Positive transcriptional regulation of an iron-regulated virulence gene in *Vibrio cholerae*

(transcriptional activation/LysR family/Fur/IrgA/IrgB)

MARCIA B. GOLDBERG*, STEPHANIE A. BOYKO*, AND STEPHEN B. CALDERWOOD*†‡

*Infectious Disease Unit, Massachusetts General Hospital, Boston, MA 02114; and [†]Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02215

Communicated by A. M. Pappenheimer, Jr., November 19, 1990

ABSTRACT We have previously described a virulence gene in Vibrio cholerae (irgA) that is more than 850-fold regulated in response to iron. Negative regulation of irgA by iron occurred at the transcriptional level, and there was a dyad symmetric nucleotide sequence in the vicinity of the irgA promoter homologous to the Fur binding site in Escherichia coli. When ireA was cloned into E. coli, we showed that transcription of irgA required 900 base pairs of DNA upstream of the irgA promoter that contained an open reading frame in inverse orientation to irgA. In the present study, we show that this upstream region of DNA encodes a gene in inverse orientation to *irgA* (named *irgB*) that is also negatively regulated by iron. Insertional inactivation of irgB on the V. cholerae chromosome leads to loss of expression of a chromosomal irgA'-' phoA fusion (in which the primes indicate truncated genes), which is restored to normal by provision of irgB on a plasmid in trans. DNA sequencing of irgB shows that the protein product (IrgB) is homologous to the LysR family of positive transcriptional activators, and secondary structure analysis of IrgB predicts a helix-turn-helix DNA binding motif. The promoters of *irgB* and *irgA* are divergent but overlap each other and the previously defined Fur-binding site. We propose a model for iron regulation of irgA expression in V. cholerae. In the presence of sufficient iron, transcription of both irgA and irgB is negatively regulated by a Fur-like protein. In low iron conditions, negative regulation of transcription is removed. and production of IrgB leads to positive transcriptional activation of irgA. It seems likely that the high induction ratio of irgA expression under low- and high-iron conditions (850-fold) relates to the fact that its cognate positive transcriptional activator (irgB) is itself negatively regulated by iron.

Vibrio cholerae infection in humans may cause a severe dehydrating diarrhea. Illness occurs following ingestion of the organism in contaminated fresh or salt water. The bacteria must pass through the acid barrier of the stomach to reach the small intestine, where they penetrate the mucus gel and adhere to the brush border of intestinal epithelial cells by specific adhesins, including toxin coregulated pilus (TcpA) and other accessory colonization factors. At the intestinal brush border, V. cholerae produces a number of virulence factors, including cholera toxin, neuraminidase, hemolysin, and other extracellular secreted proteins. The genetic regulation of cholera toxin and TcpA in response to environmental stimuli has been described (1, 2). The genetic regulation of other virulence determinants in response to environmental signals is less well understood.

Many bacterial virulence determinants are regulated by the environmental concentration of free iron, with increased expression occurring under low-iron conditions. Examples of iron-regulated virulence determinants include diphtheria toxin of Corynebacterium diphtheriae (3), Shiga toxin of Shigella dysenteriae 1 (4), Shiga-like toxin I of enterohemorrhagic Escherichia coli (5), exotoxin A of Pseudomonas aeruginosa (6, 7), and several outer membrane proteins of Vibrio anguillarum (8) and Yersinia species (9).

We have described an iron-regulated virulence determinant in V. cholerae (10). Strain MBG40 of V. cholerae, which contains a TnphoA 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 (10). The deduced amino acid sequence of the amino terminus of IrgA is homologous to FepA, the ferrienterochelin outer membrane receptor of E. coli, suggesting that IrgA could be the iron-vibriobactin outer membrane receptor of V. cholerae (11). Expression of the irgA'-'phoA gene fusion (in which the primes indicate truncated genes), as measured by alkaline phosphatase activity, increases 850-fold in lowversus high-iron media (10). Negative regulation of irgA by iron occurs at the transcriptional level, and its promoter region, defined by primer extension, contains a dyad symmetric sequence homologous to dyad elements in E. coli that bind the Fur protein, an iron-responsive repressor of transcription (11).

When the irgA'-'phoA gene fusion was cloned from V. cholerae onto a plasmid in E. coli, expression of alkaline phosphatase activity and transcription of the gene fusion required an additional 900 base pairs (bp) of DNA upstream of the *irgA* promoter (11); this region of DNA, which we call *irgB*,[§] contains an open reading frame in inverse orientation to *irgA* (Fig. 1). We wished to test the model in which this open reading frame encodes a positive activator for *irgA* transcription.

MATERIALS AND METHODS

Bacterial Strains. Bacterial strains and plasmids used in this study are shown in Table 1. *V. cholerae* strain MBG260 and plasmids pSBC45 and pMBG111 were constructed as described below.

Media. Luria-Bertoni (LB) medium with or without the addition of the iron chelator 2,2-dipyridyl (final concentration, 0.2 mM) was used to evaluate the effect of iron concentration on gene expression as described (10).

Construction of plasmids. Strain MBG40 contains a chromosomal gene fusion between *irgA* and *phoA*, constructed by Tn phoA mutagenesis (10). Plasmid pMBG59 contains a subclone of the intact *irgA*::Tn phoA gene fusion from MBG40 into plasmid pBR322. Plasmid pMBG59 also contains all of the open reading frame upstream of *irgA* (subsequently named *irgB*) (see *Results*) (Fig. 1).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

[‡]To whom reprint requests should be addressed.

[§]The sequence reported in this paper has been deposited in the GenBank data base (accession no. M55988).



Plasmids pSBC45 and pSBC46 are derivatives of pA-CYC184, a plasmid encoding chloramphenicol resistance. Plasmids pSBC45 and pSBC46 were constructed by isolating the Nru I-Sma I fragment of plasmid pMBG59 by electroelution from a gel and ligation into Nru I-digested pA-CYC184; the two plasmids differ only in the orientation of the inserted fragment, which contains the intact gene *irgB* (Fig. 1; see *Results*).

Plasmid pMBG111 was derived from pGP704, a broadhost-range plasmid containing the ampicillin-resistance gene from pBR322, the mobilization domain of plasmid RP4 (14), the origin of replication from plasmid R6K (15), and a polylinker from phage M13 tg131 (Amersham). Plasmid pGP704 was a gift of Gregory D. N. Pearson and is itself derived from plasmid pJM703.1 (13). Plasmid pGP704 and its derivatives are able to replicate only in strains containing the *pir* gene, which encodes the π protein necessary for the function of the R6K origin (15). To construct pMBG111, a 676-bp *HincII-Bgl* II fragment of pMBG59 internal to *irgB* (Fig. 1) was ligated into the *Eco*RV and *Bgl* II sites of the pGP704 polylinker.

Genetic Methods. V. cholerae strain MBG260, which contains an insertion mutation in *irgB*, was constructed from strain MBG40 in the following manner. Plasmid pMBG111 was transferred from strain SY327 λ pir into SM10 λ pir by transformation. SM10 λ pir contains a chromosomallyintegrated RP4-2 (Tc::Mu), which encodes trans-acting fac-

FIG. 1. Partial restriction map of the chromosomal DNA upstream of the TnphoA fusion junction in V. cholerae strain MBG40, subcloned as described (11) to make plasmid pMBG59. Locations of relevant restriction enzyme sites are indicated, as well as fragments used to construct plasmids pMBG111 and pSBC45. The solid bar represents MBG40 chromosomal DNA. The diagonally cross-hatched bar represents TnphoA DNA. The stippled bar indicates the position and direction of transcription of *irgA* upstream of the fusion with phoA. The striped bar indicates the position and direction of transcription of irgB.

tors necessary to mobilize pGP704 derivatives into a broad range of recipients without RP4 itself being transferred (13). SM10 λ pir containing pMBG111 was conjugated with MBG40, with double selection for ampicillin resistance (encoded by pMBG111) and streptomycin resistance (encoded by MBG40). Because MBG40 does not contain the pir gene, pMBG111 is unable to replicate in this strain, so that doubly resistant colonies arise by homologous recombination between the internal fragment of irgB on pMBG111 and the corresponding chromosomal gene on the recipient, causing insertional inactivation of irgB. To confirm that chromosomal integration occurred within irgB on the recipient, we performed Southern hybridization of chromosomal DNA digests with EcoRV, an enzyme that does not cut within either irgB or pMBG111. The blot was probed with the HincII-Bgl II fragment of *irgB*, radioactively labeled by random primer extension with a commercial kit (Prime Time, International Biotechnologies).

Plasmids were transformed into *E. coli* strains by standard techniques (16). Plasmids pACYC184 and pSBC45 were introduced into *V. cholerae* strains by electroporation by using the protocol of the manufacturer (Gene Pulser, Bio-Rad), with the exception of substitution of 2 mM CaCl₂ as the buffer for resuspending cells during preparation, rather than water or Hepes buffer. Plasmid content of the electroporants was confirmed by restriction enzyme digestion of plasmid

Table 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant genotype or phenotype	Ref. or source
V. cholerae strains		
0395	Sm ^r	27
MBG40	irgA::TnphoA, Sm ^r Km ^r	10
MBG260	irgA::TnphoA irgB::pMBG111,	This study
E. coli strains		
CC118	Δ (ara leu)7697 Δ (lac)X74 araD139 phoA Δ 20 galE galK thi rpsE rpoB argE(Am) recA1	12
SY327 λ pir	Δ(lac pro) nalA recA56 araD argE(Am) λ pirR6K	13
SM10 λ pir	thi thr leu tonA lacY supE recA::RP4-2-Tc::Mu λ pirR6K, Km ^r	13
Plasmids		
pMBG59	pBR322 with 4.6-kb irgB irgA::TnphoA insert from MBG40	11
pJM703.1	oriR6K mobRP4, Ap ^r	13
pGP704	pJM703.1 with 1.5-kb <i>Sph</i> I– <i>Pvu</i> II deletion and 75-bp insert of M13 tg131 polylinker	G. D. N. Pearson
pMBG111	pGP704 with 676-bp HincII-Bgl II insert of pMBG59	This study
pSBC45	pACYC184 with 1.6-kb Nru I-Sma I insert of pMBG59	This study

Ap^r, ampicillin resistance; Km^r, kanamycin resistance; Sm^r, streptomycin resistance.

minipreps, analyzed by agarose gel electrophoresis as described (11).

Assays. The enzymatic activity of alkaline phosphatase encoded on Tn phoA permitted the comparison of fusion gene expression when strains were grown in low- versus high-iron media. Strains were grown overnight in LB medium with or without added 2,2-dipyridyl. Alkaline phosphatase activity was determined as described (10).

DNA and RNA Analysis. Analysis of DNA and RNA, including DNA and RNA preparation, restriction mapping, DNA sequencing, RNA (Northern) blot analysis, and primer extension were performed as described (11). For Northern blot analysis, an equivalent quantity of RNA, as calculated from OD_{260} , was loaded into each lane.

Synthetic oligonucleotides used as probes for Northern blot analysis and as primers for DNA sequencing and primer extension were the gift of Brian Seed (Massachusetts General Hospital).

Protein Analysis and Protein Data Base Searches. Protein analysis and protein data base searches were performed by using IBI-Pustell sequence analysis software (International Biotechnologies). The hydropathicity index profile of IrgB was calculated by the formula of Kyte–Doolittle (17). The protein secondary structure prediction of IrgB was calculated by the algorithm of Chou–Fasman (18). Data base searches and protein alignments were performed by searching the Protein Identification Resource, National Biomedical Research Foundation data base (release 19) using the FASTP algorithm for protein homology (19).

RESULTS

Northern Blot Analysis of the Open Reading Frame Upstream of irgA. To determine whether an RNA transcript was associated with the upstream open reading frame, and if present, to determine the size of the transcript and whether or not transcription was regulated by iron, we performed Northern blot analysis of RNA prepared from strains 0395 and MBG40 following growth in low- and high-iron media (Fig. 2). The blot was probed with an oligonucleotide complementary to the DNA sequence near the 5' terminus of the open reading frame. A single band of approximately 1.1 kilobases (kb) in size was seen in RNA prepared from 0395 and MBG40 grown under low-iron conditions (Fig. 2, lanes 2 and 4), showing that a transcript was associated with the open reading frame. No bands were seen in either strain grown under high-iron conditions (Fig. 2, lanes 1 and 3), showing that transcription of the open reading frame was negatively regulated by iron. The size of the transcript, in relation to the size of the open reading frame (894 bp, see below), suggests



FIG. 2. Northern blot analysis of RNA from 0395 and MBG40 probed with an oligonucleotide complementary to the 5' terminus of *irgB*. Lanes: 1, 0395 grown in high-iron medium; 2, 0395 grown in low-iron medium; 3, MBG40 grown in high-iron medium; 4, MBG40 grown in low iron medium. The positions of single-stranded RNA molecular weight markers (in kilobases) are indicated on the left.

that the transcript is monocistronic. The iron-regulated gene encoded by this open reading frame was designated *irgB*.

Trans-Complementation of *irgB* **and** *irgA*. We have previously demonstrated, by deletion subcloning of pMBG59 in an *E. coli* background, that deletion of any portion of *irgB* eliminates transcription of *irgA* (11). We wished to examine whether *irgB* could restore expression of *irgA* in trans.

To examine trans-complementation of irgB and irgA in the V. cholerae background, we constructed the irgB mutant strain MBG260 from MBG40. Highly expressed, ironregulated alkaline phosphatase activity seen in strain MBG40 was almost completely eliminated with the introduction of the irgB mutation to make strain MBG260 (Table 2). Ironregulated, alkaline phosphatase activity was completely restored by the introduction of *irgB* in trans on plasmid pSBC45 (Table 2) or pSBC46 (data not shown), while the introduction of the vector pACYC184 had no effect (data also not shown). Strain 0395, which does not contain an irgA'-' phoA fusion, had negligible alkaline phosphatase activity, with or without the introduced plasmids (data not shown). The higher alkaline phosphatase activities seen in MBG260(pSBC45) and MBG40(pSBC45), as compared with MBG40 (Table 2), may result from the high copy number of *irgB* carried on pSBC45. These data suggest that irgB is a trans-acting factor that positively regulates irgA.

DNA Sequence of *irgB* and Deduced Protein Sequence of IrgB. Fig. 3 shows the DNA sequence of the chromosomal insert of pMBG59 (reading 5' to 3' from right to left in Fig. 1), starting approximately 60 bp beyond the transcription start site of *irgA* (11) and extending up to the *Cla* I site of pMBG59, including the 894-bp open reading frame of *irgB*. A Shine-Dalgarno sequence is indicated just upstream of the initiating methionine (20). A perfect inverted repeat, suggestive of a transcription terminator, is indicated just beyond the termination codon of the *irgB* open reading frame. The deduced protein sequence for IrgB is shown below the nucleotide sequence.

Primer-Extension Analysis of the Start Site of irgB Transcription. Primer-extension analysis of RNA prepared from MBG40 and 0395 following growth in low-iron medium was done with a synthetic oligonucleotide complementary to the DNA sequence located between 3 bases upstream and 17 bases downstream of the methionine start codon (data not shown). The same transcription start site was identified in both MBG40 and 0395 and is indicated by an asterisk in Fig. 3. A promoter homologous to the E. coli consensus sequence (21) was located upstream of the transcription start site (Fig. 3). The 19-bp interrupted dyad symmetric sequence that is homologous to the Fur binding consensus sequence of E. coli (5, 22) and is located immediately downstream of the *irgA* transcription start site (11) also overlaps the irgB transcription start site and -10 box (Fig. 3). Fig. 4 shows the overlapping but divergent *irgA* and *irgB* promoters and the location of the Fur-like box in relation to each.

Table 2.	Complementation	of irgB	and	irgA'-'	phoA	in
V. choler	ae					

Strain MBG40 MBG260 MBG40(pSBC45)	Alkaline phosphatase activity, units						
Strain	Low iron	High iron					
MBG40	683	3					
MBG260	4	1					
MBG40(pSBC45)	1420	11					
MBG260(pSBC45)	1573	8					

Alkaline phosphatase assays were in LB medium with and without added 2,2-dipyridyl which chelates iron. Units of alkaline phosphatase were normalized by using the absorbance of the cell culture at 600 nm.

		10		20		-	30		40			50	_*.	60			70	80
GGGG	GAIG	GATT	GAAI	-3	5	TIAC	CACI	œn	1.944	IGAL	MALT I	ATTC		TTIC	AGAN	GOGG	ATTA	TICATATA
	-10																	
											-							
	-	90 1022a	C3778	100		1 (2)10	10 Сарт		120	~~~	[Геге	.30 1723 A	~~~~	140		1	50	160
1101	GIAL	1001				CALG	Grant	AGII		unn		Gunn	COIC	1919	GIAA	TIAL	ICH	SD
		170		Bal	<u>11</u>			1	.90 <u>H</u>	indI	11	200		_	210		_	220
AAAD	ACCI	Met	Gln	A GALL		Ser	alla	Val	AAA 13/10	a GCT	The	Hie	GCG 1 Alla		- TGC	CAA	Ula	ANG AGC
																GII		- Lys Ser
	2	30			240			25	60		2	60			270			280
CIG	ACT	200	GCT	000	YYY	COC	CIT	GAA	CAG	CT	YYY	TCC	W C	CIG	AGT	œc	CGT	TTG GCG
Ter .	inr	ALA	ALA	ALA	Був	AIS	Leu	GIU	GIN	Pro	LYS	ser	Inr	Leu	Ser	Arg	Arg	Leu Ala
	2	:90			300		н	lindi	II		2	20			330			340
CAA	CTT	GAA	GAG	GAC	TTG	GGA	CAA	AGC	TIG	TTG	ATG	OUT	CAA	GGC	AAC	œc	TTA	ACG CTC
Gln	Leu	Glu	Glu	уvado	Leu	Gly	Gln	Ser	Leu	Leu	Met	Arg	Gln	Gly	λen	λrg	Leu	Thr Leu
		50			260			2.	20						200		-17	400
ACC	,	GCA	GGA	GAA	GTG	TTT	GOG	GTT	TAC	TOG	GAG	CAA	СТА	CTT	GAA	circ ¹	an	AAT AAA
Thr	Lys	Ala	Gly	Glu	Val	Phe	Ala	Val	Tyr	Ser	Glu	Gln	Leu	Leu	Glu	Leu	Ala	Asn Lys
																		-
2011	ac'	10 CN	~		420	~~~		4	30	~	~	140	~	~	450		_	460
Ser	Gin	Glu	Ala	Les	Gln	Glu	Ten	Asn	Asn.	Gln	Val	Thr	Glv	Glu	Tan	Thr	Ten	Val Val
														014				vur vur
		470			480			- 49	90		1	500			510			520
CAC	<u>@</u>	AAT	TIG	ATC		000	TOG	CIC	AGC	CAA	GIA	TIG	GAT	GAG	TIT	ATG	CAG	CAGCAT
ша	PIO	ABC	Lieu	шe	Arg	GIY	пp	Lieu	Ser	GIN	vai	Lieu	vab	GIU	me	MEC	GIN	GIN HIS
	1	530				ΩCI	_	5	50		:	560			570			580
TCG	ЪCЪ	TTG	ANG	ATC	CT	CTA	CIC	AGC	CAG	TIT	CAA	CAC	AGT	GAT	GAG	GIG	TTT	GAG CCC
Ser	Thr	Leu	Lys	Ile	Arg	Leu	Leu	Ser	Gln	Phe	Gln	His	Ser	Ъзр	Glu	Val	Phe	Glu Pro
		590			600			6	10	,	iT				630			640
GAT	TTG	ATC	ATT	TGG	ATT	GAA	CAC	œ	õœ	œ	ATG	GGT	TAT	œc	222	GAA	œc	TTA GGC
Asp	Leu	Ile	Ile	Trp	Ile	Glu	His	Ala	Ala	Pro	Met	Gly	Tyr	Arg	Lys	Glu	Arg	Leu Gly
								-										
тат	m 22	650 mm	TAT	\sim	660	mac	~	6	⁷⁰			-	~	~~~	690 m	~~~		700
Tvr	Tro	Arg	Tvr	Ala	Thr	Tvr	Ala	Ser	Pro	Ivs	Tvr	Iau	Ala	His	Ana	Aso	Lvs	Pro Thr
-4-		,	-4-			-4-				-2-	-1-				,		-1-	
		710			720			_N	ΩI			740			750			760
CAT	œ	CGT	GAG	CIG	ATT	CAT	CAC	ΩCA.	TOG	ATT	GAT	TTT	ATT	GCT	TGT	ŝ	000	GOG GAG
nis	PIO	Arg	GIU	neu	116	nis	nus	PIO	пр	шe	vab	Me	шe	Ala	Сув	Arg	Arg	Ala Giu
		770			780			7	90			B00			810			820
CIT	GAA	CIG	CAC	CAT	CCA	GAG	TTC	GCC	AGT	TAT	TCA	CIG	CCA.	GCA	CTA	GAG	AGT	ogt tta
Leu	Glu	Leu	His	His	Pro	Glu	Phe	Gly	Ser	Tyr	Ser	Leu	Pro	Ala	Leu	Glu	Ser	Arg Leu
		830			840			8	50			860			870			880
CAA	AGC	GAT	AAT	CIT	ŝ	ATG	CAA	õ	GAT	GCCG	ATT	GCT	AAA	GGT	œT	GGT	ATT	GGT TIG
Gln	Ser	Asp	Asn	Leu	Ala	Met	Gln	Ala	Asp	Ala	Ile	Ala	Lys	Gly	Arg	Gly	Ile	Gly Leu
								-										
~	~	890	~~		900	2.20	007	9	10	200	~	920	m	~	930	ŝ	ATT	940
Ten	Pm	Thr	100 Tm	Phe	Ala	λsn	Glv	Phe	Glu	Thr	Ala	His	Pro	Glv	Ser	Leu	Tle	Pro Cvs
							1							1				
Hind	ш	950			960			9	70			980			990	_	-	1000
GTC	AAC	GGA	TGG	CAA	TCA	CAG	CCA D~~	ACA	GAA	ATC	AAC	TGC	TIC	TAT	00G	CIC	GGT	OUT CAC
val	ASN	GIY	np	GIN	Ser	GIU	PIO	mr	GIU	шe	ASD	Cys	me	TÅL	10	neu	σιγ	vid ute
	1	010			1020			10	30		1	040		:	1050			1060
CCA	CIT	œ	CIG	œc	CTA	TIT	ATT	GAT	GOG	CIC	œ	CAA	GCA	AGG	∞	GAT	GAG	TGG CAA
Pro	Leu	Arg	Leu	Arg	Leu	Phe	Ile	Asp	Ala	Leu	Arg	Gln	Ala	Arg	Pro	Asp	Glu	Trp Gln
	10	70		1080		10	90	~	laT									
TAA	AÃA	ATCO	œœ	GCAG	œœ	ATTT	TTTA	TCAT	CGAT									
			_			-	-											
			-	-														

FIG. 3. Nucleotide sequence of the chromosomal DNA in pMBG59 (reading 5' to 3' from right to left in Fig. 1), starting downstream of the start site of *irgA* transcription and extending up to the *Cla* I restriction site in pMBG59. The locations of restriction enzyme sites are noted. The deduced amino acid sequence of IrgB is shown in three-letter code. The approximate start site of transcription (*), the -10 and -35 boxes of the promoter, and the putative Shine-Dalgarno sequence (SD) are indicated. A 19-bp interrupted dyad symmetric element homologous to the Fur box of *E. coli* is indicated by heavy horizontal arrows below the sequence in the vicinity of the promoter. The termination codon of IrgB is indicated (. . .), followed by a probable transcription terminator (light horizontal arrows at end of sequence).

IrgB Protein Analysis. The predicted amino acid sequence of IrgB is shown below the nucleotide sequence in Fig. 3.

helix-turn-helix

IrgB	(1)	MODISANNAFHAICOHKSITAAANAILEOPKSTISRILAGIEEDIGGSILMAOGNRITIIMAGEV
IlvY	(1)	MDIRDINTELILAESRHFGRSARAMHVSPSTLSRQIQRLEEDLOGHLFVRDNRTVTLTBAGEE
LysR	(3)	AVALRHIEIFHAVMTAGELTEAAHLIHTSQPTVSRELARHEKVIGEACHHHVVGER
AmpR	(5)	YLELNSLRAFEMAARHLSETHAAIELNVTHSAISOHVKTLEOHUNCOUFVEVSRGLATTEGEN
NodD	(5)	GIDINILVALDAIMTERNETAAARKINLSQPANSAAIARIRSYFRDELFTMRGRELVEIPGABA
CysB	(3)	LQCTRYI-V-EVVNHNLNVSSTARECTLYTSQPGTSKQVRMLEDETGIQIFSRSCKHTQVTPACQ

							- 35			-10	*
5'	GGG	GGA	TGG	ATT	GAA	TCT	GGA	CAT	TTACCACTCCTTTAAAT	GATAATTA	TTCTTAATTTCA
3'	CCC	CCT	ACC	TAA	CTT	AGA	CCT	GTA	AATGGT <u>GAGGA</u> AATTTA	CTATTAAT	AAGAATTAAAGT
	Pro	Ser	Pro	Asn	Phe	Arg	Ser	Met	SD		*

GAAGCGGATTATTCATATATTCTGTACTGGAGATAGATCGACACATCGAATAGTTCGGTCCAAATATGG CTTCGCCTAATAAGTATATAAGACATGACCTCTATCTAGCTGTGTACCTTATCAAGCCAGGTTTATACC -10

	IrgB>											
SD	Met	Gln	Asp	Leu	Ser	Ala	Val					
AACGTCTGTGGTAATTATTCTTTAAGGGTCAAATACCT	ATG	CAA	GAT	CTC	AGC	GCC	GTA	3′				
TTGCAGACACCATTAATAAGAAATTCCCAGTTTATGGA	TAC	GTT	CTA	GAG	TCG	CGG	CAT	5'				

FIG. 4. Detail of the overlapping, divergent promoters of *irgA* and *irgB*. The promoter of *irgB* (-35, -10), start site of transcription (*), Shine-Dalgarno sequence (SD), and open reading frame are indicated on the upper strand, while the corresponding features of *irgA* are noted on the bottom strand. Note that *irgA* does not have a consensus -35 box. The dyad symmetric element homologous to the *E. coli* Fur-binding site is enclosed within a box.

Hydropathicity profile. The hydropathicity profile of IrgB showed no stretches of hydrophobic residues that would be suggestive of either a signal sequence or a transmembrane domain (data not shown).

Homology of IrgB to the LysR family of positive transcriptional activators. Comparison of IrgB to the Protein Identification Resource, National Biomedical Research Foundation (release 19) data base demonstrated significant homology between IrgB and the LysR family of positive transcriptional activator proteins in bacteria (23). The best match in this family was to the E. coli positive activator protein IlvY; the optimized score between IrgB and IlvY was 217. The regions of highest homology between IrgB and the LysR family were near the amino terminus and were at roughly the same positions in each protein. The homology near the amino terminus of IrgB with several members of the LysR family is shown in Fig. 5. All of the proteins shown are of approximately the same size (IrgB is 298 amino acids long; IlvY, LysR, AmpR, NodD, and CysB are 297, 300, 291, 300, and 301 amino acids long, respectively).

Prediction of the secondary structure of IrgB. The secondary structure of IrgB was predicted by using the Chou-Fasman algorithm (18). A helix-turn-helix motif was seen in the same region of IrgB as in the other members of the LysR family (Fig. 5) (23).

DISCUSSION

The expression of many bacterial virulence determinants is regulated by the concentration of iron in the environment, with increased expression occurring under low-iron conditions. We have previously described such an iron-regulated virulence determinant in V. cholerae named irgA (10). V. cholerae strain MBG40, which contains a TnphoA gene fusion with the amino terminus and promoter of irgA, shows several-hundred-fold regulation of alkaline phosphatase activity in response to iron (10). We demonstrated that transcription of irgA was negatively regulated by iron and that there was a 19-bp dyad symmetric sequence homologous to E. coli Fur-binding sites present in the promoter region of irgA (11). Questions about the regulation of irgA expression,

(64)

(63) FIG. 5. Homology between the amino terminus
(66) of IrgB and several members of the LysR family of positive transcriptional activators. Amino acids are
(68) identified in single letter code and identical residues
(68) are enclosed in boxes. The conserved helix-turnhelix domain of these proteins is indicated above
(64) the sequences.

however, remained. (i) Why was the induction ratio of *irgA* expression in relation to iron so large compared with ironregulated genes in other bacteria? (*ii*) Why was there no consensus -35 box in the *irgA* promoter, a feature seen in genes requiring a positive activator? (*iii*) Why was 900 bp of DNA upstream of the *irgA* promoter required for transcription of the *irgA'-' phoA* fusion when cloned on a plasmid into *E. coli*?

The data in the present study suggest that the open reading frame upstream of irgA, now named irgB, is a positive activator of *irgA* transcription that is itself negatively regulated by iron at the transcriptional level (Fig. 2). We have previously shown that E. coli CC118 (pMBG59) expresses iron-regulated alkaline phosphatase activity, while deletion subclones into any portion of irgB, which still maintain an intact irgA'-' phoA fusion, lose alkaline phosphatase activity (11). We examined the effect of irgB on irgA expression in single copy in the V. cholerae background by introducing an irgB mutation into the chromosome of strain MBG40. The resulting mutant, MBG260, showed marked reduction of irgA'-' phoA fusion activity that was fully restored by complementation with irgB in trans (Table 2). In addition, transcomplementation with *irgB* in V. cholerae also restored full iron regulation to expression of the irgA'-' phoA fusion.

The deduced amino acid sequence of IrgB is homologous to the LysR family of bacterial transcriptional activator proteins (Fig. 5). Homologous proteins have been identified as members of this family from several species of Enterobacteriacae, including *E. coli*, *Salmonella typhimurium*, and *Enterobacter cloacae* (23). All proteins in this family have a helix-turn-helix motif near the amino terminus, and several of the proteins are transcribed divergently from an operon or gene that is regulated by the protein (23). IrgB is similar in both of these respects.

The genes *irgA* and *irgB* are divergently transcribed from overlapping promoters that contain a single 19-bp dyad symmetric element homologous to E. coli Fur-binding sites (Fig. 4), suggesting that a V. cholerae Fur-like protein might simultaneously repress the transcription of both irgA and *irgB* in the presence of iron. We have shown previously that the amino-terminal portion of IrgA is homologous to FepA, the E. coli ferrienterochelin receptor (11). irgA is similar to fepA in that its transcription is negatively regulated by iron and there is a Fur-like box in its promoter region. In addition, each of these genes has another iron-regulated gene that is divergently transcribed from a promoter that overlaps the same Fur box (24). irgA differs from fepA, however, in that the gene divergently transcribed from it (irgB) is itself a trans-acting positive regulator of irgA transcription, whereas the gene that is divergently transcribed from fepA (fes) does not appear to be involved in *fepA* regulation (24).

A trans-acting factor, angR, has recently been described in the plasmid-mediated iron-regulated anguibactin iron-uptake system of V. anguillarum (25). angR is a positive regulator of this system and is itself negatively regulated by iron at the transcriptional level (26). angR is not divergently transcribed from any of the genes it is known to regulate (26). It is not yet known whether or not AngR has any homology to IrgB or other members of the LysR family.

We have developed the following model for regulation of *irgA* expression in V. *cholerae*. In the presence of sufficient iron, transcription of both *irgA* and *irgB* is negatively regulated by a Fur-like protein. In low-iron conditions, negative regulation of transcription is removed, and production of IrgB leads to positive transcriptional activation of *irgA*. We do not yet know whether transcription of *irgB* might also be

positively autoregulated. It seems likely that the explanation of the very high induction ratio of irgA expression under lowand high-iron conditions relates to the fact that its cognate positive transcriptional activator (irgB) is itself negatively regulated by iron. We are currently investigating whether other iron-regulated proteins in V. cholerae are also positively activated by IrgB. In addition, we are examining other bacterial species to look for further examples of dual positive and negative iron regulatory systems.

We gratefully acknowledge John J. Mekalanos for helpful discussions, Gregory D. N. Pearson for providing the pGP704 vector, and Brian Seed for providing synthetic oligonucleotides. This work was supported by a 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 the recipient of a National Research Service Award from the National Institute of Allergy and Infectious Diseases.

- Taylor, R. K., Miller, V. L., Furlong, D. B. & Mekalanos, J. J. (1987) Proc. Natl. Acad. Sci. USA 84, 2833–2837.
- Peterson, K. M. & Mekalanos, J. J. (1988) Infect. Immun. 56, 2822-2829.
- Pappenheimer, A. M., Jr. (1955) Symp. Soc. Gen. Microbiol. 5, 40-56.
- 4. Dubos, R. J. & Geiger J. W. (1946) J. Exp. Med. 84, 143-156.
- Calderwood, S. B. & Mekalanos, J. J. (1987) J. Bacteriol. 169, 4759–4764.
- Bjorn, M. J., Iglewski, B. H., Ives, S. K., Sadoff, J. C. & Vasil, M. L. (1978) Infect. Immun. 19, 785-791.
- Bjorn, M. J., Sokol, P. A. & Iglewski, B. H. (1979) J. Bacteriol. 138, 193-200.
- Actis, L. A., Potter, S. A. & Crosa, J. H. (1985) J. Bacteriol. 161, 736–742.
- 9. Carniel, E., Mazigh, D. & Mollaret, H. H. (1987) Infect. Immun. 55, 277-280.
- Goldberg, M. B., DiRita, V. J. & Calderwood, S. B. (1990) Infect. Immun. 58, 55-60.
- Goldberg, M. B., Boyko, S. A. & Calderwood, S. B. (1990) J. Bacteriol. 172, 6863–6870.
- Manoil, C. & Beckwith, J. (1985) Proc. Natl. Acad. Sci. USA 82, 8129–8133.
- 13. Miller, V. I. & Mekalanos, J. J. (1988) J. Bacteriol. 170, 2575–2583.
- 14. Simon, R., Priefer, U. & Puhler, A. (1983) *Bio/Technology* 1, 784–791.
- Kolter, R., Inuzuka, M. & Helsinki, D. R. (1978) Cell 15, 1199-1208.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
- 17. Kyte, J. & Doolittle, R. F. (1982) J. Mol. Biol. 157, 109-132.
- 18. Chou, P. & Fasman, G. D. (1974) Biochemistry 13, 222-245.
- Lipman, D. J. & Pearson, W. R. (1985) Science 227, 1435– 1441.
- 20. Shine, J. & Dalgarno, L. (1975) Nature (London) 254, 34-38.
- 21. Hawley, D. K. & McClure, W. R. (1983) Nucleic Acids Res. 11, 2237-2255.
- Calderwood, S. B. & Mekalanos, J. J. (1988) J. Bacteriol. 170, 1015–1017.
- Henikoff, S., Haughn, G. W., Calvo, J. M. & Wallace, J. C. (1988) Proc. Natl. Acad. Sci. USA 85, 6602–6606.
- Pettis, G. S., Brickman, T. J. & McIntosh, M. A. (1988) J. Biol. Chem. 263, 18857-18863.
- Salinas, P. C., Tomalsky, M. E. & Crosa, J. H. (1989) Proc. Natl. Acad. Sci. USA 86, 3529–3533.
- Farrell, D. H., Mikesell, P., Actis, L. A. & Crosa, J. H. (1990) Gene 86, 45-51.
- Mekalanos, J. J., Swartz, D. J., Pearson, G. D. N., Harford, N., Groyne, F. & de Wilde, M. (1983) Nature (London) 306, 551-557.