid fever. Mutations in *mviA*$^+$ may there-

fore be useful in the rational construction of attenuated strains for use as live vaccines.

Nucleotide sequence accession number. All sequences obtained in this study have been submitted to GenBank under accession number U53173.

This work was supported by a grant from the National Institutes of Health (RO1 AI15946) to William H. Benjamin. W. Edward Swords was supported in part by training grant T32 AI107041 from the NIH. Genetics Computer Group sequence analysis was made available through the computer resources of the University of Alabama at Birmingham Center for AIDS Research (P30 AI27767).

We thank David E. Briles for assistance and Rebecca C. Tart and Melissa Caimano for critical reading of the manuscript.

REFERENCES

1. **Albright, L. M., E. Huala, and F. M. Ausubel.** 1989. Prokaryotic signal transduction mediated by sensor and regulator protein pairs. Annu. Rev. Genet. **23:**311–336.
2. **Alexeyev, M. F.** 1995. Three kanamycin resistance gene cassettes with different polylinkers. BioTechniques **18:**52–55.
3. **Aricó, B., V. Scarlato, D. M. Monack, S. Falkow, and R. Rappuoli.** 1991. Structural and genetic analysis of the *bvg* locus in *Bordetella* species. Mol. Microbiol. **5:**2481–2491.
4. **Benjamin, W. H., Jr., C. L. Turnbough, Jr., J. D. Goguen, B. S. Posey, and D. E. Briles.** 1986. Genetic mapping of novel virulence determinants of *Salmonella typhimurium* to the region between *trpD* and *supD*. Microb. Pathog. **1:**115–124.
5. **Benjamin, W. H., Jr., C. L. Turnbough, Jr., B. S. Posey, and D. E. Briles.** 1986. *Salmonella typhimurium* virulence genes necessary to exploit the $Ity^{s/s}$ genotype of the mouse. Infect. Immun. **51:**872–878.
6. **Benjamin, W. H., Jr., J. Yother, P. Hall, and D. E. Briles.** 1991. The *Salmonella typhimurium* locus *mviA* regulates virulence in Ity^{s} but not Ity^{r} mice: functional *mviA* results in avirulence; mutant (nonfunctional) *mviA* results in virulence. J. Exp. Med. **174:**1073–1083.
7. **Bösl, M., and H. Kersten.** 1994. Organization and functions of genes in the upstream region of *tyrT* of *Escherichia coli*: phenotypes of mutants with partial deletions of a new gene (*tgs*). J. Bacteriol. **176:**221–231.
8. **Bösl, M. R.** 1993. Genetic map of the *tyrT* region of *Escherichia coli* from 27.1 to 27.7 minutes based exclusively on sequence data. J. Bacteriol. **175:**7751–7753.
9. **Cozzone, A.** 1993. ATP-dependent protein kinases in bacteria. J. Cell. Biochem **51:**7–13.
10. **Dahl, M. K., T. Msadek, F. Kunst, and G. Rapoport.** 1991. Mutational analysis of the *Bacillus subtilis* DegU regulator and its phosphorylation by the DegS protein kinase. J. Bacteriol. **173:**2539–2547.
11. **Dorman, C. J., S. Chatfield, C. F. Higgins, C. Hayward, and G. Dougan.** 1989. Characterization of porin and *ompR* mutants of a virulent strain of *Salmonella typhimurium*: *ompR* mutants are attenuated in vivo. Infect. Immun. **57:**2136–2140.
12. **Dziejman, M., and J. J. Mekalanos.** 1995. Two-component signal transduction and its role in the expression of bacterial virulence factors, p. 305–317. *In* J. A. Hoch and T. J. Silhavy (ed.), Two-component signal transduction. ASM Press, Washington, D.C.
13. **Fields, P. I., E. A. Groisman, and F. Heffron.** 1989. A *Salmonella* locus that controls resistance to microbiocidal proteins from phagocytic cells. Science **243:**1059–1062.
14. **Gay, N. J.** 1984. Construction and characterization of an *Escherichia coli* strain with a *uncI* mutation. J. Bacteriol. **158:**820–825.
15. **Groisman, E. A., E. Chiao, C. J. Lipps, and F. Heffron.** 1989. *Salmonella typhimurium phoP* virulence gene is a transcriptional regulator. Proc. Natl. Acad. Sci. USA **86:**7077–7081.
16. **Harr, R., M. Häggström, and P. Gustafsson.** 1983. Search algorithm for pattern match analysis of nucleic acid sequences. Nucleic Acids Res. **11:**2943–2957.
17. **Mekalanos, J. J.** 1992. Environmental signals controlling expression of virulence determinants in bacteria. J. Bacteriol. **174:**1–7.
18. **Parkinson, J. S., and E. C. Kofoid.** 1992. Communication modules in bacterial signaling proteins. Annu. Rev. Genet. **26:**71–112.
19. **Pazour, G. J., C. N. Ta, and A. Das.** 1992. Constitutive mutations of *Agrobacterium tumefaciens* transcriptional activator *virG*. J. Bacteriol. **174:**4169–4174.
20. **Stibitz, S., A. A. Weiss, and S. Falkow.** 1988. Genetic analysis of a region of the *Bordetella pertussis* chromosome encoding filamentous hemagglutinin and the pleiotropic regulatory locus *vir*. J. Bacteriol. **170:**2904–2913.
21. **Stock, J. B., A. J. Ninfa, and A. M. Stock.** 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. **53:**450–490.
22. **Stock, J. B., A. M. Stock, and J. M. Mottonen.** 1990. Signal transduction in bacteria. Nature (London) **344:**395–400.
23. **Zubay, G.** 1973. In vitro synthesis of protein in microbial systems. Annu. Rev. Genet. **7:**267–287.

Editor: A. O'Brien