JOURNAL OF BACTERIOLOGY, May 1991, p. 2842–2851
0021-9193/91/092842-10$02.00/0

Vol. 173, No. 9

Expression of the *Vibrio cholerae* Gene Encoding Aldehyde Dehydrogenase Is under Control of ToxR, the Cholera Toxin Transcriptional Activator

CLAUDE PARSOT† AND JOHN J. MEKALANOS*

Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115

Received 10 September 1990/Accepted 16 February 1991

The *toxR* gene of *Vibrio cholerae* encodes a transcriptional activator required for the expression of the cholera toxin genes (*ctxAB*) and more than 15 other genes encoding secreted or membrane proteins. The latter group includes virulence genes involved in the biogenesis of the TCP pilus, the accessory colonization factor, and such ToxR-activated genes as *tagA*, mutations in which cause no detectable virulence defect in the suckling mouse model. To analyze the regulation of expression and the structure of *tagA*, we have cloned and sequenced about 2 kb of DNA upstream from a *tagA*::Tn*phoA* fusion. While the portion of the *tagA* gene product examined presented no extensive similarity to any known protein, the amino acid sequence deduced from an open reading frame (designated *aldA*) located upstream from and in opposite orientation to *tagA* was highly similar to the sequences of eukaryotic aldehyde dehydrogenases. An assay of aldehyde dehydrogenase activity in extracts of a wild-type *V. cholerae* strain and an *aldA* mutant confirmed that *aldA* encodes an aldehyde dehydrogenase. Expression of the *aldA* gene was studied together with that of *tagA* in both *V. cholerae* and *Escherichia coli*. The expression of both *tagA* and *aldA* was environmentally regulated and dependent on a functional *toxR* gene in *V. cholerae*, but neither promoter was activated by ToxR in *E. coli*, suggesting that expression of *tagA* and *aldA* requires an additional transcriptional activator besides ToxR. The *aldA* gene is the first example of a gene encoding a cytoplasmic protein that is under the control of ToxR, and this suggests that metabolic enzymes may constitute novel members of virulence regulons in bacteria.

Vibrio cholerae is the causative agent of the diarrheal disease cholera, which is induced largely by the action of cholera toxin, a heat-labile enterotoxin encoded by the *ctxA* and *ctxB* genes (14). Expression of the *ctxAB* operon is under the control of the transcriptional activator ToxR. The *toxR* gene has been cloned through its ability to activate the expression of a *ctx-lacZ* transcriptional fusion in *Escherichia coli* (17, 19); *toxR* encodes a transmembrane, DNA-binding protein that recognizes a heptanucleotide sequence tandemly repeated in the *ctx* regulatory region (20).

The ToxR regulon, defined as the set of genes whose expression is under the control of ToxR, comprises, in addition to *ctxAB*, the *tcp* genes required for the production of the TCP pilus, the genes encoding the outer membrane proteins OmpU and OmpT, the *acf* genes specifying an accessory colonization factor (ACF), and a group of genes, the *tag* genes (ToxR-activated genes), which have not yet been associated with a definite virulence property (23). These *tag* genes were identified through Tn*phoA* mutagenesis of the *V. cholerae* chromosome and subsequent screening for the active alkaline phosphatase (PhoA) hybrid proteins whose expression was modulated by the environmental growth conditions known to regulate toxin production (23). Because expression of *tag* genes is dependent on an intact *toxR* locus, it is a formal possibility that ToxR might directly bind to and activate *tag* promoters.

To understand the structure and the regulation of expression of these *tag* genes, we cloned the *tagA*::*phoA* fusion from *V. cholerae* KP8.56 (23) and investigated its expression in both *V. cholerae* and *E. coli*. We present evidence that the ToxR regulon includes *aldA*, a gene encoding aldehyde dehydrogenase (Ald dehydrogenase), that was found upstream from and in opposite orientation to *tagA*. Although the expression from both the *tagA* and *aldA* promoters was under the control of ToxR in *V. cholerae*, these promoters were not activated by ToxR in *E. coli*.

MATERIALS AND METHODS

Bacterial strains and growth media. Derivatives of *V. cholerae* Ogawa 395 (O395) and *E. coli* K-12 were maintained at −70°C in Luria broth (LB) medium (16) containing 25% glycerol (vol/vol) or on LB plates. Methods for cultivation of *V. cholerae* strains under high and low pHs were as described by Miller and Mekalanos (19). Antibiotics were used at the following concentrations: ampicillin (Amp), 50 mg/ml; gentamicin (Gen), 30 mg/ml; kanamycin (Kan), 30 mg/ml; streptomycin (Sm), 100 mg/ml; and tetracycline (Tet), 15 mg/ml. *E. coli* GT869 [*thrB1004 pro thi rpsL hsdS lacZΔM15* F′(*lacZΔM15 lacI*^{q} *traD36 proA*$^+$ *proB*$^+$)] (21) was used for cloning experiments, *E. coli* SM10 (*thi thr leu tonA lacY supE recA*::RP4-2Tc::Mu Km) (28) was used to transfer plasmids to *V. cholerae* by conjugation, and *E. coli* KS272 (F$^-$ Δ*lacX74 galE galK thi rpsL* Δ*phoA*) (29) was used to assay PhoA and β-galactosidase activities expressed from the recombinant plasmids. *V. cholerae* O395 Sm is a streptomycin-resistant derivative of wild-type O395 (30), JJM43 is a *toxR* derivative of O395 Sm (30), and KP8.56 is a derivative of O395 Sm that carries a Tn*phoA* insertion in *tagA* (23).

Molecular cloning procedures. Plasmid DNA purification,

* Corresponding author.
† Present address: Unité de Pathogénie Microbienne Moléculaire, Département de Bactériologie et de Mycologie, Institut Pasteur, 75724 Paris Cedex 15, France.

DNA restriction and separation by gel electrophoresis, transfer and hybridization, ligation, and transformation of *E. coli* strains were performed by the standard methods described by Maniatis et al. (12). Nucleotide sequences were determined by the dideoxy-chain termination procedure (26) on double-stranded plasmid DNA.

SDS-PAGE and immunoblotting. Polyacrylamide gel electrophoresis (PAGE) in 8% polyacrylamide gels in the presence of sodium dodecyl sulfate (SDS) was performed as described by Laemmli (10). After electrophoresis, proteins were either stained with Coomassie brilliant blue or transferred by electrophoresis to nitrocellulose sheets (31). The preparation and use of the anti-PhoA antiserum has been previously described (23).

β-Galactosidase and PhoA assays. β-Galactosidase activity was assayed as described by Miller (16), by using the substrate *o*-nitrophenol-β-D-galactoside; PhoA activity was assayed as described by Peterson and Mekalanos (23), by using the substrate *p*-nitrophenyl phosphate. Both activities are expressed in Miller units (16).

Ald dehydrogenase assay. *V. cholerae* cells were grown to late exponential phase in LB medium (pH 7.0) at 30°C, collected by centrifugation, resuspended in 1/25 volume of H_2O, and sonicated. The coenzyme A (CoA)-independent Ald dehydrogenase activity was assayed in a buffer containing 100 mM Tris-HCl (pH 8.0), 100 mM KCl, 1 mM mercaptoethanol, 0.5 mM NAD, and 0.17 mM acetaldehyde as described previously (1). The CoA-dependent activity was assayed in the presence of 23 mM oxidized CoA. One unit of activity is defined as the amount of enzyme producing a change of 0.001 optical density at 340 nm per min and mg of protein.

Construction of an *aldA-lacZ* transcriptional fusion. The *lacZ* reporter gene (lacking its own promoter) from the pMLB1010 vector (27) was isolated as a *Bam*HI-*Dra*I fragment and cloned into the *Eco*RV site located within the *aldA* coding sequence on pVC7.1 (see Fig. 4, nucleotide 1,238). The *Bam*HI fragment of the pVC112 plasmid thus constructed, carrying both the *tagA*::*phoA* translational and *aldA-lacZ* transcriptional fusions, was then cloned into the *Bam*HI site of pLAFR2, giving rise to pVC115 (Fig. 1).

Construction of an *aldA* mutant of *V. cholerae* O395 Sm. The *Stu*I-*Bam*HI fragment of pVC7.1 (see Fig. 4, nucleotides 59 to 2,068) was cloned between the filled-in *Sal*I site and the *Bam*HI site of pUC18 (34) to give rise to pVC97, in which the *V. cholerae* DNA insert is flanked by *Sal*I and *Bam*HI restriction sites. The 1.4-kb *Hin*cII fragment from pUC71K (kindly given by J. Vieira and J. Messing), carrying the *aphA-1* gene (which encodes 3′-aminoside phosphotransferase type I, conferring resistance to Kan), was inserted into the filled-in *Cla*I site of pVC97 (located within the *aldA* coding sequence at position 1,115 [see Fig. 4]) to give rise to pVC97.1. The *Sal*I-*Bam*HI fragment of pVC97.1 was then cloned into the filled-in *Hin*dIII site of pLAFR2 (Tet^r) (4) to give pVC97.3 (Fig. 1). This last plasmid was mobilized into *V. cholerae* O395 Sm, selecting the transconjugants on LB plates containing Sm, Tet, and Kan. Marker exchange of the wild-type chromosomal *aldA* gene for the insertion-inactivated copy (the *aldA1* allele) carried by pVC97.3 was selected by incompatibility with plasmid pPH1JI (Gen^r) as described previously (15). After curing of the pPH1JI plasmid by passage at 42°C, Southern blot analysis confirmed the structure of the chromosome in mutant strain RC1134 carrying the *aldA1* allele.

In vivo assay for *V. cholerae* intraintestinal growth. The competitive index for the *V. cholerae aldA* mutant was determined in 3-day-old suckling CD1 mice inoculated orally as previously described (3, 30). Viable bacteria were recovered by plating dilutions of intestinal homogenates on LB agar containing Sm, and the ratio of *aldA* mutant (Sm^r Kan^r) to wild-type O395 Sm (Sm^r) was calculated after screening for the Kan^r clones.

Nucleotide sequence accession number. The nucleotide sequence reported in this article has been submitted to GenBank and assigned the accession number M60658.

RESULTS

Cloning of the *tagA*::*phoA* fusion. The first step in cloning the *tagA*::*phoA* fusion from *V. cholerae* KP8.56 (23) was to determine the physical map of the Tn*phoA* insertion site in this strain, by using a DNA fragment of Tn*phoA* as a probe. From this analysis (data not shown), we deduced the positions of restriction sites for *Stu*I, *Xba*I, *Eco*RV, and *Bam*HI, located 5.1, 5.6, 6.4, and 7.1 kb, respectively, upstream from the *Bam*HI site of Tn*phoA* (Fig. 1). We then cloned the 7.1-kb *Bam*HI fragment containing 2.1 kb of DNA upstream from the junction with *phoA* and 5 kb of Tn*phoA* (including the *aphA-2* gene conferring resistance to Kan). KP8.56 chromosomal DNA was digested by *Bam*HI and ligated with *Bam*HI-linearized pUC18 DNA (34), and the ligation mixture was used to transform *E. coli* GT869 and select the transformants on LB plates containing Amp and Kan. This allowed us to isolate plasmid pVC7.1, which carries a 7.1-kb *Bam*HI insertion whose restriction map (Fig. 1) is consistent with Southern blot analysis of KP8.56.

Expression of the cloned *tagA*::*phoA* fusion in *V. cholerae*. To study expression of the cloned *tagA*::*phoA* fusion, the 7.1-kb *Bam*HI insertion of pVC7.1 was cloned into the *Bam*HI site of the low-copy-number vector pLAFR2 (Tet^r) (4). Having thus obtained the recombinant plasmids pVC8 and pVC9, (Fig. 1), representing the two possible orientations of the *Bam*HI insertion, the plasmids were then mobilized by conjugation into *V. cholerae* derivatives O395 Sm (wild type) and JJM43 (*toxR*). PhoA activity expressed from these plasmids was assayed in the transconjugants inoculated in LB at pH 6.5 or 8.4, growth conditions known to modulate the expression of the chromosomally located *tagA*::*phoA* fusion (23). As shown in Table 1, expression of the cloned *tagA*::*phoA* fusion was regulated by the starting pH of the growth medium and was dependent on an intact *toxR* locus. These results indicated that the *tagA* regulatory region was present and functional on the cloned fragment.

To localize the *tagA* regulatory region, we constructed *Xba*I deletion derivatives of pVC8 and pVC9. Plasmid DNA was partially digested by *Xba*I and religated, giving rise to plasmids pVC51, pVC52, pVC55, and pVC58 (Fig. 2). These plasmids were then mobilized into *V. cholerae* O395 Sm, and the PhoA activity was assayed after growth of the transconjugants under inducing (pH 6.5) or repressing (pH 8.4) conditions. In all cases, PhoA activity was modulated by the starting pH of the growth media, indicating that the *tagA* promoter and regulatory region was located in the 600-bp region extending from the first *Xba*I site to the *phoA* junction (Fig. 2). We then took advantage of the presence of several *Nsi*I sites in the *tagA* upstream region to further localize the regulatory region. The DNA of pVC7.1 was digested by *Nsi*I and religated to construct pVC72, in which the *phoA* upstream region was reduced to 300 bp. The *Bam*HI insert of pVC72 was then cloned into the *Bam*HI site of pLAFR2 to give rise to pVC80 and pVC81 (Fig. 2), representing the two possible orientations of the *Bam*HI fragment. As shown in

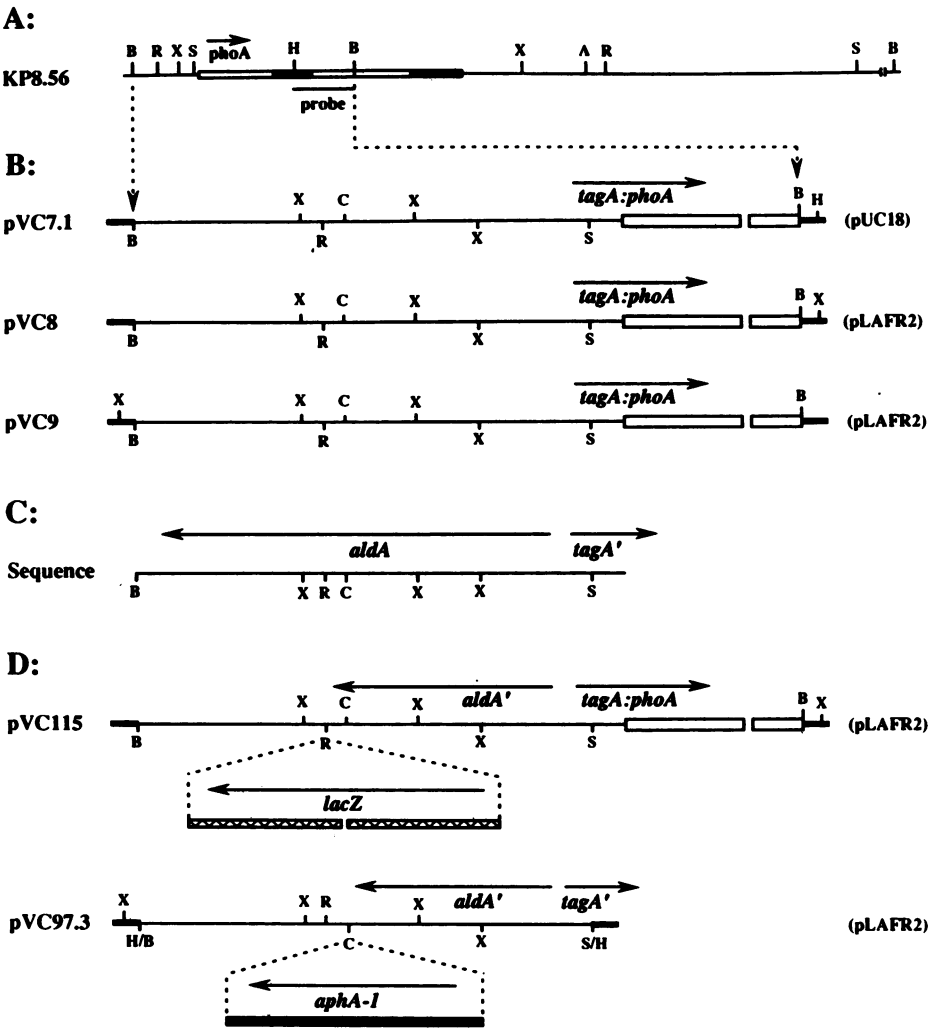

FIG. 1. Restriction maps of the Tn*phoA* insertion site in KP8.56 and of the plasmids carrying the *tagA*::*phoA* fusion. (A) Restriction map of the chromosome around the insertion site of Tn*phoA* in KP8.56 (deduced from Southern blot analysis). The chromosome (thin line) and Tn*phoA* (open bar) are shown. The position of the DNA fragment used as a probe is indicated below the part of Tn*phoA* on which the IS*50* sequences are indicated by dashed arrows. This restriction map was determined after digestion of KP8.56 chromosomal DNA by *Bam*HI, which cuts once in Tn*phoA*, or by *Bam*HI plus either *Eco*RV, *Sca*I, *Stu*I, or *Xba*I, restriction enzymes that do not recognize the transposon sequence. Accordingly, for each of these enzymes, only the position of the restriction site proximal to Tn*phoA* could be determined. (B) Recombinant plasmid pVC7.1, including the 7.1-kb *Bam*HI fragment carrying the *tagA*::*phoA* fusion and part of Tn*phoA* cloned into pUC18. Also shown are pVC8 and pVC9, resulting from the cloning of the same *Bam*HI fragment into the pLAFR2 vector. (C) Positions of *aldA* and the 5′ part of *tagA*, as deduced from the sequence shown in Fig. 4. (D) Schematic structure of the pVC115 plasmid, which carries the *aldA-lacZ* transcriptional fusion, along with that of the pVC97.3 plasmid used to inactivate the chromosomal *aldA* gene (see text for details). Restriction site abbreviations: A, *Sca*I; B, *Bam*HI; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; R, *Eco*RV; S, *Stu*I; X, *Xba*I.

Fig. 2, both plasmids still expressed PhoA activity, but regulation of the *tagA*::*phoA* fusion was almost abolished, showing only a 2.5 U increase under inducing conditions. These results indicated that the beginning of the *tagA* gene was still present on the cloned fragment and that a regulatory element was probably located between the *Xba*I and *Nsi*I sites (Fig. 2).

Characterization of the TagA::PhoA hybrid protein. The size of the protein encoded by the *tagA*::*phoA* fusion was analyzed by immunoblot with antibodies directed against PhoA. Crude extracts of *V. cholerae* strains expressing the *tagA*::*phoA* fusion from either the chromosome or the pVC8 plasmid were prepared from cells grown under inducing conditions. In both cases, two species of 58 and 56 kDa were recognized by the anti-PhoA antiserum, with a stronger signal for the strain carrying the *tagA*::*phoA* fusion on the plasmid (Fig. 3). Knowing that the PhoA part of the hybrid molecule represents a polypeptide of 48 kDa (13), it can deduced from the size of the larger species (58 kDa) that the TagA moiety of the hybrid protein consists of about 90 residues. The 56-kDa polypeptide might represent the mature form of the TagA::PhoA hybrid protein after processing of a signal sequence of about 20 amino acid residues or a degradation product, as previously observed for other PhoA fusions in *V. cholerae* (23).

Nucleotide sequence of the *tagA*::*phoA* fusion. The se-

TABLE 1. PhoA and β-galactosidase activities expressed from the cloned *tagA::phoA* and *aldA-lacZ* fusions in *V. cholerae* *toxR*$^+$ and *toxR* mutant strains

Plasmid (fusion)[a]	Background[b]	Activity (U)[c]			
		PhoA		β-Gal	
		pH 6.5	pH 8.4	pH 6.5	pH 8.4
pVC8 (*tagA::phoA*)	*toxR*$^+$	4,400	80	ND	ND
pVC8 (*tagA::phoA*)	*toxR*	60	70	ND	ND
pVC9 (*tagA::phoA*)	*toxR*$^+$	3,200	100	ND	ND
pVC9 (*tagA::phoA*)	*toxR*	60	60	ND	ND
pVC115 (*tagA::phoA* + *aldA-lacZ*)	*toxR*$^+$	5,020	60	5,210	370
pVC115 (*tagA::phoA* + *aldA-lacZ*)	*toxR*	40	40	300	340

[a] pVC8, pVC9, and pVC115 are shown in Fig. 1.
[b] O395 Sm and JJM43 were used as the *V. cholerae* *toxR*$^+$ and *toxR* mutant strains, respectively.
[c] PhoA and β-galactosidase (β-Gal) activities were assayed as described previously (16, 23). Shown are the starting pHs of the growth media. ND, not determined.

quence of the 2.1-kb DNA fragment located upstream from the junction with *phoA* was determined by using deletion derivatives of pVC7.1 and synthetic oligonucleotides. Both strands of DNA were completely sequenced. The nucleotide sequence of this fragment as well as the deduced sequences of the TagA and AldA polypeptides is shown in Fig. 4. This sequence is presented from 5′ to 3′ in the direction of transcription of *aldA*, i.e., opposite to the orientation of *tagA*.

The *tagA* reading frame was identified as the one in frame with the *phoA* coding sequence. Within the *tagA* open reading frame (ORF), extending from nucleotide 238 to nucleotide 1 at the junction with Tn*phoA*, there are three potential translation start sites: two ATG codons at nucleotides 238 and 181 (shown as CAT codons in Fig. 4) and a GTG codon at nucleotide 205 (shown as a CAC codon in Fig. 4). We have tentatively assigned the beginning of the *tagA* gene to the GTG codon which is preceeded by a potential ribosome binding site (5′-GAGGAG-3′) and presented the translation of the *tagA* ORF accordingly (Fig. 4). This is

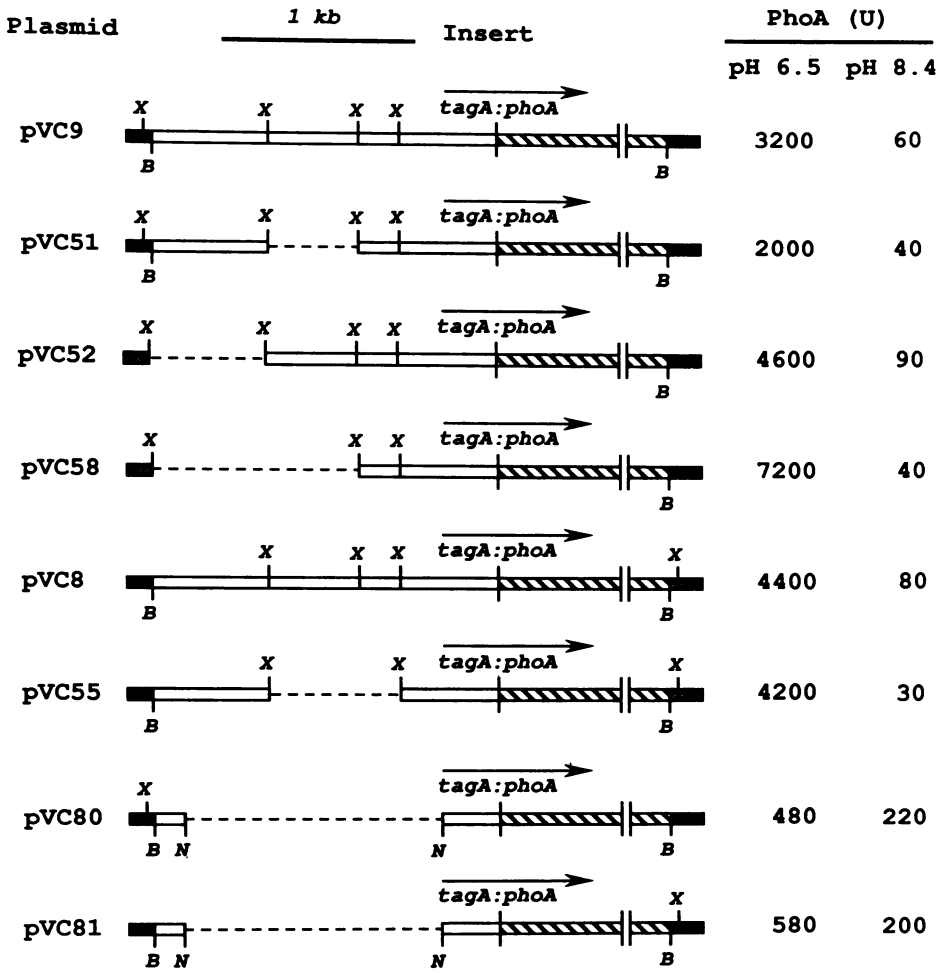

FIG. 2. Structures of the deletion derivative plasmids constructed to localize the *tagA* regulatory region and PhoA activity expressed from these plasmids in *V. cholerae*. The restriction map of the insert carried by each plasmid is indicated in front of the plasmid name. *V. cholerae* DNA (□), Tn*phoA* (▨), and the two sides of the pLAFR2 polylinker (▩ and ■) are shown. pVC51, pVC52, and pVC58 are derivatives of pVC9, and pVC55 is a derivative of pVC8. The *Bam*HI insert carried by pVC80 and pVC81 was constructed by deleting the *Nsi*I fragments in the *tagA* upstream region of pVC7.1 (see text for details). DNA fragments that have been deleted are also shown (– – –). Restriction site abbreviations: B, *Bam*HI; N, *Nsi*I; X, *Xba*I. PhoA activities expressed from each plasmid in transconjugants of *V. cholerae* O395 Sm grown to late exponential phase at 30°C in LB at pH 6.5 or 8.4 are shown on the right.

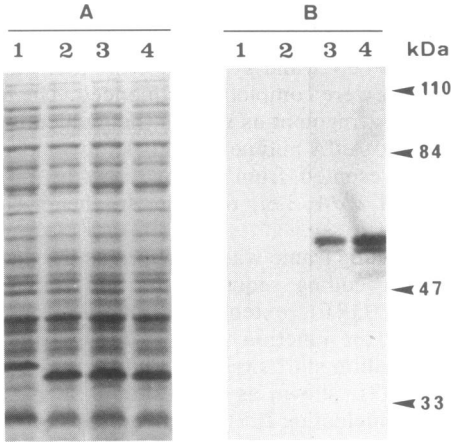

FIG. 3. Immunoblot analysis of the TagA::PhoA fusion protein. Crude extracts of *V. cholerae* cells grown in LB at pH 6.5 at 30°C were analyzed by SDS-PAGE and either stained with Coomassie brilliant blue (A) or transferred to nitrocellulose and probed with anti-PhoA antibodies (B). Lanes: 1, JJM43 (*toxR*); 2, O395 Sm (wild type); 3, KP8.56 (carrying the *tagA*::*phoA* fusion on the chromosome); 4, O395 Sm with the *tagA*::*phoA* fusion cloned into pVC8 (Fig. 1). Arrowheads indicate the positions and apparent molecular masses (kilodaltons) of prestained molecular mass markers.

consistent with the immunoblot analysis (Fig. 3) that suggested the TagA part of the hybrid protein to be ca. 90 amino acid residues long and with the deletion analysis (Fig. 2) that localized the beginning of the *tagA* gene downstream from the *Nsi*I site (nucleotide 284 in Fig. 4).

The amino acid sequence deduced from the *tagA* ORF shows a stretch of hydrophobic and nonpolar residues that is likely to be the signal sequence responsible for the periplasmic localization and PhoA activity of the fusion protein (13). The processing of this signal sequence would give rise to a mature protein that might correspond to the 56-kDa polypeptide detected with the anti-PhoA antibodies (Fig. 3). No significant similarity was detected between the sequence of the N-terminal part of TagA and any protein sequence contained in the National Biomedical Research Foundation library (release 21) by using the FASTP computer program of Lipman and Pearson (11).

Identification of the gene for Ald dehydrogenase. Upstream from and in opposite orientation to *tagA*, we detected an ORF extending from nucleotide 305 to 1,936 (Fig. 4). Assuming that the translation start site is the first ATG codon of the ORF (nucleotide 419 in Fig. 4), the encoded protein would consist of a 505-amino-acid-residue polypeptide. Downstream from the ORF stop codon, there is a region of dyad symmetry similar to a transcription termination signal, suggesting that the ORF belongs to a monocistronic transcription unit.

The amino acid sequence deduced from that ORF was compared with the protein sequences contained in the National Biomedical Research Foundation library by using the FASTP program (11); this comparison revealed an extensive similarity between the ORF sequence and the following sequence(s): (i) the sequences of the Ald dehydrogenases (EC 1.2.1.3) from humans (2, 6), horses (32), and *Aspergillus nidulans* (24), and (ii) the sequence of the *d*-1-pyrroline-5-carboxylate (P5C) dehydrogenase (EC 1.5.1.12) from *Saccharomyces cerevisiae* (9). The alignment of the ORF sequence with those of Ald dehydrogenase from *A. nidulans* and of P5C dehydrogenase from *S. cerevisiae* is shown in Fig. 5. Except for an N-terminal extension of about 50 residues which is present only in the P5C dehydrogenase sequence, the three proteins are very similar in size and the sequences can be aligned over their entire lengths. Pairwise comparisons indicate that 220 positions (46%) are occupied by identical residues in the *V. cholerae* ORF and *A. nidulans* Ald dehydrogenase sequences; this is twice as much as the number of identities detected by comparing either the *V. cholerae* ORF or the *A. nidulans* Ald dehydrogenase to the *S. cerevisiae* P5C dehydrogenase (111 and 135 identical positions, respectively). Such an extensive similarity between the *V. cholerae* ORF and Ald dehydrogenase sequences strongly suggests that this ORF corresponds to the *V. cholerae* gene encoding Ald dehydrogenase; this ORF was therefore named *aldA*, and the encoded protein referred to as Ald dehydrogenase.

Assay of Ald dehydrogenase activity in *V. cholerae*. To confirm that the *aldA* gene was encoding an Ald dehydrogenase, we constructed an *aldA* mutant derivative (RC1134) of *V. cholerae* O395 Sm (see Materials and Methods) and assayed Ald dehydrogenase activity in extracts of both strains. Whereas 11.2 U of activity was measured in the extract of O395 Sm, no activity (i.e., less than 0.2 U) was detected in the extract of RC1134. These results provide definitive evidence that *aldA* specifies an Ald dehydrogenase.

Since two types of Ald dehydrogenases, the CoA-dependent (EC 1.2.1.10) and CoA-independent (EC 1.2.1.3) enzymes, have been detected in some *E. coli* strains (7, 8), we tested the effect of CoA on the activity of the *V. cholerae* Ald dehydrogenase. Addition of 23 mM CoA in the assay did not enhance Ald dehydrogenase activity in the extracts of either O395 Sm or RC1134, suggesting that (i) the *aldA* gene product is a CoA-independent Ald dehydrogenase and (ii) no CoA-dependent Ald dehydrogenase is expressed in *V. cholerae*, at least under these growth conditions.

ToxR regulates *aldA* expression in *V. cholerae*. The close proximity of the *aldA* and *tagA* genes raised the question whether the expression of *aldA* was also under the control of ToxR. We therefore constructed an *aldA-lacZ* transcriptional fusion on the plasmid already carrying the *tagA*::*phoA* fusion so that the expression from both the *tagA* and *aldA* promoters could be investigated simultaneously (see Materials and Methods). The pVC115 plasmid (Fig. 1), carrying both the *tagA*::*phoA* and *aldA-lacZ* fusions, was mobilized into *V. cholerae* derivatives O395 Sm (wild type) and JJM43 (*toxR*), and the β-galactosidase and PhoA activities were assayed in transconjugants grown under either inducing (pH 6.5) or repressing (pH 8.4) conditions. As shown in Table 1, the expression of both the *tagA*::*phoA* and the *aldA-lacZ* fusions was (i) modulated by the starting pH of the culture and (ii) under the control of ToxR, indicating a similar control on the *tagA* and *aldA* promoters.

Inactivation of *aldA* does not affect the colonization properties of *V. cholerae*. To investigate the importance of the ToxR-regulated *aldA* gene in the virulence properties of *V. cholerae*, we used the *aldA* mutant RC1134 (see Materials and Methods). This strain was tested for a colonization defect in the infant mouse intestinal competition assay (3, 30). The *aldA* mutant was not displaced by the wild-type strain O395 Sm (data not shown) and therefore appeared to be unaffected in its colonization properties, at least in this animal model.

Expression of *tagA* and *aldA* is not activated by ToxR in *E. coli*. As the expression of both the *tagA*::*phoA* and *aldA*-

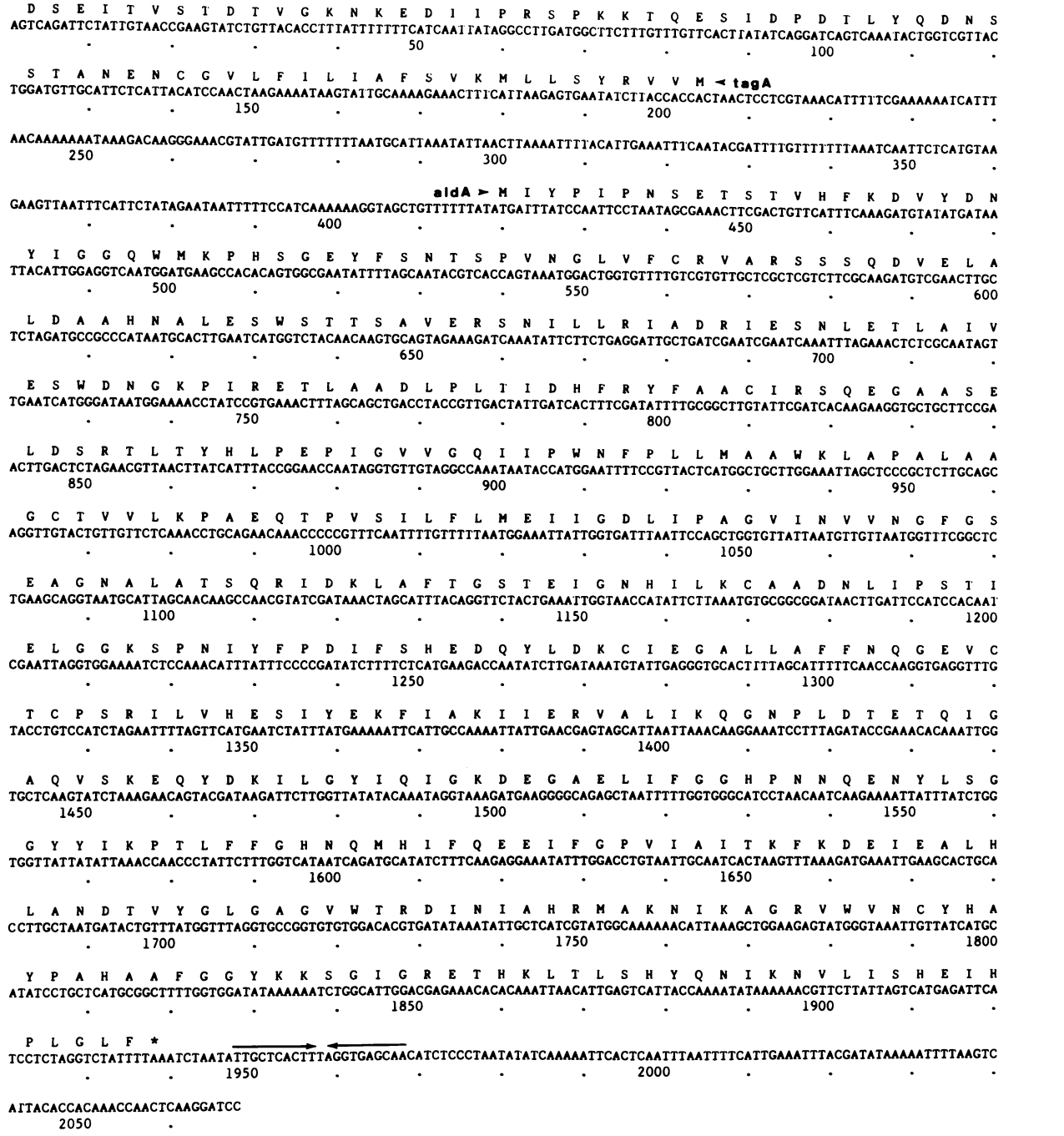

FIG. 4. Nucleotide sequence of the *tagA*::*phoA* upstream region. The nucleotide sequence is shown from 5′ to 3′ in the direction of transcription of *aldA*, along with the deduced amino acid sequences of AldA and the N-terminal part of TagA. As *tagA* is oriented opposite to *aldA*, the TagA sequence should be read from right to left and from bottom to top. Nucleotide 1 is the first nucleotide of the *V. cholerae* DNA at the junction with Tn*phoA*. A region of dyad symmetry located downstream from *aldA* is indicated by arrows, and its stop codon is indicated by an asterisk.

lacZ fusions was under the control of ToxR in *V. cholerae*, we tested the effect of ToxR on the expression of these fusions in *E. coli*. In these experiments, the ToxR protein is expressed from the plasmid pVM25 (Cm^r) (17), a pACYC184 derivative that carries *toxR* and is compatible with pLAFR2 derivatives. As a control for the ability of ToxR to activate the transcription of a promoter cloned into pLAFR2, we constructed plasmid pVC42, a pLAFR2 derivative carrying a *ctxA*::*phoA* fusion. Plasmid pVC42 contains a *Bam*HI insertion corresponding to the *ctx* regulatory region (starting at the *Hha*I site, i.e., 189 bp upstream from the *ctx* transcription start point) and the first 60 codons of *ctxA* fused in frame to *phoA*, as well as a 5 kb fragment of Tn*phoA*. This construction is thus very similar to the *tagA*::*phoA* fusion

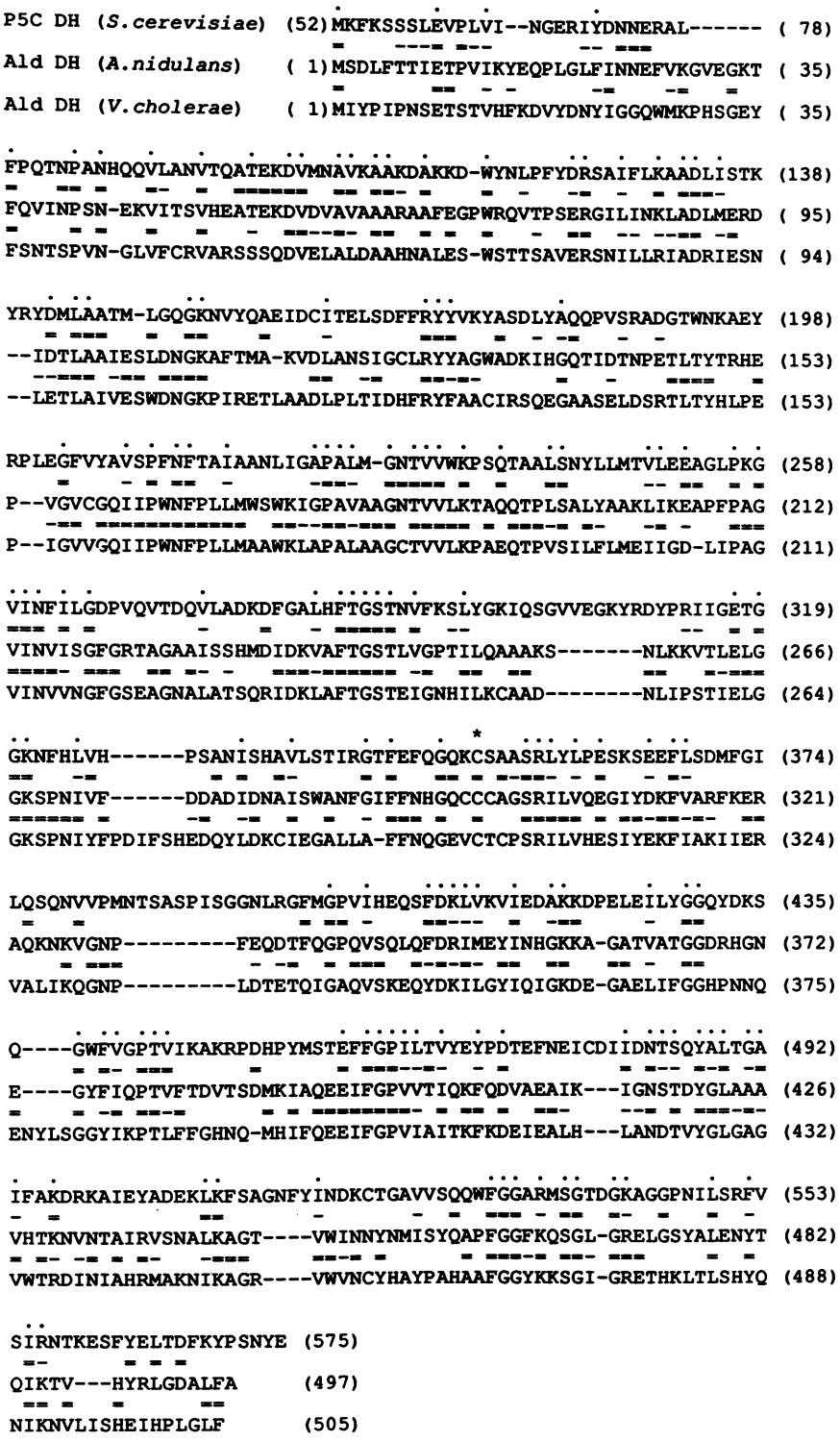

FIG. 5. Amino acid sequence comparison of the *V. cholerae* Ald dehydrogenase with the *A. nidulans* Ald dehydrogenase and the *S. cerevisiae* P5C dehydrogenase. The amino acid sequence deduced from the *V. cholerae* ORF (Fig. 4) is aligned with the sequence of the Ald dehydrogenase from *A. nidulans* (24) and with part of the sequence (the first 51 residues are not shown) of the P5C dehydrogenase from *S. cerevisiae* (9). Identical (=) and similar (−) residues (I-L-V-M-, D-E, R-K, S-T, G-A, and F-Y) between pairs of sequences, the Cys residue proposed to be part of the active site (5, 6, 24) (*), and positions which are occupied by identical or similar residues in the three sequences (·) are shown. Numbers indicate the positions of the residues in the original sequences.

TABLE 2. PhoA and β-galactosidase activities expressed from the cloned *ctxA::phoA*, *tagA::phoA*, and *aldA-lacZ* fusions in *E. coli*

Reporter plasmid[a]	Activator plasmid[b]	Activity (U)[c]	
		PhoA	β-Gal
pVC42 (*ctxA::phoA*)	pACYC184	260	ND
pVC42 (*ctxA::phoA*)	pVM25 (*toxR*$^+$)	3,330	ND
pVC115 (*tagA::phoA* + *aldA-lacZ*)	pACYC184	6	280
pVC115 (*tagA::phoA* + *aldA-lacZ*)	pVM25 (*toxR*$^+$)	7	230

[a] pVC115 is shown in Fig. 1, and pVC42 is described in the text.
[b] pVM25 expresses *toxR* from the *tet* promoter of the pACYC184 vector.
[c] PhoA and β-galactosidase (β-Gal) activities were assayed as described previously (16, 23). ND, not detected.

present on pVC8 and pVC115 and is in the same orientation with respect to the pLAFR2 moiety.

E. coli KS272 *phoA lacZ* was first transformed by pVC115 (*tagA::phoA* and *aldA-lacZ*) or by pVC42 (*ctxA::phoA*), and the transformants were then used as the recipients for another transformation by pVM25 (*toxR*$^+$) or pACYC184 (the vector part of pVM25, i.e., the *toxR* mutant). PhoA and β-galactosidase activities were assayed after growth of the double transformants in LB, pH 6.5, at 30°C. As shown in Table 2, the expression of the *ctxA::phoA* fusion was markedly increased in the presence of pVM25, indicating that ToxR was able to activate transcription of the *ctx* promoter cloned into pLAFR2. On the other hand, neither the expression of the *tagA::phoA* fusion nor that of the *aldA-lacZ* fusion was activated by ToxR in *E. coli*.

To test whether the lack of activation of the *tagA* and *aldA* promoters was due to a titrating effect of the *tagA-aldA* regulatory region cloned into pLAFR2, we took advantage of the *ctxA-lacZ* fusion present in the chromosome of *E. coli* VM2 (18); this strain was transformed by pVC8 (*tagA::phoA*) and then by either pACYC184 or pVM25 (*toxR*$^+$). PhoA and β-galactosidase activities were assayed in the double transformants grown at 30°C in LB, pH 6.5; β-galactosidase activity reflects the expression of the chromosomal *ctxA-lacZ* fusion, and PhoA activity reflects the expression of the *tagA::phoA* fusion carried by the pVC8 plasmid. As shown in Table 3, pVM25 led to a 10-fold increase in expression of the *ctxA-lacZ* fusion. Activation of the expression of this chromosomal *ctxA-lacZ* fusion was not impaired by the presence of the plasmid carrying the *tagA-aldA* regulatory region, indicating that lack of activation of the *tagA* and *aldA* promoters by ToxR was not the result of ToxR titration.

TABLE 3. PhoA activity expressed from the cloned *tagA::phoA* fusion and β-galactosidase activity expressed from the chromosomal *ctxA::lacZ* fusion in *E. coli*

Reporter plasmid[a]	Activator plasmid[b]	Activity (U)[c]	
		PhoA	β-Gal
None	pACYC184	8	60
None	pVM25 (*toxR*$^+$)	8	800
pVC8 (*tagA::phoA*)	pACYC184	10	60
pVC8 (*tagA::phoA*)	pVM25 (*toxR*$^+$)	10	800

[a] pVC8 is shown in Fig. 1.
[b] pVM25 expressed *toxR* from the *tet* promoter of the pACYC184 vector.
[c] PhoA and β-galactosidase (β-Gal) activities were assayed as described previously (16, 23).

DISCUSSION

In an effort to understand the nature of the genes controlled by ToxR in *V. cholerae* and the molecular mechanism of that control, we have cloned and characterized the ToxR-regulated *tagA::phoA* fusion from strain KP8.56. Nucleotide sequence analysis of the 5′ part of *tagA*, i.e., upstream from the junction with Tn*phoA*, indicates that the TagA moiety of the hybrid protein should consist of 69 amino acid residues, the sequence of which does not present a statistically significant similarity with any other protein sequence. The N-terminal part of TagA shows a stretch of hydrophobic and nonpolar residues, which is typical of a signal sequence for exported proteins. At the end of this potential signal sequence, there is the motif Leu-Val-Gly-Cys, which is similar to the processing site of lipoproteins (33). The presence of this motif suggests that *tagA* encodes a lipoprotein, which has been confirmed by [^{3}H]palmitate labeling studies (22a).

Upstream from and in opposite orientation to *tagA*, we have detected an ORF whose deduced amino acid sequence is highly similar to the sequences of Ald dehydrogenase from humans, horses, and *A. nidulans*, and, to a lesser extent, to the sequence of P5C dehydrogenase from *S. cerevisiae*. Since the number of identical residues detected between the sequences of the *V. cholerae* ORF and the *A. nidulans* Ald dehydrogenase (taken here as a representative of any Ald dehydrogenase) is twice that obtained when comparing the *V. cholerae* ORF (or any Ald dehydrogenase) and the P5C dehydrogenase sequences, we concluded that this ORF corresponds to the *V. cholerae* gene encoding Ald dehydrogenase and have therefore named it *aldA*. Insertion inactivation of the *aldA* gene led to an undetectable level of Ald dehydrogenase activity, demonstrating that this gene indeed encodes an Ald dehydrogenase. Since addition of CoA in the assay did not enhance Ald dehydrogenase activity, we conclude that the *aldA* gene product is a CoA-independent enzyme, which is consistent with the sequence similarity detected with the eukaryotic, CoA-independent enzymes. Whether there are some sequence similarities between the CoA-dependent and CoA-independent Ald dehydrogenases is an open question, as the sequence of a CoA-dependent Ald dehydrogenase has not been determined. Inasmuch as Southern blot analysis did not reveal the presence of a second copy of *aldA* in the *V. cholerae* chromosome (data not shown) and no Ald dehydrogenase activity (whether CoA independent or CoA dependent) was detected in the extract of the *aldA1* mutant, the *aldA* gene seems to be the only *V. cholerae* gene encoding an Ald dehydrogenase.

The similarity detected here between the sequences of Ald dehydrogenase and P5C dehydrogenase indicates that these two enzymes share a common evolutionary origin. Such amino acid sequence similarities between enzymes catalyzing analogous reactions on different substrates are generally interpreted in terms of the substrate ambiguity of the common ancestor (for a review, see reference 22). Consistent with this evolutionary scheme is the presence of a Cys residue (Fig. 5) at the same position in the *V. cholerae* ORF, *A. nidulans* Ald dehydrogenase, and *S. cerevisiae* P5C dehydrogenase sequences; the corresponding Cys residue (Cys-302) of the human Ald dehydrogenase is sensitive to the inhibitor disulfiram, reacts with iodoacetamide, and has therefore been proposed to be part of the enzyme active site (5, 6).

By use of an *aldA-lacZ* transcriptional fusion, we have shown that the expression of *aldA* is under the control of

ToxR in *V. cholerae*. Between bacteria grown at pHs 6.5 and 8.4, we observed an induction ratio of 14 for the *aldA-lacZ* fusion and an induction ratio of 84 for the *tagA*::*phoA* fusion, thus demonstrating a similar environmental control on the *tagA* and *aldA* promoters. That this regulation is under the general control of ToxR is indicated by the lack of expression of the *aldA-lacZ* and *tagA*::*phoA* fusions in the *toxR* strain JJM43.

Deletion analysis located the *tagA* regulatory region between the junction with Tn*phoA* and the first *Xba*I site, a restriction site that turned out to be within the *aldA* coding sequence (nucleotide 601 in Fig. 4). The *tagA-aldA* intergenic region, which most probably contains the two divergent *tagA* and *aldA* promoters, consists of 212 bp. The heptanucleotide 5′-TTTTGAT-3′ that was identified as the tandemly repeated target of ToxR in the *ctx* regulatory region (20) is present as a single copy in the *tagA-aldA* intergenic region (shown as 5′-ATCAAAA-3′, nucleotides 394 to 400 in Fig. 4); however, as it is located only 18 nucleotides upstream from the *aldA* translation start site, it is not likely to be involved in the activation of the *aldA* promoter. Since the *tagA-aldA* intergenic region is 79% A-T, the presence of this heptanucleotide could be merely due to the overall A-T richness of this region.

These data raise several questions on the nature and number of targets and regulatory molecules involved in the control of *tagA* and *aldA* expression. That these genes should be dependent on a transcriptional activator is suggested by the lack of expression of the *tagA*::*phoA* and *aldA-lacZ* fusions in a *V. cholerae toxR* strain and in *E. coli*. Concerning the nature of this putative transcriptional activator, the lack of activation of the expression of the *tagA*::*phoA* and *aldA-lacZ* fusions by ToxR in *E. coli* suggests that, contrary to the *ctx* promoter, the *tagA* and *aldA* promoters are not directly activated by ToxR or are not activated by ToxR alone. In this regard, we have recently identified a locus called *toxT*, whose product(s) activates the *tagA* and *aldA* promoters in *E. coli* and which is itself controlled by ToxR in *V. cholerae* (2a). Elucidation of the *cis*-acting element(s) involved in the regulation of *tagA* and *aldA* expression will require a fine deletion analysis and the characterization of regulatory mutations.

The methodology previously used to detect ToxR-regulated genes involved the screening of Tn*phoA* insertion mutants expressing active PhoA hybrid proteins, thereby focusing the search on the genes encoding membrane or exported proteins. The rationale for this strategy was that the proportion of virulence genes should be enriched in that gene subset, and, indeed, the expression of almost 10% of the active PhoA hybrid proteins thus screened turned out to be under the control of ToxR; characterization of the Tn*phoA* insertion sites in these mutants allowed the identification of 17 different loci (23). The *aldA* gene described here thus appears as the first ToxR-regulated gene which does not encode a membrane or exported protein. The total number of genes which belong to the ToxR regulon may therefore be more numerous than the 17 identified so far. It is also interesting that *aldA* may be the first example of a gene encoding a cytoplasmic, metabolic enzyme the expression of which is coordinately regulated with virulence factors.

The importance of *aldA* in the virulence properties of *V. cholerae* was tested by using the *V. cholerae aldA* strain constructed by inserting a 1.4-kb DNA fragment into the *aldA* coding sequence. The mutant should express a polypeptide truncated at residue 233, i.e., before the Cys residue of the proposed active site of Ald dehydrogenase. The *aldA* mutant was not inhibited by the wild-type strain O395 Sm when tested in the infant mouse model. The colonization phenotype of the *aldA* mutant thus appears to be similar to that of the mutant with the Tn*phoA* insertion in *tagA* (23). *S. typhimurium* mutants lacking the CoA-dependent Ald dehydrogenase can no longer use ethanolamine as a carbon source but retain their ability to use it as a nitrogen source (25). On the other hand, the metabolic role of the CoA-independent Ald dehydrogenase, which is the eukaryotic enzyme involved in the assimilation of ethanol in bacteria, has not been characterized. In a small survey, we did not find that the *aldA1* mutation affected the growth of *V. cholerae* on 16 different substrates (including ethanolamine, acetate, glycerol, glucose, mannitol, pyruvate, and a variety of amino acids) under anaerobic or aerobic conditions (data not shown).

The fact that the *V. cholerae aldA* mutant was not impaired in its colonization properties when tested in the infant mouse model could be due to a number of reasons: (i) the presence of another gene encoding an Ald dehydrogenase, the expression of which would be induced in vivo but not in vitro, (ii) host specificity, such that an intact *aldA* gene would be required for colonization of mature animals or humans but not of suckling mice, (iii) the irrelevance of *aldA* expression to *V. cholerae* pathogenesis, although this would be surprising since expression of this gene is regulated by the same system that controls expression of the virulence genes in the *ctx*, *tcp*, and *acf* operons. We also have to keep in mind that *V. cholerae* is not an obligate pathogen but can survive and multiply in aquatic environments in which the expression of some genes could also be under the control of ToxR. Perhaps expression of *aldA* and *tagA* contributes to the colonization properties of *V. cholerae* in an as yet unidentified aquatic host or even to the utilization of unusual environmental subtrates.

ACKNOWLEDGMENTS

We are pleased to acknowledge A. Roberts for technical assistance in the assay of Ald dehydrogenase activity and V. DiRita, K. Peterson, and I. Saint-Girons for helpful discussions.

This work was supported by Public Health Service grants AI-18405 and AI-26289. C.P. was a Research Fellow from the Institut Pasteur.

REFERENCES

1. **Black, S.** 1955. Potassium-activated yeast aldehyde dehydrogenase. Methods Enzymol. **1**:508–511.
2. **Braun, T., E. Bober, S. Singh, P. A. Agarwal, and H. W. Goedde.** 1987. Isolation and sequence analysis of a full length cDNA clone coding for human mitochondrial aldehyde dehydrogenase. Nucleic Acids Res. **15**:3179.

2a. **DiRita, V. J., C. Parsot, G. Jander, and J. J. Mekalanos.** Proc. Natl. Acad. Sci., in press.

3. **Freter, R., P. C. M. O'Brien, and M. M. S. Macsai.** 1981. Role of chemotaxis in the association of motile bacteria with intestinal mucosa: in vivo studies. Infect. Immun. **34**:234–240.
4. **Friedman, A. M., S. R. Long, S. E. Brown, W. J. Buikema, and F. M. Ausubel.** 1982. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. Gene **18**:289–296.
5. **Hempel, J., R. Pietruszko, P. Fietzek, and H. Jörnvall.** 1982. Identification of a segment containing a reactive cysteine residue in human liver cytoplasmic aldehyde dehydrogenase (isoenzyme E_1). Biochemistry **21**:6834–6838.
6. **Hempel, J., H. von Bahr-Lindström, and H. Jörnvall.** 1984. Aldehyde dehydrogenase from human liver. Primary structure of the cytoplasmic enzyme. Eur. J. Biochem. **141**:21–35.

7. **Jones, P. W., and J. M. Turner.** 1984. Interrelationships between the enzymes of ethanolamine metabolism in *Escherichia coli*. J. Gen. Microbiol. **130**:299–308.
8. **Jones, P. W., and J. M. Turner.** 1984. A model for the common control of enzymes of ethanolamine catabolism in *Escherichia coli*. J. Gen. Microbiol. **130**:849–860.
9. **Krzywicki, K. A., and M. C. Brandriss.** 1984. Primary structure of the nuclear *PUT2* gene involved in the mitochondrial pathway for proline utilization in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:2837–2842.
10. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
11. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
12. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Manoil, C., and J. Beckwith.** 1985. Tn*phoA*: a transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA **82**:8129–8133.
14. **Mekalanos, J. J.** 1985. Cholera toxin: genetic analysis, regulation and role in pathogenesis. Curr. Top. Microbiol. Immunol. **118**:97–118.
15. **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.
16. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Miller, V. L., V. J. DiRita, and J. J. Mekalanos.** 1989. Identification of *toxS*, a regulatory gene whose product enhances ToxR-mediated activation of the cholera toxin promoter. J. Bacteriol. **171**:1288–1293.
18. **Miller, V. L., and J. J. Mekalanos.** 1984. Synthesis of cholera toxin is positively regulated at the transcriptional level by *toxR*. Proc. Natl. Acad. Sci. USA **81**:3471–3475.
19. **Miller, V. L., and J. J. Mekalanos.** 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. J. Bacteriol. **170**:2575–2583.
20. **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.
21. **Parsot, C.** 1986. Evolution of biosynthetic pathways: a common ancestor for threonine synthase, threonine dehydratase and D-serine dehydratase. EMBO J. **5**:3013–3019.
22. **Parsot, C., I. Saint-Girons, and G. N. Cohen.** 1987. Enzyme specialization during the evolution of amino acid biosynthetic pathways. Microbiol. Sci. **4**:258–262.
22a. **Parsot, C., E. Taxman, and J. J. Mekalanos.** 1991. ToxR regulates the production of lipoproteins and expression of serum resistance in *Vibrio cholerae*. Proc. Natl. Acad. Sci. USA **88**:1641–1645.
23. **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.
24. **Pickett, M., D. I. Gwynne, F. P. Buxton, R. Elliott, R. W. Davies, R. A. Lockington, C. Scazzochio, and H. M. Sealy-Lewis.** 1987. Cloning and characterization of the *aldA* gene of *Aspergillus nidulans*. Gene **51**:217–226.
25. **Roof, D. M., and J. R. Roth.** 1988. Ethanolamine utilization in *Salmonella typhimurium*. J. Bacteriol. **170**:3855–3863.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
27. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Simon, R., U. Priefer, and A. Puhler.** 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. Biotechnology **1**:784–791.
29. **Strauch, K. L., and J. Beckwith.** 1988. An *Escherichia coli* mutation preventing degradation of abnormal periplasmic proteins. Proc. Natl. Acad. Sci. USA **85**:1576–1580.
30. **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 cholerae toxin. Proc. Natl. Acad. Sci. USA **84**:2833–2837.
31. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
32. **von Bahr-Lindström, H., J. Hempel, and H. Jörnvall.** 1984. The cytoplasmic isoenzyme of horse liver aldehyde dehydrogenase: relationship to the corresponding human isoenzyme. Eur. J. Biochem. **141**:37–42.
33. **Wu, H. C., and M. Tokunaga.** 1986. Biogenesis of lipoproteins in bacteria. Curr. Top. Microbiol. Immunol. **125**:127–157.
34. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–109.