JOURNAL OF BACTERIOLOGY, Dec. 1990, p. 6863–6870
0021-9193/90/126863-08$02.00/0

Vol. 172, No. 12

Transcriptional Regulation by Iron of a *Vibrio cholerae* Virulence Gene and Homology of the Gene to the *Escherichia coli* Fur System

MARCIA B. GOLDBERG,[1] STEPHANIE A. BOYKO,[1] AND STEPHEN B. CALDERWOOD[1,2]*

Infectious Disease Unit, Massachusetts General Hospital, Boston, Massachusetts 02114,[1] and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02215[2]

Received 4 May 1990/Accepted 7 September 1990

We have previously described an iron-regulated virulence determinant in *Vibrio cholerae*. Strain MBG40, which contains a Tn*phoA* insertion mutation in the iron-regulated gene *irgA*, has reduced virulence in a newborn mouse model and has lost the major 77-kDa iron-regulated outer membrane protein. We report here the cloning of the *irgA'*-*'phoA* gene fusion, the sequencing of the 5'-proximal portion of *irgA*, and the definition of its promoter region by primer extension. The deduced amino acid sequence of the amino-terminal portion of IrgA is homologous to the ferrienterochelin receptor of *Escherichia coli* (FepA), suggesting that IrgA may be the iron-vibriobactin outer membrane receptor. Iron regulation of *irgA* in an *E. coli* background and that of the *E. coli* gene *slt*-IA in a *V. cholerae* background are reciprocal, suggesting a common mechanism of iron regulation. Regulation of *irgA* by iron in *V. cholerae* occurs at the transcriptional level, and there is an interrupted dyad symmetric sequence in the vicinity of the promoter that is homologous to Fur binding sites of *E. coli*. Unlike iron-regulated genes in *E. coli*, however, transcription of *irgA* requires an additional 900 bp of upstream DNA that contains an open reading frame in inverse orientation to *irgA*.

Vibrio cholerae is the causative agent of a severe dehydrating and occasionally fatal diarrhea that occurs primarily in developing countries of Asia and Africa. Disease develops after ingestion of the organisms in contaminated fresh or salt water. The organisms pass through the stomach into the small intestine, where they penetrate the mucus gel overlying the intestinal mucosa, adhere to intestinal epithelial cells, and produce several extracellular secreted proteins, including cholera toxin, neuraminidase, and hemolysin. Cholera toxin genes are present in all pathogenic strains; the genetic regulation of these genes and a toxin-coregulated pilus (TcpA) in response to environmental signals has been described before (29, 37). The pathogenic role and genetic regulation of other virulence determinants in *V. cholerae* are less well understood.

Several virulence determinants in bacteria are regulated by the concentration of free iron in the environment, with increased expression occurring with low iron concentrations. These virulence determinants include Shiga toxin of *Shigella dysenteriae* 1 (9), Shiga-like toxin I of *Escherichia coli* (5), diphtheria toxin of *Corynebacterium diphtheriae* (28), and exotoxin A of *Pseudomonas aeruginosa* (3, 4). We have previously described an iron-regulated virulence determinant in *V. cholerae* (10). *V. cholerae* MBG40, which contains a Tn*phoA* insertion mutation in the iron-regulated gene *irgA*, has reduced virulence in a newborn mouse model and has lost the 77-kDa major iron-regulated outer membrane protein. Several other outer membrane proteins in *V. cholerae* have been shown to be regulated by iron (17, 32, 33), as has a hemolysin (36), but no role in virulence has yet been described for these other proteins. The molecular mechanism of iron regulation of any of these genes in *V. cholerae* has not been determined.

Two systems of iron assimilation from the environment have been described in *V. cholerae*, one involving the siderophore vibriobactin and a second that appears to utilize ferric citrate as a source of iron (11, 34). Vibriobactin is a phenolate-type siderophore that has structural similarities to other siderophores, including enterochelin, ferrichrome, and agrobactin (11). Work by Griffiths et al. (11) suggests that vibriobactin and enterochelin have similar synthetic pathways and that vibriobactin and agrobactin may recognize a shared membrane receptor or uptake mechanism.

In *E. coli*, several genes in addition to those encoding Shiga-like toxin I are regulated by iron, including the genes encoding the ferrienterochelin receptor (*fepA*) (30), the aerobactin uptake outer membrane protein (*iucA*) (2), and other genes involved in iron assimilation (26). Coordinate regulation of these diverse genes occurs by a regulatory locus, *fur*, whose protein product acts as a repressor at the transcriptional level in the presence of sufficient iron (1, 12, 13). Calderwood and Mekalanos have identified and characterized a 21-bp dyad symmetric nucleotide sequence, overlapping the Shiga-like toxin I promoter, that serves as an operator binding site for Fur in the presence of iron (5, 6). This operator binding site is homologous to similar elements in the promoters of the other iron-regulated genes in *E. coli* (5, 30).

In this paper, we report the cloning and sequencing of the 5'-proximal portion of *irgA* and the definition of its promoter region. The deduced amino acid sequence of the amino-terminal portion of IrgA is homologous to the ferrienterochelin receptor (FepA) of *E. coli*, suggesting that IrgA may be the iron-vibriobactin outer membrane receptor. Regulation of *irgA* by iron in *V. cholerae* occurs at the transcriptional level and involves an operator binding site homologous to Fur binding sites of *E. coli*.

---

* Corresponding author.

MATERIALS AND METHODS

Bacterial strains. The *V. cholerae* wild-type strain used in this study was classical Ogawa strain 0395 Sm^r (22). *V. cholerae* MBG40 is 0395 *irgA*::Tn*phoA* (10). *E. coli* CC118 is a *phoA* deletion derivative of MC1000 (21), and *E. coli* SM10 has been described previously (35). *E. coli* DH5α was obtained from Bethesda Research Laboratories Life Technologies, Inc., Gaithersburg, Md.

Media. Two types of liquid media were used to assess the effect of iron concentration on gene expression as described previously (10): (i) LB medium with or without the addition of the iron chelator 2,2-dipyridyl (final concentration, 0.2 mM) and (ii) Tris-buffered medium (T medium) with or without the addition of 36 μM $FeSO_4$. For *E. coli* strains, T medium was supplemented with thiamine (10 μg/ml) and the L-amino acids arginine and leucine (40 μg/ml each). For citrate utilization assays, sodium citrate was added to T media at a final concentration of 10 mM.

Solid media containing a normal concentration of iron were made by using LB agar. Low-iron LB plates supplemented with 0.3% glucose were made after overnight treatment with Chelex-100 (Bio-Rad Laboratories, Richmond, Calif.), as described previously (10). Ampicillin (100 μg/ml), tetracycline (15 μg/ml), kanamycin (45 μg/ml), streptomycin (100 μg/ml), and 5-bromo-4-chloro-3-indolyl phosphate (XP) (40 μg/ml; Amresco) were added to the media as appropriate. XP is a chromogenic substrate that allows blue-white colony screening on agar plates for alkaline phosphatase activity.

Construction of plasmids. pSBC34 is a derivative of the broad-host-range plasmid pRK290 (8) modified to contain a gene fusion between the iron-regulated Shiga-like toxin I A gene (*slt*-IA) from *E. coli* and *phoA*, the *E. coli* gene for alkaline phosphatase. To construct pSBC34, the 5.5-kbp *Bam*HI fragment of pSC105 ($\Delta 2$) (5), encompassing the *slt*-I promoter, *slt*-IA fused to *phoA*, and the downstream kanamycin resistance marker, was purified by gel electroelution. This fragment was ligated into the unique *Bgl*II restriction site of pRK290, with selection for tetracycline resistance (encoded by pRK290) and kanamycin resistance. pRK290 and its derivatives are not self-transmissible, but can be mobilized at high frequency if supplied with mobilization functions on RP4 in *trans*. pSBC34 was therefore transformed into SM10, a derivative of *E. coli* C600 that contains a chromosomally integrated RP4-2 (Tc::Mu). SM10 (pSBC34) was conjugated with *V. cholerae* 0395, with double selection for tetracycline (encoded by pSBC34) and streptomycin (encoded by 0395) resistance.

Strain MBG40 contains a chromosomal gene fusion between *irgA* and *phoA*, constructed by Tn*phoA* mutagenesis (10). This fusion is contained within a 10-kbp *Bam*HI restriction fragment, extending from a *Bam*HI site in the chromosome approximately 5 kbp upstream of the fusion joint to the unique *Bam*HI site within Tn*phoA*, approximately 5 kbp downstream of the fusion joint. The *Bam*HI site within Tn*phoA* is located downstream of the end of *phoA* and the kanamycin resistance marker of the transposon.

This gene fusion was cloned into pUC19 as follows. Chromosomal DNA from MBG40 was digested with *Bam*HI and ligated into the *Bam*HI site on pUC19. The ligation mix was electroporated into DH5α according to methods described below. Colonies containing the correct clone (pMBG47) were isolated by selecting for ampicillin (carried on pUC19) and kanamycin (carried on Tn*phoA*) resistance. pMBG47 was then digested with *Bam*HI, and the 10-kbp fragment was separated from the parent plasmid by gel electroelution and ligated into the *Bam*HI site on pBR322, creating plasmid pMBG53 (Fig. 1). Presence of the correct insert in pMBG53 was confirmed by sequencing the junction between pBR322 and the Tn*phoA* end of the fragment. Subsequent subcloning of pMBG53 to make pMBG55 and pMBG59 was performed as shown in Fig. 1, pMBG59 contains an insert of approximately 1.8 kbp of chromosomal DNA fused to 2.9 kbp of Tn*phoA*. All plasmids were

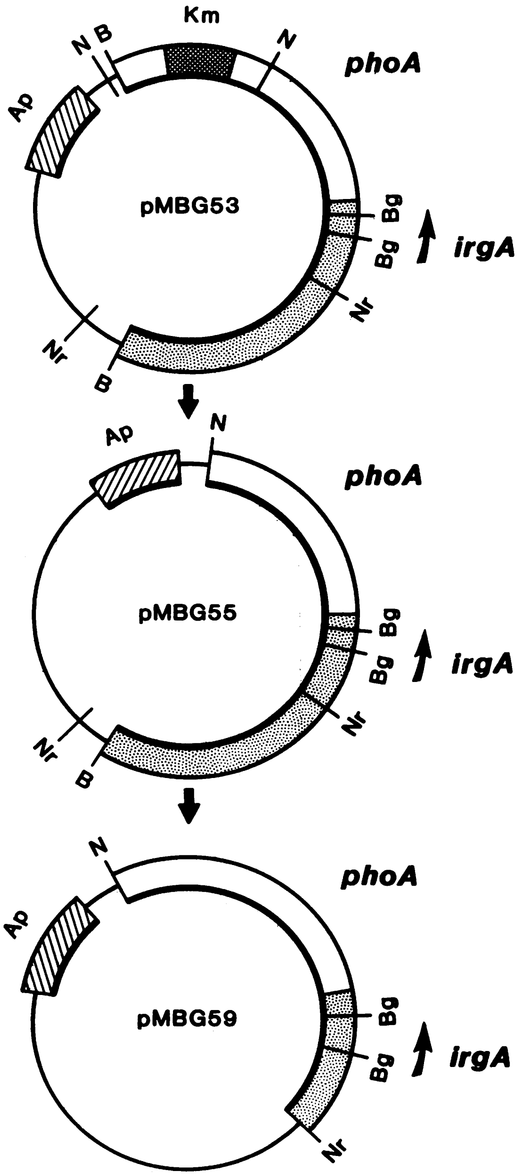


FIG. 1. Construction of plasmids used in this study. pMBG53 was constructed by cloning the *irgA'*-*'phoA* fusion gene from MBG40 into pBR322. pMBG55 was derived from pMBG53 by deletion of the fragment between the two *Nhe*I restriction sites. pMBG59 was derived from pMBG55 by deletion of the fragment between the two *Nru*I restriction sites. The lightly stippled bar represents MBG40 chromosomal DNA contained in the insert. The open bar represents DNA from Tn*phoA*. The darkly stippled bar (Km) represents the kanamycin resistance gene carried on Tn*phoA*. The hatched bar (Ap) represents the ampicillin resistance gene carried on pBR322. The arrow indicates the direction of transcription of *irgA*. B, *Bam*HI; Bg, *Bgl*II; N, *Nhe*I; Nr, *Nru*I.

propagated in *E. coli* CC118, which contains a chromosomal deletion of alkaline phosphatase.

Plasmid subclones pMBG57 and pMBG58 were derived from pMBG55 by deletion of portions of the chromosomal insert from each of two *Bgl*II restriction sites within the insert to the *Bam*HI restriction site at the junction of the insert with pBR322. Plasmid subclones pMBG102, pMBG105, pMBG109, and pMBG110 were similarly derived from pMBG59 by deletion of portions of the chromosomal insert from *Acc*I, *Nco*I, *Cla*I, and *Hinc*II restriction sites respectively within the insert to the *Nru*I site in pBR322. Plasmid subclone pMBG103 was derived from pMBG59 by deletion from the *Bal*I site within the chromosomal insert to the *Bal*I site in pBR322.

Genetic methods. Except as noted, plasmid constructs were transformed into *E. coli* strains and conjugated into *V. cholerae* 0395 by mobilization from the *E. coli* SM10 by standard techniques (31). Electroporation used in the transformation of pMBG47 into CC118 was performed in a Gene Pulser (Bio-Rad) according to the manufacturer's protocol. Electroporation conditions were 2,500 V at 25-μF capacitance, producing a time constant of 4.8 ms.

Assays. The enzymatic activity of alkaline phosphatase encoded on Tn*phoA* permitted screening of plasmid subclones containing the *irgA'-'phoA* gene fusion for the presence or absence of fusion gene expression under low-iron growth conditions by streaking *E. coli* CC118 containing subclones onto Chelex-LB agar containing XP. Colonies having a PhoA$^+$ phenotype are blue on XP, whereas colonies having a PhoA$^-$ phenotype are white.

Similarly, the enzymatic activity of alkaline phosphatase encoded on Tn*phoA* permitted the comparison of fusion gene expression in liquid media when *V. cholerae* MBG40 or *E. coli* CC118 carrying plasmid subclones of the gene fusion were grown in low versus high iron conditions. Strains were grown overnight in T medium with or without added $FeSO_4$. Alkaline phosphatase activity was determined as described previously (23) from measurement of hydrolysis of *p*-nitrophenyl phosphate (Amresco) by permeabilized cells. Activity was normalized to the A_{600} of the bacterial cells and defined as described by O'Callaghan et al. (27).

DNA and RNA analyses. DNA and RNA analyses, including preparation of DNA and RNA, restriction mapping, and Northern (RNA) blot analysis, were performed by standard molecular biological techniques (31). For Northern blot analysis, an equivalent amount of RNA, as calculated from optical density at 260 nm, was loaded into each lane. DNA sequencing was performed with the Sequenase DNA Sequencing Kit (United States Biochemical Corp., Cleveland, Ohio). Primer extension and primer extension sequencing were performed essentially as described by Miller et al. (24), except that the oligonucleotide primers were hybridized to RNA in 0.4 M NaCl and 40 mM PIPES [piperazine-*N*,*N'*-bis (2-ethanesulfonic acid)], pH 6.4, without formamide, at 60°C for 2 h. RNasin and avian myeloblastosis virus reverse transcriptase were obtained from Bethesda Research Laboratories Life Technologies, Inc.

Synthetic oligonucleotides for use both as probes for Northern blot analysis and as primers for DNA sequencing and primer extension were the generous gift of Brian Seed.

Protein analysis. Whole-cell and periplasmic proteins were prepared following growth in low- and high-iron media as described previously (16). Proteins were separated on a sodium dodecyl sulfate–12.5% polyacrylamide gel and transferred to a NitroScreen West membrane (Dupont, Boston, Mass.) with a semidry blotting apparatus (Hoefer Scientific Instruments, San Francisco, Calif.). Immunoreactive proteins were visualized by sequential incubation with polyclonal rabbit anti-PhoA antibody (a generous gift of John J. Mekalanos) and goat anti-rabbit immunoglobulin-conjugated alkaline phosphatase (Sigma Chemical Co., St. Louis, Mo.), followed by staining for alkaline phosphatase activity as described previously (25).

TABLE 1. Alkaline phosphatase assays in T media with and without added iron

Strain	Alkaline phosphatase activity (U/A_{600})	
	With iron	Without iron
V. cholerae 0395(pSBC34)	82	496
E. coli CC118(pMBG53)	63	277

Protein data base searches. Data base searches and protein alignments were done by using IBI-Pustell Sequence Analysis software (International Biotechnologies, Inc., New Haven, Conn.) to search the NBRF Protein Database (Release 19), using the FASTP algorithm for protein homology (19). The hydropathicity index profile of IrgA was calculated by the technique of Kyte-Doolittle (18), using IBI-Pustell Sequence Analysis software.

RESULTS

Reciprocity of iron regulation between *V. cholerae* and *E. coli*. We compared iron regulation of the *V. cholerae* gene *irgA* in an *E. coli* background with iron regulation of the *E. coli* gene *slt*-IA in a *V. cholerae* background, using alkaline phosphatase activities of the respective gene fusions following growth in low- and high-iron media (Table 1). These assays demonstrated reciprocal iron regulation of each gene fusion in the heterologous background, suggesting that *V. cholerae* and *E. coli* may share a common mechanism of iron regulation. We attempted to transform the *irgA'-'phoA* gene fusion on pMBG53 into an *E. coli fur*^{0} background, but were unable to establish a stable transformant, perhaps due to toxicity of the overexpressed fusion protein in the *fur*^{0} background.

Mapping of irgA by analysis of protein products and RNA transcripts. (i) Western blot analysis of IrgA'-'PhoA fusion protein. We performed Western blot (immunoblot) analysis of whole-cell and periplasmic proteins from MBG40 and 0395 following growth in low- and high-iron media, using an anti-PhoA antibody probe (Fig. 2). As expected, strain 0395 produced no immunoreactive protein and production of IrgA'-'PhoA by MBG40 was iron regulated. The IrgA'-'PhoA fusion protein had an apparent molecular mass of 70 kDa in whole-cell extracts (Fig. 2, lane 4), with several smaller immunoreactive species probably representing fusion protein breakdown products. The fusion protein was transported to the periplasmic space (Fig. 2, lane 3), but we were unable to resolve with certainty whether the largest band in periplasmic extracts was the same size as in whole-cell extracts or slightly smaller, as would occur with proteolytic processing by signal peptidase. Since the PhoA portion of the fusion protein has a predicted molecular mass of approximately 50 kDa, the upstream IrgA portion of the fusion protein would have a predicted molecular mass of approximately 20 kDa. Therefore, plasmid pMBG59, which contains approximately 1.8 kbp of DNA upstream of the fusion joint, should contain the 5' terminus of *irgA*, plus more than 1.2 kbp of additional upstream DNA.

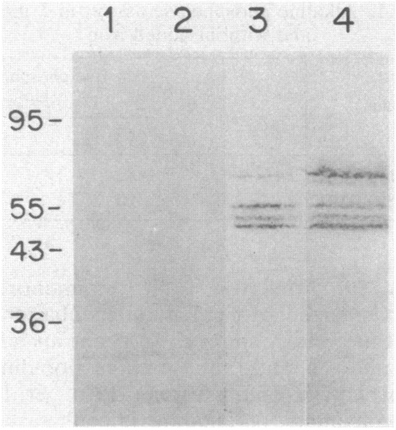


FIG. 2. Western blot analysis of proteins from MBG40 and 0395 prepared after growth in low- and high-iron media and probed with anti-PhoA antibody. Lanes: 1, whole-cell proteins of 0395 grown in high iron; 2, whole-cell proteins of MBG40 grown in high iron; 3, periplasmic extract of MBG40 grown in low iron; 4, whole-cell proteins of MBG40 grown in low iron. The sizes (in kilodaltons) of molecular weight standards are indicated on the left.

(ii) Verification that pMBG59 contains those sequences necessary for the expression and iron regulation of *irgA*. We determined alkaline phosphatase activities in low- and high-iron media of *V. cholerae* MBG40 and *E. coli* CC118 containing plasmid subclones pMBG47, pMBG53, pMBG55, and pMBG59. Alkaline phosphatase activity of MBG40 increased from 1 U after growth in iron-supplemented T medium to 126 U after growth in T medium without added iron, while activity of CC118 containing each of the plasmid subclones increased from approximately 70 U after growth in iron-supplemented T medium to approximately 200 U after growth in T medium without iron. Therefore, expression and iron regulation of the gene fusion on each of these plasmid subclones are similar, suggesting that they each contain the DNA sequences necessary for the expression and iron regulation of *irgA*. The difference in induction ratio of iron regulation between CC118 containing the plasmid subclones and MBG40 is probably due to a plasmid copy number effect and less efficient regulation of a *V. cholerae* gene in an *E. coli* background.

(iii) Deletion subcloning and mapping of fusion gene expression and iron regulation. Alkaline phosphatase assays were performed on CC118 containing plasmid deletion subclones pMBG109, pMBG110, pMBG105, pMBG102, pMBG103, pMBG58, and pMBG57 after growth in low- and high-iron T media (Fig. 3). Alkaline phosphatase activity of CC118 containing pMBG109 was comparable to that of pMBG59, but alkaline phosphatase activity of all subclones having <1.5 kbp of chromosomal DNA upstream of the fusion joint was minimal. This amount of upstream chromosomal DNA is substantially more than that predicted to encode IrgA based on Western blot analysis.

(iv) Northern blot analysis of the *irgA* transcript. RNA was prepared from strains MBG40 and 0395 following growth in low- and high-iron media and run on a Northern blot. The blot was probed with the restriction fragment located between the *Hin*dIII and *Sma*I sites on the insert of pMBG59 (Fig. 3). A single band is seen at 2.2 kb in RNA from 0395 grown under low-iron conditions (Fig. 4, lane 1b), and two less intense bands are seen at 2.5 and 2.1 kb in RNA from MBG40 grown under low-iron conditions (lane 2b). No bands are seen in either strain grown under high-iron conditions (lanes 1a and 2a), demonstrating that iron regulation occurs at the transcriptional level. After DNA sequencing, these results were confirmed by probing with a synthetic oligonucleotide shown to be entirely internal to *irgA* (data not shown).

To investigate the basis of the two bands seen in MBG40, we synthesized oligonucleotides complementary to *phoA* and to an open reading frame immediately downstream of *phoA*, as reported by Chang et al. (7). Probing additional Northern blots with these two oligonucleotides demonstrated that the two bands seen in MBG40 are not a result of two transcriptional start sites for *irgA*, but a result of two downstream transcription termination sites within the Tn*phoA* portion of the insert, one after *phoA* and the second located 400 bp further downstream, after the next open reading frame beyond *phoA* (data not shown). The 2.1-kb band from MBG40 is the transcript formed when termination occurs at a site located approximately 1.4 kb downstream from the fusion joint, and the 2.5-kb band is the transcript formed when termination occurs at a site approximately 1.8 kbp downstream from the fusion joint.

Comparison of the size of the IrgA'-'PhoA fusion protein by Western blot analysis (70 kDa) with the size of the smaller transcript seen in MBG40 by Northern blot analysis (2.1 kb) suggests that the RNA transcript for *irgA* is monocistronic and has a transcription start site located near the 5' *Bgl*II restriction site in the DNA upstream of the fusion joint (Fig. 3). Localization of the transcription start site to this region was confirmed by probing Northern blots of MBG40 RNA with synthetic oligonucleotides complementary to DNA sequence located between the *Hin*dIII and upstream *Bgl*II restriction sites (*irgA* transcript not seen) and to a DNA sequence located approximately 220 bp downstream of this *Bgl*II site (*irgA* transcript seen; data not shown). These data, coupled with the alkaline phosphatase assays of the plasmid deletion subclones (Fig. 3), suggest that approximately 900 bp of DNA upstream of the predicted transcription start site is required for production of an active IrgA'-'PhoA fusion protein.

(v) Northern blot analysis of *irgA* transcription from plasmid deletion subclones. RNA was prepared from CC118 containing the plasmid deletion subclones in parallel with MBG40 RNA, following growth in low-iron media, and run on a Northern blot. The blot was probed with the *Hin*dIII-*Sma*I restriction fragment used previously. Identical double bands, representing the two *irgA* transcripts, were seen in lanes containing RNA from MBG40 and RNA from CC118 containing pMBG53, pMBG59, and pMBG109, but were absent in all other lanes (Fig. 3). Thus, plasmid deletion subclones having full iron-regulated alkaline phosphatase activity also had the *irgA* transcripts, and those lacking alkaline phosphatase activity lacked the *irgA* transcripts. The presence of approximately equal amounts of RNA in the lanes of this Northern blot was confirmed for the plasmid deletion subclones by reprobing the membrane with a restriction fragment internal to the ampicillin resistance gene from pBR322 (data not shown). These data suggest that 900 bp of DNA upstream of the predicted transcription start site is necessary for transcription of *irgA*, rather than for a posttranscriptional event. That the *irgA* transcripts seen in *E. coli* CC118 containing the active plasmid deletion subclones are identical to the transcripts seen with the *irgA*::*phoA* fusion on the chromosome of *V. cholerae* MBG40 suggests that the results with CC118 (Fig. 3) are not

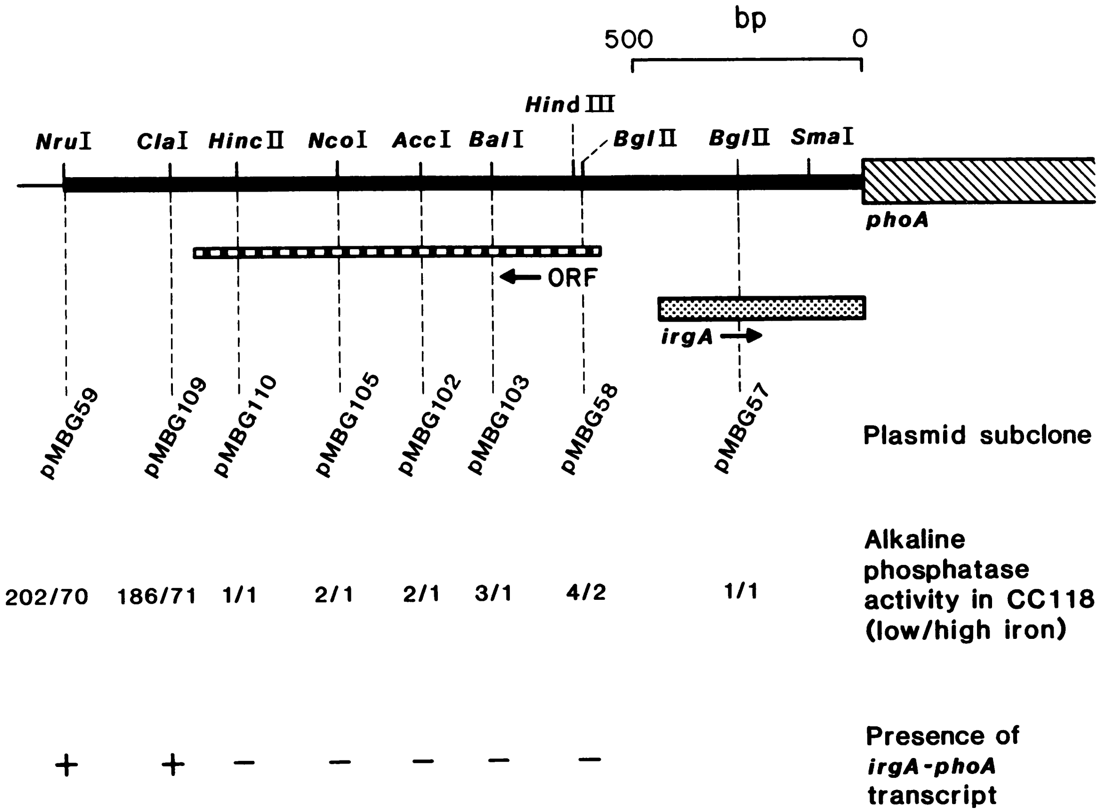


FIG. 3. Restriction map of *irgA* and upstream chromosomal DNA in pMBG59. The locations of restriction enzyme sites are indicated, as well as mapping of the plasmid subclones, assays of alkaline phosphatase activity of plasmid subclones, and the presence (+) or absence (−) of an *irgA'-'phoA* transcript from the plasmid subclones in CC118 after growth in low iron. The indicated *Nco*I and *Hin*dIII restriction sites are not unique in the insert. The solid bar represents MBG40 chromosomal DNA. The diagonally hatched bar represents Tn*phoA* DNA. The stippled bar indicates the position of *irgA* upstream of the fusion with *phoA*. The striped bar indicates the location of the 894-bp open reading frame (ORF) in inverse orientation to *irgA*.

a plasmid-related artifact or unique to *E. coli* rather than *V. cholerae*.

DNA sequence of the *irgA'-'phoA* fusion and upstream DNA. Figure 5 shows the DNA sequence of the chromosomal insert of pMBG59 from the *Cla*I site to the fusion with *phoA*. A 453-bp open reading frame, which is in-frame with *phoA*, begins at position 1077, representing the 5' portion of *irgA*. A Shine-Dalgarno sequence is located just upstream of the initiating methionine. The size of *irgA* upstream of the fusion joint with *phoA* is slightly smaller than we had predicted by Western blot of IrgA'-'PhoA and by Northern blot of the *irgA* transcript, but within reasonable experimental error.

The 900-bp region of DNA upstream of the *Bgl*II site that is required for transcription of *irgA* includes an 894-bp open reading frame in inverse orientation to *irgA* (Fig. 3). This open reading frame begins at position 932, 144 bp upstream of the *irgA* open reading frame, and terminates at position 39, just downstream of the *Cla*I restriction site.

Primer extension analysis of the start site of *irgA* transcription. Primer extension analysis of RNA from MBG40 grown in low iron was done with two distinct synthetic oligonucleotides complementary to DNA sequences located 20 and 64 bases downstream of the methionine start codon (data not shown). Both oligonucleotides identified the same approximate transcription start site, which is indicated by an asterisk in Fig. 5. A −10 box, reasonably homologous to the *E. coli* consensus sequence (14), was located approximately 7

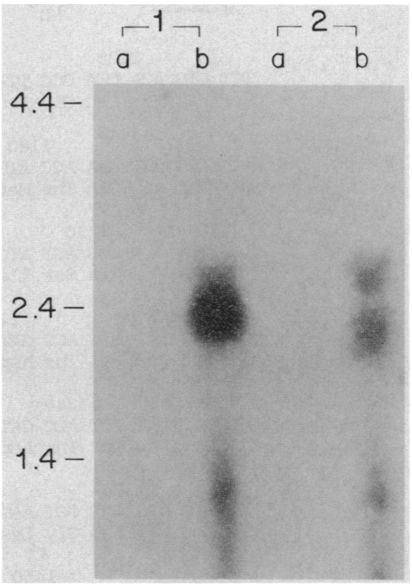
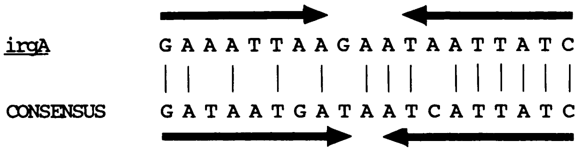


FIG. 4. Northern blot analysis of RNA from MBG40 and 0395, probed with a ^{32}P-labeled fragment of *irgA*. Lanes: 1a, 0395 grown in high iron; 1b, 0395 grown in low iron; 2a, MBG40 grown in high iron; 2b, MBG40 grown in low iron. The position of single-stranded RNA molecular weight markers (in kilobases) are indicated on the left.

```
 ClaI   10         20         30         40         50         60         70
ATCGATGATAAAAAATCCCGCTGCGGCGGGATTTTTTATTGCCACTCATCGGGCCTTGCTTGGCGGAGCG

        80         90        100        110        120        130        140
CATCAATAAATAGGCGCAGCCGAAGTGGTGACGACCGAGCGGATAGAAGCAGTTGATTTCTGTTGGCTG

       150  HincII           170        180        190        200        210
TGATTGCCATCCGTTGACGCAAGGAATGAGGCTGCCCGGATGCGCCGTTTCAAAACCATTGGCAAACCAA

       220        230        240        250        260        270        280
GTGGGAAGCAAACCAATACCACGACCTTTAGCAATCGCATCGGCTTGCATGGCAAGATTATCGCTTTGTA

       290        300        310        320        330        340        350
AACGACTCTCTAGTGCTGGCAGTGAATAACTGCCGAACTCTGGATGGTGCAGTTCAAGCTCCGCGCGCCG

       360          NcoI     380        390        400        410        420
ACAAGCAATAAAATCAATCCATGGGTGATGAATCAGCTCACGAGGATGGGTCGGTTTATCTCGATGGGCC

       430        440        450        460        470        480        490
AAATATTTGGGAGAGGCGTAAGTGGCATAGCGCCAATAGCCTAAGCGTTCTTTGCGATAACCCATGGGG

       500        510        520        530        540        550       AccI
CGGCGTGTTCAATCCAAATGATCAAATCGGGCTCAAACACCTCATCACTGTGTTGAAACTGGCTGAGTAG

__     570        580        590        600        610        620        630
ACGGATCTTCAATGTCGAATGCTGCTGCATAAACTCATCCAATACTTGGCTGAGCCAGCCGCGGATCAAA

       640        650        660        670        680        690        700
TTGGGGTGTACCACCAGCGTGAGTTCGCCAGTCACTTGATTGTTCAATTCTTGCAACGCTTCCTGACTTT

   BalI           720        730        740        750        760        770
TATTGGCCAGTTCAAGTAGTTGCTCCGAGTAAACCGCAAACACTTCTCCTGCTTTGGTGAGCGTTAAGCG

       780        790        800        810        820        830        840
GTTGCCTTGACGCATCAACAAGCTTTGTCCCAAGTCCTCTTCAAGTTGCGCCAAACGGCGACTCAGGGTG

       850        860        870        880        890        900    HindIII
GATTTAGGCTGTTCAAGCGCTTTGGCAGCGGCAGTCAGGCTCTTATGTTGGCAAAGCGCATGGAAAGCTT

       920 BglII 930        940        950        960        970        980
TTACGGCGCTGAGATCTTGCATAGGTATTTGACCCTTAAAGAATAATTACCACAGACGTTCCATATTTGG

       990       1000       1010       1020       1030          *      1050
ACCGAACTATTCCATGTGTCGATCTATCTCCAGTACAGAATATATGAATAATCCGCTTCTGAAATTAAGA
                                                 -10            ──────→

      1060       1070       1080         1090         1100
ATAATTATCATTTAAAGGAGTGGTAA ATG TCC AGA TTC AAT CCA TCC CCC GTC AGT TTA
←─────          SD          Met Ser Arg Phe Asn Pro Ser Pro Val Ser Leu

        1120       1130       1140        1150         1160
TCT GTG ACA CTA GGC TTA ATG TTT TCG GCT AGC GCT TTT GCT CAA GAC GCG ACG
Ser Val Thr Leu Gly Leu Met Phe Ser Ala Ser Ala Phe Ala Gln Asp Ala Thr

    1170       1180       1190        1200        1210
AAA ACG GAT GAA ACC ATG GTG GTC ACT GCG GCG GGA TAC GCG CAA GTG ATT CAA
Lys Thr Asp Glu Thr Met Val Val Thr Ala Ala Gly Tyr Ala Gln Val Ile Gln

          1230       1240       1250   BglII 1260        1270
AAT GCA CCA GCC AGT ATC AGT GTG ATT TCA AGA GAA GAT CTG GAA TCT CGC TAT
Asn Ala Pro Ala Ser Ile Ser Val Ile Ser Arg Glu Asp Leu Glu Ser Arg Tyr

        1280        1290       1300        1310        1320
TAC CGT GAT GTG ACC GAT GCG CTA AAA AGC GTA CCG GGT GTG ACA GTC ACC GGA
Tyr Arg Asp Val Thr Asp Ala Leu Lys Ser Val Pro Gly Val Thr Val Thr Gly

   1330       1340        1350        1360        1370
GGG GGC GAT ACT ACC GAT ATC AGC ATT CGT GGT ATG GGA TCA AAC TAT ACT CTT
Gly Gly Asp Thr Thr Asp Ile Ser Ile Arg Gly Met Gly Ser Asn Tyr Thr Leu

          1390       1400        1410        1420        1430
ATC TTG GTG GAT GGT AAG CGC CAA ACC TCA CGC CAG ACC CGT CCA AAC AGC GAT
Ile Leu Val Asp Gly Lys Arg Gln Thr Ser Arg Gln Thr Arg Pro Asn Ser Asp

    SmaI          1450        1460        1470        1480
GGC CCG GGC ATT GAG CAA GGT TGG TTA CCG CCA CTG CAA GCG ATT GAA CGT ATC
Gly Pro Gly Ile Glu Gln Gly Trp Leu Pro Pro Leu Gln Ala Ile Glu Arg Ile

           1500       1510        1520        1530
GAG GTG ATC CGT GGC CCG ATG TCT ACG CTG TAC GGC TCG GAT Gct gac ...
Glu Val Ile Arg Gly Pro Met Ser Thr Leu Tyr Gly Ser Asp Ala Asp
```



```
irgA        G A A A T T A A G A A T A A T T A T C
            | |   |   |   |   | | |   | | | | | |
CONSENSUS   G A T A A T G A T A A T C A T T A T C
```

FIG. 6. Nucleotide homology between the proposed iron regulatory region of *irgA* and the Fur box consensus sequence of *E. coli* (5, 6). Arrows refer to areas of interrupted dyad symmetry.

```
IrgA   132   Q A I E R I E V I R G P M S
                 | | | | | | | | | |
FepA   139   E M I E R I E V I R G P A R
```

FIG. 7. One region of homology between the peptide sequences of IrgA and FepA. The number of residues from the amino terminus of each precursor protein to the start of each sequence is indicated to the left.

bases upstream of the transcription start site, but no consensus −35 box was identified. A 19-bp interrupted dyad symmetric sequence, homologous to the Fur binding consensus sequence of *E. coli* (5, 6), was located immediately downstream of the transcription start site (Fig. 6; indicated by converging horizontal arrows in Fig. 5).

IrgA protein analysis. The predicted amino acid sequence encoded by the portion of *irgA* located upstream of the fusion joint with *phoA* is shown in Fig. 5.

(i) Hydropathicity index. The hydropathicity index profile of this portion of IrgA suggested a signal sequence at the amino terminus. A second stretch of hydrophobic residues further downstream, 24 amino acids in length, was suggestive of a possible transmembrane domain.

(ii) Homology of IrgA to *E. coli* ferrienterochelin receptor. The 151 amino acids at the amino terminus of IrgA were analyzed for homologous proteins in the NBRF Protein Database. The best match was to the *E. coli* ferrienterochelin receptor, FepA, an 80-kDa iron-regulated outer membrane protein (20). The optimized score between the amino terminus of IrgA and FepA was 185. There were several regions of amino acid homology at roughly the same positions in each protein, including one stretch of 10 consecutive identical amino acid residues (Fig. 7).

Citrate utilization by MBG40. Cell density and alkaline phosphatase activity were measured for MBG40 after growth in low- and high-iron media with and without the addition of citrate (Table 2). The addition of citrate to iron-deficient media partially restores growth of the mutant, demonstrating that the mutant is able to utilize citrate. Further, alkaline phosphatase activity of the *irgA'-'phoA* fusion was not regulated in response to citrate. These data suggest that the ferric citrate iron uptake system is intact in the mutant MBG40 and that *irgA* is not involved in ferric citrate uptake.

DISCUSSION

Using Tn*phoA* mutagenesis of *V. cholerae* 0395, we have previously constructed the mutant strain MBG40, which contains a Tn*phoA* insertion in the iron-regulated gene *irgA* and has reduced virulence in a newborn mouse model (10). By using the FASTP algorithm, the 151 amino acids at the amino terminus of IrgA were found to have significant homology to the *E. coli* ferrienterochelin receptor, FepA. FepA, an 80-kDa iron-regulated outer membrane protein, is quite similar in size and localization to the 77-kDa major iron-regulated outer membrane protein lost in MBG40. These findings suggest that IrgA may be the receptor for the related *V. cholerae* siderophore vibriobactin, and *irgA* is most likely the structural gene for this 77-kDa outer membrane protein. Mutant strain MBG40 has significantly reduced virulence in a newborn mouse 50% lethal dose assay, suggesting that the vibriobactin-iron assimilation system may be essential to full virulence of this organism. Sigel et al. (34) have previously described a *V. cholerae* mutant defective in the iron-vibriobactin transport system that had normal virulence in a mouse intestinal fluid accumulation assay. It is unclear whether these differences in observed virulence are based on the use of different virulence assays, the nature of the defects of the two mutants, or other as yet unknown factors. Our data also suggest that *irgA* does not have a role in the ferric citrate iron uptake system, the other iron assimilation system that has been described in *V. cholerae* (34), since growth of MBG40 was restored by the addition of citrate to low-iron media.

In this study, we have analyzed the molecular mechanism of iron regulation of the *V. cholerae* gene *irgA* and have demonstrated its similarity to the mechanism of iron regulation in *E. coli*. *irgA* is negatively regulated at the transcriptional level in the presence of sufficient iron. There is reciprocity of iron regulation of the *V. cholerae* gene *irgA* in *E. coli* and of the *E. coli* gene *slt*-IA in *V. cholerae*. Stoebner and Payne have previously demonstrated that the cloned *E. coli fur* gene, introduced in *trans*, restores normal iron regulation to a *V. cholerae* mutant that constitutively synthesizes vibriobactin and hemolysin, providing further evidence that a Fur-like mechanism may be operative in *V. cholerae* (36). We have identified a 19-bp interrupted dyad symmetric nucleotide sequence in the promoter region of *irgA*, with significant homology to the *E. coli* Fur binding consensus sequence, which may represent a binding site for a Fur-like protein in *V. cholerae*. In contrast to iron-regulated genes in *E. coli*, however, the *irgA* promoter lacks a consensus −35 box.

Our data suggest that approximately 900 bp of DNA upstream of the *irgA* transcription start site is required for transcription of *irgA*. Within this region is an 894-bp open reading frame, in inverse orientation to *irgA*, that begins just upstream of the *irgA* promoter and terminates just downstream of the *Cla*I restriction site (Fig. 3). When any portion of this open reading frame was deleted, transcription of *irgA* was lost (Fig. 3). Further experiments investigating the nature of this open reading frame are under way.

FIG. 5. Nucleotide sequence of *irgA* upstream of its fusion with *phoA*, the promoter region of *irgA*, and the large region of upstream DNA required for its transcription, starting at the *Cla*I site. The first five nucleotides of the *phoA* sequence after the fusion joint with *irgA* are indicated by lowercase letters at the end of the sequence. The locations of restriction enzyme sites are noted. The deduced amino acid sequence of the amino-terminal portion of IrgA is shown in three-letter code. The approximate transcription start site (*), −10 box (−10), and Shine-Dalgarno sequence (SD) are indicated. A 19-bp interrupted dyad symmetric element homologous to the Fur box of *E. coli* is indicated by inverted horizontal arrows below the sequence.

TABLE 2. Growth and alkaline phosphatase activity of *V. cholerae* MBG40 grown overnight in low- and high-iron T media, with and without added citrate

Growth conditions	A_{600}	Alkaline phosphatase activity (U/A_{600})
Low iron, without citrate	0.397	320
Low iron, with citrate	0.801	167
High iron, without citrate	1.431	1
High iron, with citrate	1.663	1

ACKNOWLEDGMENTS

We gratefully acknowledge John J. Mekalanos and Samuel I. Miller for many helpful discussions and Brian Seed for providing synthetic oligonucleotides.

This work was supported by a Public Health Service grant from the National Institute of Allergy and Infectious Diseases (AI27329) to S.B.C. Partial support was also provided by a Chugai Faculty Scholar Award to S.B.C. M.B.G. is a recipient of a National Research Service Award from the National Institutes of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Bagg, A., and J. B. Neilands.** 1985. Mapping of a mutation affecting regulation of iron uptake systems in *Escherichia coli* K-12. J. Bacteriol. **161:**450–453.
2. **Bindereif, A., and J. B. Neilands.** 1985. Promoter mapping and transcriptional regulation of the iron assimilation system of plasmid ColV-K30 in *Escherichia coli* K-12. J. Bacteriol. **162:**1039–1046.
3. **Bjorn, M. J., B. H. Iglewski, S. K. Ives, J. C. Sadoff, and M. L. Vasil.** 1978. Effect of iron on yields of exotoxin A in cultures of *Pseudomonas aeruginosa* PA-103. Infect. Immun. **19:**785–791.
4. **Bjorn, M. J., P. A. Sokol, and B. H. Iglewski.** 1979. Influence of iron on yields of extracellular products in *Pseudomonas aeruginosa* cultures. J. Bacteriol. **138:**193–200.
5. **Calderwood, S. B., and J. J. Mekalanos.** 1987. Iron regulation of Shiga-like toxin expression in *Escherichia coli* is mediated by the *fur* locus. J. Bacteriol. **169:**4759–4764.
6. **Calderwood, S. B., and J. J. Mekalanos.** 1988. Confirmation of the Fur operator site by insertion of a synthetic oligonucleotide into an operon fusion plasmid. J. Bacteriol. **170:**1015–1017.
7. **Chang, C. N., W.-J. Kuang, and E. Y. Chen.** 1986. Nucleotide sequence of the alkaline phosphatase gene of *Escherichia coli*. Gene **44:**121–125.
8. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **77:**7347–7351.
9. **Dubos, R. J., and J. W. Geiger.** 1946. Preparation and properties of Shiga toxin and toxoid. J. Exp. Med. **84:**143–156.
10. **Goldberg, M. B., V. J. DiRita, and S. B. Calderwood.** 1990. Identification of an iron-regulated virulence determinant in *Vibrio cholerae*, using Tn*phoA* mutagenesis. Infect. Immun. **58:**55–60.
11. **Griffiths, G. L., S. P. Sigel, S. M. Payne, and J. B. Neilands.** 1984. Vibriobactin, a siderophore from *Vibrio cholerae*. J. Biol. Chem. **259:**383–385.
12. **Hantke, K.** 1981. Regulation of ferric iron transport in *Escherichia coli* K12: isolation of a constitutive mutant. Mol. Gen. Genet. **182:**288–292.
13. **Hantke, K.** 1984. Cloning of the repressor protein gene of iron-regulated systems in *Escherichia coli* K12. Mol. Gen. Genet. **197:**337–341.
14. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11:**2237–2255.
15. **Henikoff, S., G. W. Haughn, J. M. Calvo, and J. C. Wallace.** 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA **85:**6602–6606.
16. **Hovde, C. J., S. B. Calderwood, J. J. Mekalanos, and R. J. Collier.** 1988. Evidence that glutamic acid 167 is an active-site residue of Shiga-like toxin I. Proc. Natl. Acad. Sci. USA **85:**2568–2572.
17. **Jonson, G., A.-M. Svennerholm, and J. Holmgren.** 1989. *Vibrio cholerae* expresses cell surface antigens during intestinal infection which are not expressed during in vitro culture. Infect. Immun. **57:**1809–1815.
18. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**109–132.
19. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227:**1435–1441.
20. **Lundrigan, M. D., and R. J. Kadner.** 1986. Nucleotide sequence of the gene for the ferrienterochelin receptor FepA in *Escherichia coli*. J. Biol. Chem. **261:**10797–10801.
21. **Manoil, C., and J. Beckwith.** 1985. Tn*phoA*: a transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA **82:**8129–8133.
22. **Mekalanos, J. J., D. J. Swartz, G. D. N. Pearson, N. Harford, F. Groyne, and M. de Wilde.** 1983. Cholera toxin genes: nucleotide sequence, deletion analysis and vaccine development. Nature (London) **306:**551–557.
23. **Michaelis, S., H. Inouye, D. Oliver, and J. Beckwith.** 1983. Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli*. J. Bacteriol. **154:**366–374.
24. **Miller, S. I., S. M. Landfear, and D. F. Wirth.** 1986. Cloning and characterization of a Leishmania gene encoding a RNA spliced leader sequence. Nucleic Acids Res. **14:**7341–7360.
25. **Miller, V. L., R. K. Taylor, and J. J. Mekalanos.** 1987. Cholera toxin transcriptional activator ToxR is a transmembrane DNA binding protein. Cell **48:**271–279.
26. **Neilands, J. B.** 1982. Microbial envelope proteins related to iron. Annu. Rev. Microbiol. **36:**285–309.
27. **O'Callaghan, C. H., A. Morris, S. M. Kirby, and A. H. Shingler.** 1972. Novel method for detection of β-lactamases by using a chromogenic cephalosporin substrate. Antimicrob. Agents Chemother. **1:**283–288.
28. **Pappenheimer, A. M., Jr.** 1955. The pathogenesis of diphtheria. Symp. Soc. Gen. Microbiol. **5:**40–56.
29. **Peterson, K. M., and J. J. Mekalanos.** 1988. Characterization of the *Vibrio cholerae* ToxR regulon: identification of novel genes involved in intestinal colonization. Infect. Immun. **56:**2822–2829.
30. **Pettis, G. S., T. J. Brickman, and M. A. McIntosh.** 1988. Transcriptional mapping and nucleotide sequence of the *Escherichia coli fepA-fes* enterobactin region. J. Biol. Chem. **263:**18857–18863.
31. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
32. **Sciortino, C. V., and R. A. Finkelstein.** 1983. *Vibrio cholerae* expresses iron-regulated outer membrane proteins in vivo. Infect. Immun. **42:**990–996.
33. **Sigel, S. P., and S. M. Payne.** 1982. Effect of iron limitation on growth, siderophore production, and expression of outer membrane proteins of *Vibrio cholerae*. J. Bacteriol. **150:**148–155.
34. **Sigel, S. P., J. A. Stoebner, and S. M. Payne.** 1985. Iron-vibriobactin transport system is not required for virulence of *Vibrio cholerae*. Infect. Immun. **47:**360–362.
35. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-negative bacteria. Bio/Technology **1:**784–791.
36. **Stoebner, J. A., and S. M. Payne.** 1988. Iron-regulated hemolysin production and utilization of heme and hemoglobin by *Vibrio cholerae*. Infect. Immun. **56:**2891–2895.
37. **Taylor, R. K., V. L. Miller, D. B. Furlong, and J. J. Mekalanos.** 1987. Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. Proc. Natl. Acad. Sci. USA **84:**2833–2837.