JOURNAL OF BACTERIOLOGY, Jan. 1992, p. 530–540
0021-9193/92/020530-11$02.00/0

Vol. 174, No. 2

Identification of Elements Involved in Transcriptional Regulation of the *Escherichia coli cad* Operon by External pH

NED WATSON,[1] DONNA S. DUNYAK,[1] EVERETT L. ROSEY,[1] JOAN L. SLONCZEWSKI,[2] AND ERIC R. OLSON[1]†*

Department of Molecular Biology, The Upjohn Company, Kalamazoo, Michigan 49001,[1] and Department of Biology, Kenyon College, Gambier, Ohio 43022[2]

Received 7 August 1991/Accepted 16 November 1991

Expression of the lysine decarboxylase gene (*cadA*) of *Escherichia coli* is induced upon external acidification. To dissect the molecular mechanisms responsible for this regulation, we analyzed a 4.2-kbp region upstream from *cadA*. DNA sequencing revealed two long open reading frames upstream of and on the same strand as *cadA*. One of these, *cadB*, is 444 codons long and is situated immediately upstream of *cadA*. Transcriptional fusions between fragments upstream of *cadA* and *lacZ*, Northern (RNA) hybridization, primer extension, and site-directed mutagenesis experiments defined a promoter, P*cad*, upstream of *cadB* that was responsible for pH-regulated expression of *cadA*. Upstream of P*cad* is an open reading frame, *cadC*, consisting of 512 codons. The predicted amino terminal region of the *cadC* gene product (CadC) resembles the carboxy-terminal domain of prokaryotic transcriptional activators involved in environmental sensing. Tn*10* insertions within or immediately upstream of *cadC* abolished P*cad* activity, suggesting that *cadC* encodes a positive transcription factor. Expression of plasmid-borne *cadC* in the Tn*10* mutants restored P*cad* activity, while introduction of a plasmid expressing truncated CadC resulted in the inability to complement. The presence of P*cad* on a multicopy plasmid was found to lower expression arising from chromosomal P*cad*, suggesting that a positive-acting factor is limiting. Our data suggests that *cadA*, *cadB*, and the acid-inducible P*cad* comprise, at least in part, the *cad* operon which is under control of the *cadC* product.

External pH has been implicated as a signal in a growing number of genetic responses in enteric bacteria. In a number of cases, acid-induced gene expression functions to decrease the acidity of bacterial products in an acidic environment, thus enhancing growth at low pH. For example, expression of genes encoding lysine decarboxylase (*cadA*) and arginine decarboxylase (*adi*) in *Escherichia coli* is dramatically increased at low external pH (2). These decarboxylases remove acid groups, resulting in less external acidification from metabolic processes. Similar effects may exist for deaminases at high external pH (11). Lactate dehydrogenase (encoded by *ldh*) activity is increased during fermentation, resulting in enhanced production of neutral end products (18). Induction of decarboxylases (2), lactate dehydrogenase (11), *aniG* and the *hyd* locus in *Salmonella typhimurium* (1) is enhanced during anaerobic growth, a condition that produces organic acids. In *Vibrio cholerae*, activity of the transcriptional regulator ToxR is, in part, regulated by changes in external pH (19).

The mechanism(s) employed by bacteria to sense and respond to external pH have not been elucidated. However, since the internal pH of *E. coli* remains relatively constant over a wide range of external pH (for reviews, see references 6 and 26) a transmembrane communication system coupled to factors involved in gene expression must exist.

Mud-*lacZ* fusions were used to isolate loci that respond transcriptionally to an external pH change, including the acid-inducible *exa-1* locus (30), the pH-dependent *ina* locus responsive to membrane-permeable weak acids (30), and the alkaline-inducible *alx* locus (4). Expression of β-galactosidase in the *exa-1* Mud fusion strain was maximal at an external pH around 5.8 and repressed above pH 6.8 (30). We report here the physical and genetic characterization of the region upstream of the Mud fusion in the *exa-1* strain, JLS8602. The *exa-1* allele was found to be a Mud insertion within the *cadA* gene, which encodes lysine decarboxylase. Similar fusions to *cadA* have been described and shown to be induced by low external pH (2). The promoter responsible for pH-regulated expression and two previously unidentified open reading frames, one of which encodes a product required for expression of *lacZ* in the *exa-1* strain, were identified. A working model is presented to explain how a change in external pH affects expression from the *cadA* promoter.

MATERIALS AND METHODS

Strains, growth conditions, and media. *E. coli*, phages, and plasmids used are shown in Table 1. LB (Luria-Bertani), per liter, was 10 g of Bacto Tryptone, 10 g of sodium chloride, and 5 g of yeast extract. LB-7.6 medium was LB buffered at pH 7.6 with a final concentration of 100 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS), and LB-5.8 was LB buffered at pH 5.8 with a final concentration of 100 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) (30). λ and P1 lysates were grown in LB plus 5 mM $CaCl_2$. The following antibiotics and concentrations were used: ampicillin sodium salt, 50 μg/ml; kanamycin sulfate, 50 μg/ml; chloramphenicol, 34 μg/ml; and tetracycline, 10 μg/ml. Antibiotics were present in all media in which cultures containing plasmids were grown. Cells harboring plasmids containing fragments from the pH region were grown and maintained on LB-7.6 with the appropriate antibiotic.

DNA biochemistry. Plasmid DNA was isolated as described previously (5) and in some instances purified by

* Corresponding author.
† Present address: Parke-Davis Research Division, Warner Lambert, Ann Arbor, MI 48105.

TABLE 1. Bacteria, phages, and plasmids

Strain, phage, or plasmid	Genotype and/or relevant characteristics[a]	Source or reference
E. coli		
W3110	F⁻ IN(*rrnD-rrnE*)	F. Neidhardt; 13
XL-1 Blue	*recA1 endA1 gyrA96 thi hsdR17 supE44 relA1 lac* F′ *proAB lacI*^{q} ZΔM15 Tn*10*	Stratagene Inc.
JLS821	F⁻ *araD139* Δ(*ara-leu*)*7697* Δ(*proAB-argF-lacIPOZYA*)XIII *rpsL* Nal^r	30
JLS8602	JLS821 *exa-1*::MudI1734	30
E608	*lac* F′lacI^q Cam^r	S. Rockenbach (The Upjohn Co.)
E2088	JLS8602 *cadC1*::Tn*10*	This work
E2110	JLS8602 *cadC2*::Tn*10*	This work
E2229	JLS8602 F′ *lacI*^{q} Cam^r (E608 × JLS8602)	This work
E2230	E2088 F′ *lacI*^{q} Cam^r	This work
Phages		
P1*vir*		D. Friedman
λEMBL3	Cloning vector	10
λA1-4	λEMBL containing *Sau*3A junction fragment from JLS8602	This work
λ5G7, λ21H11	Phage from Kohara collection with fragments from min 93 region	14
Plasmids		
pBS		Stratagene Inc.
pRS415		R. Simons (29)
pRS528		R. Simons (29)
pSR-P*lac*		S. Rockenbach (The Upjohn Co.)
pPH20	[pBS] *Eco*RV$^*_{2018}$-*Eco*RV$^*_{2579}$ [λ5G7]	This work
pPH32	[pBS] *Eco*RV$^*_{2579}$ *Pst*I$_{3751}$ [λ5G7]	This work
pPH33	[pBS] *Pvu*II$^*_{2227}$-*Pvu*II$^*_{3391}$ [λA1-4]	This work
pBS2133	[pBS] As pPH33, with −10 mutation	This work
pBS2138	[pBS] As pPH33, with −35 mutation	This work
pPH2241	[pBS] *Eco*RV$^*_{2579}$-*Pvu*II$^*_{3391}$ [λA1-4]	This work
pPH2242	[pBS] *Pvu*II$^*_{2277}$-*Eco*RV$^*_{2579}$ [λA1-4]	This work
pER2269	[pBS]*Rsa*I$^*_{2375}$-*Rsa*I$^*_{2809}$ [λA1-4]	This work
pER2270	[pBS]*Rsa*I$^*_{2375}$-*Eco*RV$^*_{2579}$ [λA1-4]	This work
pER2271	[pBS] *Eco*RV$^*_{2579}$-*Rsa*I$^*_{2809}$ [λA1-4]	This work
pPH1647	[pRS415] *Eco*RI-*Eco*RI$_{5024}$ [λ21H11]	This work
pPH1832	[pRS528] *Eco*RV$^*_{2579}$-*Eco*RI$_{5024}$ [λ5G7]	This work
pPH1856	[pRS415] *Eco*RV$^*_{28}$-*Eco*RV$^*_{2018}$ [λ5G7]	This work
pPH2089	[pRS415] *Pvu*II$^*_{2277}$-*Pvu*II$^*_{3391}$ [λA1-4]	This work
pPH2134	[pRS415] *Pvu*II$^*_{2277}$-*Pvu*II$_{3391}$ with −10 mutation [pBS2133]	This work
pPH2143	[pRS415] *Pvu*II$^*_{2277}$-*Pvu*II$_{3391}$ with −35 mutation [pBS2138]	This work
pPH2146	[pRS415] *Sal*I$^*_1$-*Sal*I$^*_{3711}$ [λA1-4]	This work
pPH2189	[pRS415] *Csp*451$^*_{854}$-*Sal*I$^*_{3711}$ [λA1-4]	This work
pPH2190	[pRS415] *Nco*I$^*_{633}$-*Sal*I$^*_{3711}$ [λA1-4]	This work
pPH2191	[pRS415] *Nsi*I$^*_{2283}$-*Sal*I$^*_{3711}$ [λA1-4]	This work
pPH2200	[pSR-P*lac*] *Ssp*I$^*_{692}$-*Ssp*I$^*_{2438}$ [λA1-4]	This work
pPH2249	[pSR-P*lac*] *Ssp*I$^*_{692}$-*Ssp*I$^*_{2438}$ with *Cla*I$_{1031}$ frameshift [λA1-4]	This work

[a] Plasmid descriptions are indicated by the following convention: [vector] insert endpoints [insert source]. Restriction sites correspond to the sequence shown in Fig. 2. Sites destroyed during cloning are marked with asterisks. In some cases, there were intermediate plasmids between the original source and the plasmid indicated. For reasons of clarity, these intermediates are not listed. Plasmid pPH1647 contains sequences beyond nucleotide number 1 shown in Fig. 2.

cesium chloride-ethidium bromide gradients (16). λ and chromosomal DNA were isolated as described previously (28). All ligations were performed overnight at 16°C in ligation buffer as described previously (16). Purified plasmid DNAs as well as ligation mixtures were introduced into *E. coli* made competent as described previously (16). Oligonucleotides were supplied by N. Hatzenbuhler (The Upjohn Co., Kalamazoo, Mich.). DNA sequencing was done on double-stranded templates with the Sequenase 2.0 kit (U.S. Biochemicals, Cleveland, Ohio). Both strands of the sequence shown in Fig. 2 as well as all junctions from subclones were determined from plasmids containing fragments from λ5G7 and λ21H11. Southern blots and hybridizations were done by transferring the DNA to nitrocellulose filters as described previously (16), prehybridizing for 1 h at 65°C in a solution consisting of 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 5× Denhardt's solution, 100 μg of denatured salmon sperm DNA per ml, and 0.1% sodium dodecyl sulfate (SDS) and hybridizing at 65°C overnight in a solution consisting of 5× SSC, 1× Denhardt's solution, 100 μg of denatured salmon sperm DNA per ml, and 0.1% SDS plus denatured labeled DNA probe, (see reference 16 for description, preparation, and use of SSC and Denhardt's solution). Filters were washed (three 15-min washes) in a solution consisting of 0.15 M NaCl, 0.03 M Tris-HCl (pH 8.0), 1 mM EDTA, and 0.1% SDS at 65°C, air dried, and exposed to X-ray film. Site-directed mutagenesis was done as described previously (17) using plasmid pPH33

as the template and oligonucleotide 5'CTTTTTTCGTG AGTGGTCTTATCGCCAGTTTG to create the −10 AA-to-GG change and 5'ACCTAATCTTTTCCTTTAATCCTTTT TTCGTG to create the −35 AGGA-to-TCCT alteration. Following mutagenesis and transformation, colonies were selected on plates containing LB plus ampicillin sodium salt, lifted onto nitrocellulose, and hybridized to the appropriate mutagenic oligonucleotide labeled at its 5' end with [^{32}P]ATP. Prehybridizations were performed at 65°C in a solution consisting of 5× Denhardt's solution and 0.5% SDS for 1 h. Hybridizations were performed overnight at 40°C in a solution consisting of 5× SSC, 1× Denhardt's solution, 0.5% SDS, and labeled oligonucleotide (1 μg; 10^6 cpm/μg). The filters were washed (three 30-min washes) in 2× SSC–0.5% SDS at 65°C, dried, and exposed to X-ray film. Plasmids from the strongly hybridizing colonies were analyzed and shown by DNA sequence analysis to contain the desired −10 (pBS2133) and −35 (pBS2138) mutations.

β-Galactosidase assays. For data shown in Tables 2, 3, and 4 and Fig. 3 and 6, the following procedures were used. Cells from an overnight LB-7.6 plate were resuspended in 1 ml of LB-7.6 to an optical density at 550 nm between 1.5 and 2.0. One hundred microliters of resuspended cells was used to inoculate 10 ml of LB-7.6 in a 50-ml Bellco culture flask and then incubated at 37°C in a New Brunswick model G76D water bath set at 200 rpm. When the culture reached an optical density at 550 nm of around 0.1, 5 ml were removed and centrifuged, and these cells were resuspended in 5 ml of LB-5.8 and then incubated in parallel with the remaining LB-7.6 culture for 2 h. Samples (0.1 or 0.01 ml) of the culture were used in the whole-cell β-galactosidase assay as described previously (22). For oxygen-limiting conditions (Table 3) 5 ml of cells were placed into screw cap tubes (20 by 125 mm) in a stationary 37°C water bath for 3 h. Tubes were gently shaken every 30 min to prevent cells from settling.

RNA analysis. RNA was isolated from log-phase cultures grown in either LB-7.6 or LB-5.8, denatured, and electrophoresed as previously described (25) and then transferred to nitrocellulose filters. Filters were prehybridized for 1 h at 65°C in a solution consisting of 5× SSC, 5× Denhardt's, 0.5% SDS, 10 μg of denatured salmon sperm DNA per ml, and 50% formamide. The filters were hybridized overnight at 42°C in prehybridization solution containing the radiolabeled probe (6×10^7 cpm of pPH20 and 3×10^7 cpm of pPH32), washed (three 15-min washes) in 1× SSC–0.1% SDS at 65°C, dried, and exposed to X-ray film. Probes used were [^{32}P]dCTP-labeled antisense RNA molecules prepared with T3 polymerase, as suggested by the manufacturer (Stratagene, La Jolla, Calif.). The plasmids used to generate the probes were pPH32 (linearized with *Bam*HI) and pPH20 (linearized with *Eco*RI). Plasmid pPH32 is pBS containing the *Eco*RV$_{2584}$-*Pst*I$_{3756}$ fragment. Plasmid pPH20 is pBS containing the *Eco*RV$_{2018}$-*Eco*RV$_{2584}$ fragment. Primer extension experiments were carried out using essentially the reagents and protocol for first-strand cDNA synthesis supplied in the cDNA Synthesis System Plus kit (1a) from Amersham Inc., Arlington Heights, Ill., with the following exceptions: (i) 5 μg of RNA and end-labeled oligonucleotide were coprecipitated with ethanol and resuspended in 25 μl of distilled water before addition of the remaining components, ii) 5 μl of 20 U of avian myeloblastosis virus reverse transcriptase (Seikagaku America Inc., St. Petersburg, Florida) per ml was used, iii) following synthesis, the reaction mixture was phenol extracted, ethanol precipitated, and resuspended in 10 μl of DNA sequencing loading dye. The products were electrophoresed on an 8% sequencing gel along with a sequencing ladder generated by using the same oligonucleotides and plasmid pPH33. The oligonucleotides used are described in Results.

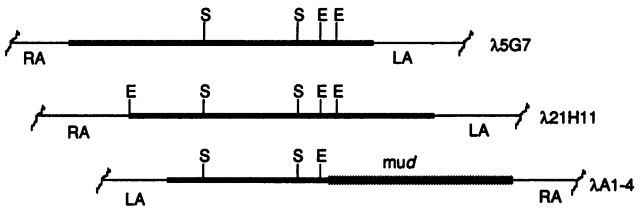

FIG. 1. Physical maps of recombinant λ derivatives carrying DNA from min 93.5 of the chromosome. Thick black lines represent chromosomal inserts. The thin black lines represent the phage arms. The stippled region represents the Mud portion of the chromosomal insert from JLS8602. RA and LA refer to the right and left arms of the λ cloning vectors, respectively. S and E, *Sal*I and *Eco*RI restriction sites, respectively. Not all of these sites are shown.

Tn*10* mutagenesis. JLS8602 was infected with λ1098 as described previously (34). Tet^r colonies were screened on LB-7.6 and LB-5.8 plates containing tetracycline (10 μg/ml) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) (40 μg/ml). Colonies showing altered phenotypes (different intensity of blue than JLS8602 on X-Gal plates) were purified on the same medium in which they exhibited their altered phenotype. Phage P1*vir* was grown in the mutants and used to transduce JLS8602 to Tet^r to verify that the phenotypes were like those of the original mutant. Two of these transductants (E2088 and E2110) from two independent mutants were analyzed further.

Nucleotide sequence accession number. The nucleotide sequence data reported in this study has been assigned GenBank accession number M67452.

RESULTS

Mapping and cloning of the *exa-1* fusion junction from JLS8602. Two approaches were taken to genetically and physically localize the *exa-1* locus (Mud transcription fusion) responsible for acid-inducible *lacZ* expression in JLS8602. The first approach was based on the observation that induction of *lacZ* in this strain resembled that of a strain containing a *cadA-lacZ* fusion (min 93.5) (2), suggesting that the Mud insertion was in the same operon as *cadA*. The *cadA* gene encodes lysine decarboxylase and has been shown to be regulated, in part, by external pH. Genetic mapping using Hfr and phage P1*vir*-mediated crosses demonstrated that the Mud insertion in JLS8602 was linked to *cadA* (data not shown). The second approach was to isolate a lambda clone that carried a fusion fragment between the chromosome and the Mud insertion and determine whether the clone contained DNA linked to *cadA*. A library of *Sau*3A partial digestion fragments from JLS8602 was constructed in λEMBL3, and a phage carrying *lacZ* was isolated by its ability to form blue plaques on a lawn of Lac$^-$ cells (JLS821) with X-Gal in the top agar. Restriction analysis of DNA from this phage, called λA1-4, showed that it carried a 15-kbp insert with 6 kbp from the chromosome flanking the end of the Mud insertion (Fig. 1).

A 3.7-kbp *Sal*I fragment from λA1-4 that mapped within the chromosomal insert and upstream of *lacZ* was isolated, nick translated, and hybridized to DNA from lambda clones that contained DNA derived from the min 93 region of the chromosome (14). This fragment hybridized to DNA from

λ5G7 and λ21H11 (data not shown). Restriction mapping and Southern hybridization experiments demonstrated that this *Sal*I fragment hybridized to a 3.7-kbp *Sal*I fragment found in both λ5G7 and λ21H11 and in chromosomal DNA from *E. coli* W3110 and JLS821 (the parent of JLS8602) (data not shown). Figure 1 shows the relationship between λ5G7, λ21H11, and λA1-4 as well as the location of the 3.7-kbp *Sal*I fragment and an adjoining *Eco*RI site common to all three clones. The *cadA* gene contains a 900-bp *Eco*RI fragment that is internal to the coding sequence (3). Lambda clones λ5G7 and λ21H11 contain this fragment, while λA1-4 does not (Fig. 1). Southern hybridization experiments of genomic DNA from JLS8602 and the parent strain JLS821 confirmed that this *Eco*RI fragment was present in the parent strain but missing in JLS8602 (data not shown). This analysis confirms that the Mud insertion lies within *cadA*.

DNA sequence analysis. The DNA sequence of the 3.7-kbp *Sal*I fragment plus 1.3 kbp of an adjoining *Sal*I-*Eco*RI fragment shown in Fig. 1 was determined and is shown in Fig. 2. The sequence begins at the *Sal*I site shown at the left in the clones shown in Fig. 1. Computer analysis revealed the presence of three long open reading frames on one strand and only small ones on the opposite strand. The three long open reading frames are shown in Fig. 2. The first open reading frame, which we have named *cadC* (see below), extends from a methionine codon at nucleotide 872 to nucleotide 2407 and would be predicted to encode a 512-amino-acid protein of 57.7 kDa. The second open reading frame extends from a methionine codon at nucleotide 2775 to nucleotide 4106 and would be predicted to encode a 444-amino-acid protein of 46.7 kDa. This open reading frame has also been identified by G. Bennett and colleagues who have proposed calling it *cadB* (3). The third open reading frame extends from a methionine codon at nucleotide 4189 past the sequence shown and corresponds to the *cadA* gene (3).

Localization of transcriptional elements by *lacZ* fusions. The promoter probe plasmids pRS415 and pRS528 (differing only in the orientation of the cloning sites upstream from *lacZYA*) were used to localize DNA fragments with promoter activity. Various fragments were cloned upstream from *lacZYA*, introduced into the Lac$^-$ strain JLS821 and the resulting strains assayed for β-galactosidase after growth at pH 5.8 and 7.6 (Fig. 3). Three plasmids, pPH2146, pPH2190, and pPH1647, produced substantial levels of β-galactosidase at pH 5.8, while plasmids pPH1856 and pPH1832 showed no induction at pH 5.8. These results suggested that sequences upstream of $Eco\text{RV}_{2579}$ were required for expression of *cadA* and that those upstream of the $Csp451_{854}$ site were needed for maximal expression in these plasmids. Three other plasmids, pPH2189, pPH2191, and pPH2089, had smaller, but reproducible, levels of low pH-induced expression.

Northern (RNA) hybridizations. The transcription pattern of the *cadC-cadB* region was examined by Northern hybridization. RNA isolated from JLS8602 and JLS821 grown at pH 7.6 and 5.8 was blotted and hybridized to two antisense RNA probes made from pBS derivatives containing fragment $Eco\text{RV}_{2018}$-$Eco\text{RV}_{2579}$ to produce pPH20 or fragment $Eco\text{RV}_{2579}$-$Pst\text{I}_{3751}$ to produce pPH32. Each of these fragments were oriented such that probes synthesized in vitro with the T3 promoter would be antisense to mRNA transcribing toward *cadA*. Following hybridization, the blots were stained and visually checked to ensure that approximately equal levels of rRNA were present in each lane. The results of the hybridization are shown in Fig. 4. This data suggests that transcription of an acid-induced RNA begins downstream (or very close) to the $Eco\text{RV}_{2579}$ site between *cadC* and *cadB*.

Primer extension. Consideration of the location of *cadB* and results of the Northern hybridizations suggested that the pH-inducible promoter was in the intercistronic region between *cadC* and *cadB*. Primer extension experiments using three oligonucleotides were carried out to test this hypothesis. Oligonucleotide 1 corresponds to nucleotides 2615 to 2635 and did not produce any detectable extension product (not shown). Oligonucleotide 2 corresponds to nucleotides 2758 to 2778 and produced an extension product ending at an A:T bp at nucleotide 2700 with template RNA from cells grown at pH 5.8 but not pH 7.6 (Fig. 5). Oligonucleotide 3 corresponds to nucleotides 2883 to 2903 and also produced a pH 5.8-specific extension product ending around nucleotide 2700 (not shown). These results suggest that differential transcription initiating around nucleotide 2700 is, at least in part, responsible for the observed pH-regulated expression of *lacZ* in JLS8602.

Examination of the sequence upstream from nucleotide 2700 revealed the presence of a hexanucleotide centered 10 bp away (5′TAATCT) that is similar to the consensus hexanucleotide (5′TATAAT) found in *E. coli* σ^{70}-dependent promoters (12). The most highly conserved bases among this class of promoters, the first, second and sixth nucleotides, are identical to those found in P*cad*. It is more difficult to assign on the basis of inspection a hexanucleotide 35 bp upstream from the start site that matches the consensus for σ^{70}-dependent promoters (5′TTGACA, with 17 bp most common between the −10 and −35 hexanucleotides) (12). The best sequence match to the consensus is 5′TTAGGA. However, the spacing between this and the −10 sequence is 19 bp.

To determine whether the −10 and −35 sequences are important for P*cad* function, the following mutations were introduced by site-directed mutagenesis: (i) the −10 sequence was changed from 5′TAATCT to 5′TGGTCT and (ii) the −35 sequence was changed from 5′TTAGGA to 5′TT TCCT. The mutations were introduced into the 1.1-kbp *Pvu*II fragment and placed upstream of *lacZ* in pRS415. Table 2 shows that both mutations resulted in the loss of pH-induced *lacZ* expression. Although the wild-type plasmid, pPH2089, produced a 3- to 4-fold induction (compared with pPH2146, which produced a 40-fold induction), we feel that this induction is physiologically significant, since a Tn*10* insertion that abolished *cadA* expression in the chromosome (*cadC1*::Tn*10*, see below) also abolished pH induction from pPH2089 (Table 2).

Taken together, the results of the Northern hybridizations, primer extension experiment, and site-directed mutagenesis led us to conclude that the sequence upstream from nucleotide 2700 defines a promoter that is responsible for pH-regulated expression of *cadA* and *cadB*. We define this promoter as P*cad*.

Isolation of Tn*10* mutants that alter regulation of P*cad*. Strain JLS8602 was mutagenized with Tn*10* and mutants that showed altered expression, as evidenced by differences in intensity of blue on LB–X-Gal plates at pH 5.8 were obtained. Two mutants that produced white colonies at pH 5.8 were analyzed further, since they represented insertion mutants that were not inducible by low pH. The Tn*10* insertions were transduced from the original isolates to JLS8602 (resulting in strains E2088 and E2110) and were found to confer the same phenotypes. Genetic mapping of the Tn*10*'s in E2088 and E2110 revealed that they were linked to the Mud insertion in *cadA* (around 50% of the Tet^r


SalI
EcoRV
GTCGACACTGCACATCGGCTGGCGGAGATATCCTGCGCTTCGCCCGCTGCCGCCAGGGT 60
GGTGCATTGACCTGACAGAAACACAGAAAAGAAGCGATTTGCCGCAATCTTAAGCAGTTG 120
AATCGCTTTTACTGAAATTAGGTTGACGAGATGTGCAGATTACGGTTTAATGCGCCCCGT 180
TGCCCGGATAGCTCAGTCGGTAGAGCAGGGGATTGAAAATCCCCGTGTCCTTGGTTCGAT 240
TCCGAGTCCGGGCACCAAATTCATATAAACGGACCTCCACGGAGGTCGTTTGCGTTTCAG 300
GGCGCCACGATTTAAGCGTTCTGCCTCCAAATCAATTCTACCGAACTCAACCAGATTCTC 360
CCCACATCACCAGCAATTTGCGGGCATATCCCAATTCGGGAAAATTTGTTCCTGAGCTAT 420
AGCACTAACTGACGTGAAATGTCGTGCGGCCCCGTGATGCTGTTGAACGTCAAATGACGT 480
CATCAGGAGCGTAACGCATCCATAAAGCACAACACCGGTCTGAACGGCGACTGATGAGAT 540
TTTCTGAATGAGAACAAAGAGAAATGTATCAGTTCGCTTGCTCATGCAAAGACTAACAAT 600
NcoI
CCATTAAAATAGTAAGTGCTCCGGACAATTTTCCATGGATTATTTTCTGAACATTTTTCT 660
SspI
TTGGCAAAGATGATGAATTTTGATGGTAAGGAATATTACTTCTGGTTCTCAGTAAAATCC 720
TTTCGTAATACTATGTAATCAAGAAGTTTATGGCTGGTAAAAATAACGTCTTGCATTCAC 780
CAATAATATGTAAATAAACCCATCTATAGATGGTAAAAATAGGTTGTGGCAATTATCATT 840
Csp451
GCATCATTCCCTTTTCGAATGAGTTTCTATTATGCAACAACCTGTAGTTCGCGTTGGCGA 900
CadC: MetGlnGlnProValValArgValGlyGl
ATGGCTTGTTACTCCGTCCATAAACCAAATTAGCCGCAATGGGCGTCAACTTACCCTTGA 960
uTrpLeuValThrProSerIleAsnGlnIleSerArgAsnGlyArgGlnLeuThrLeuGl
ClaI
GCCGAGATTAATCGATCTTCTGGTTTTCTTTGCTCAACACAGTGGCGAAGTACTTAGCAG 1020
uProArgLeuIleAspLeuLeuValPhePheAlaGlnHisSerGlyGluValLeuSerAr
ClaI
GGATGAACTTATCGATAATGTCTGGAAGAGAAGTATTGTCACCAATCACGTTGTGACGCA 1080
gAspGluLeuIleAspAsnValTrpLysArgSerIleValThrAsnHisValValThrGl
GAGTATCTCAGAACTACGTAAGTCATTAAAAGATAATGATGAAGATAGTCCTGTCTATAT 1140
nSerIleSerGluLeuArgLysSerLeuLysAspAsnAspGluAspSerProValTyrIl
CGCTACTGTACCAAAGCGCGGCTATAAATTAATGGTGCCGGTTATCTGGTACAGCGAAGA 1200
eAlaThrValProLysArgGlyTyrLysLeuMetValProValIleTrpTyrSerGluGl
AGAGGGAGAGGAAATAATGCTATCTTCGCCTCCCCCTATACCAGAGGCGGTTCCTGCCAC 1260
uGluGlyGluGluIleMetLeuSerSerProProProIleProGluAlaValProAlaTh
AGATTCTCCCTCCCACAGTCTTAACATTCAAAACACCGCAACGCCACCTGAACAATCCCC 1320
rAspSerProSerHisSerLeuAsnIleGlnAsnThrAlaThrProProGluGlnSerPr
AGTTAAAAGCAAACGATTCACTACCTTTTGGGTATGGTTTTTTTTCCTGTTGTCGTTAGG 1380
oValLysSerLysArgPheThrThrPheTrpValTrpPhePhePheLeuLeuSerLeuGl
TATCTGTGTAGCACTGGTAGCGTTTTCAAGTCTTGATACACGTCTTCCTATGAGCAAATC 1440
yIleCysValAlaLeuValAlaPheSerSerLeuAspThrArgLeuProMetSerLysSe
PvuII
GCGTATTTTGCTCAATCCACGCGATATTGACATTAATATGGTAAATAAAAGTTGTAACAG 1500
rArgIleLeuLeuAsnProArgAspIleAspIleAsnMetValAsnLysSerCysAsnSe
CTGGAGTTCCCCGTATCAGCTCTCTTACGCGATAGGCGTGGGTGATTTGGTGGCGACATC 1560
rTrpSerSerProTyrGlnLeuSerTyrAlaIleGlyValGlyAspLeuValAlaThrSe
ACTTAACACCTTCTCCACCTTTATGGTGCATGACAAAATCAACTACAACATTGATGAACC 1620
rLeuAsnThrPheSerThrPheMetValHisAspLysIleAsnTyrAsnIleAspGluPr
GAGCAGTTCCGGTAAAACATTATCTATTGCGTTTGTTAATCAGCGCCAATACCGTGCTCA 1680
oSerSerSerGlyLysThrLeuSerIleAlaPheValAsnGlnArgGlnTyrArgAlaGl
ACAATGCTTTATGTCGATAAAATTGGTAGACAATGCAGATGGTTCAACCATGCTGGATAA 1740
nGlnCysPheMetSerIleLysLeuValAspAsnAlaAspGlySerThrMetLeuAspLy
PvuII
ACGTTATGTCATCACTAACGGTAATCAGCTGGCGATTCAAAATGATTTACTGGAGAGTTT 1800
sArgTyrValIleThrAsnGlyAsnGlnLeuAlaIleGlnAsnAspLeuLeuGluSerLe
ATCAAAAGCGTTAAACCAACCGTGGCCACAACGAATGCAGGAGACGCTCCAGAAAATTTT 1860
uSerLysAlaLeuAsnGlnProTrpProGlnArgMetGlnGluThrLeuGlnLysIleLe
GCCGCATCGTGGTGCGTTATTAACTAATTTTTATCAGGCACATGATTATTTACTGCATGG 1920
uProHisArgGlyAlaLeuLeuThrAsnPheTyrGlnAlaHisAspTyrLeuLeuHisGl
CGATGATAAATCATTGAACCGTGCCAGTGAATTATTAGGTGAGATTGTTCAATCATCCCC 1980
yAspAspLysSerLeuAsnArgAlaSerGluLeuGlyGluLeuValGlnSerSerPr
EcoRV
AGAATTTACCTACGCGAGAGCAGAAAAAGCATTAGTTGATATCGTGCGCCATTCTCAACA 2040
oGluPheThrTyrAlaArgAlaGluLysAlaLeuValAspIleValArgHisSerGlnHi
TCCTTTAGATGAAAAACAATTAGCAGCACTGAACACAGAAATAGATAACATTGTTACACT 2100
sProLeuAspGluLysGlnLeuAlaAlaLeuAsnThrGluIleAspAsnIleValThrLe
GCCGGAATTGAACAACCTGTCCATTATATATCAAATAAAAGCGGTCAGTGCTCTGGTAAA 2160
uProGluLeuAsnAsnLeuSerIleIleTyrGlnIleLysAlaValSerAlaLeuValLy
AGGTAAAACAGATGAGTCTTACCAGGCGATAAATACTGGCATTGATCTTGAAATGTCCTG 2220
sGlyLysThrAspGluSerTyrGlnAlaIleAsnThrGlyIleAspLeuGluMetSerTr
PvuII
GCTAAATTATGTGTTGCTTGGCAAGGTTTATGAAATGAAGGGGATGAACCGGGAAGCAGC 2280
pLeuAsnTyrValLeuLeuGlyLysValTyrGluMetLysGlyMetAsnArgGluAlaAl
NsiI
TGATGCATATCTCACCGCCTTTAATTTACGCCCAGGGGCAAACACCCTTTACTGGATTGA 2340
aAspAlaTyrLeuThrAlaPheAsnLeuArgProGlyAlaAsnThrLeuTyrTrpIleGl
RsaI
AAATGGTATATTCCAGACTTCTGTTCCTTATGTTGTACCTTATCTCGACAAATTTCTTGC 2400
uAsnGlyIlePheGlnThrSerValProTyrValValProTyrLeuAspLysPheLeuAl
SspI
TTCAGAATAAGTAACTCCGGGTTGATTTATGCTCGGAAATATTTGTTGTTGAGTTTTTGT 2460
aSerGlu###
ATGTTCCTGTTGGTATAATATGTTGCGGCAATTTATTTGCCGCATAATTTTTATTACATA 2520
EcoRV
AATTTAACCAGAGAATGTCACGCAATCCATTGTAAACATTAAATGTTTATCTTTTCATGA 2580
TATCAACTTGCGATCCTGATGTGTTAATAAAAAACCTCAAGTTCTCACTTACAGAAACTT 2640
-35 -10 +1
TTGTGTTATTTCACCTAATCTTTAGGATTAATCCTTTTTTCGTGAGTAATCTTATCGCCA 2700
GTTTGGTCTGGTCAGGAAATAGTTATACATCATGACCCGGACTCCAAATTCAAAAATGAA 2760
RsaI
ATTAGGAGAAGAGCATGAGTTCTGCCAAGAAGATCGGGCTATTTGCCTGTACCGGTGTTG 2820
CadB: MetSerSerAlaLysLysIleGlyLeuPheAlaCysThrGlyValV
TTGCCGGTAATATGATGGGGAGCGGTATTGCATTATTACCTGCGAACCTAGCAAGTATCG 2880
alAlaGlyAsnMetMetGlySerGlyIleAlaLeuLeuProAlaAsnLeuAlaSerIleG
GTGGTATTGCTATCTGGGGTTGGATTATCTCTATTATTGGTGCAATGTCGCTGGCGTATG 2940
lyGlyIleAlaIleTrpGlyTrpIleIleSerIleIleGlyAlaMetSerLeuAlaTyrV
TATATGCCCGACTGGCAACAAAAAACCCGCAACAAGGTGGCCCAATTGCTTATGCCGGAG 3000
alTyrAlaArgLeuAlaThrLysAsnProGlnGlnGlyGlyProIleAlaTyrAlaGlyG
AAATTTCCCCTGCATTTGGTTTTCAGACAGGTGTTCTTTATTACCATGCTAACTGGATTG 3060
luIleSerProAlaPheGlyPheGlnThrGlyValLeuTyrTyrHisAlaAsnTrpIleG
GTAACCTGGCGATTGGTATTACCGCTGTATCTTATCTTTCCACCTTCTTCCCAGTATTAA 3120
lyAsnLeuAlaIleGlyIleThrAlaValSerTyrLeuSerThrPhePheProValLeuA
ATGATCCTGTTCCGGCGGGTATCGCCTGTATTGCTATCGTCTGGGTATTTACCTTTGTAA 3180
snAspProValProAlaGlyIleAlaCysIleAlaIleValTrpValPheThrPheValA
ATATGCTCGGCGGTACTTGGGTAAGCCGTTTAACCACTATTGGTCTGGTGCTGGTTCTTA 3240
snMetLeuGlyGlyThrTrpValSerArgLeuThrThrIleGlyLeuValLeuValLeuI
TTCCTGTGGTGATGACTGCTATTGTTGGCTGGCATTGGTTTGATGCGGCAACTTATGCAG 3300
leProValValMetThrAlaIleValGlyTrpHisTrpPheAspAlaAlaThrTyrAlaA
CTAACTGGAATACTGCGGATACCACTGATGGTCATGCGATCATTAAAAGTATTCTGCTCT 3360
laAsnTrpAsnThrAlaAspThrThrAspGlyHisAlaIleIleLysSerIleLeuLeuC
PvuII
GCCTGTGGGCCTTCGTGGGTGTTGAATCCGCAGCTGTAAGTACTGGTATGGTTAAAAACC 3420
ysLeuTrpAlaPheValGlyValGluSerAlaAlaValSerThrGlyMetValLysAsnP
CGAAACGTACCGTTCCGCTGGCAACCATGCTGGGTACTGGTTTAGCAGGTATTGTTTACA 3480
roLysArgThrValProLeuAlaThrMetLeuGlyThrGlyLeuAlaGlyIleValTyrI
TCGCTGCGACTCAGGTGCTTTCCGGTATGTATCCGTCTTCTGTAATGGCGGCTTCCGGTG 3540
leAlaAlaThrGlnValLeuSerGlyMetTyrProSerSerValMetAlaAlaSerGlyA
CTCCGTTGCAATCAGTGCTTCAACTATCCTCGGTAACTGGGCTGCGCCGCTGGTTTCTG 3600
laProHeAlaIleSerAlaSerThrIleLeuGlyAsnTrpAlaAlaProLeuValSerA
CATTCACCGCCTTTGCGTGCCTGACTTCTCTGGGCTCCTGGATGATGTTGGTAGGCCAGG 3660
laPheThrAlaPheAlaCysLeuThrSerLeuGlySerTrpMetMetLeuValGlyGlnA
SalI
CAGGTGTACGTGCCGCTAACGACGGTAACTTCCCGAAAGTTTATGGTGAAGTCGACAGCA 3720
laGlyValArgAlaAlaAsnAspGlyAsnPheProLysValTyrGlyGluValAspSerA
PstI
ACGGTATTCCGAAAAAAGGTCTGCTGCTGGCTGCAGTGAAAATGACTGCCCTGATGATCC 3780
snGlyIleProLysLysGlyLeuLeuLeuAlaAlaValLysMetThrAlaLeuMetIleL
TTATCACTCTGATGAACTCTGCCGGTGGTAAAGCATCTGACCTGTTCGGTGAACTGACCG 3840
euIleThrLeuMetAsnSerAlaGlyGlyLysAlaSerAspLeuPheGlyGluLeuThrG
GTATCGCAGTACTGCTGACTATGCTGCCGTATTTCTACTCTTGCGTTGACCTGATTCGTT 3900
lyIleAlaValLeuLeuThrMetLeuProTyrPheTyrSerCysValAspLeuIleArgP
TTGAAGGCGTTAACATCCGCAACTTTGTCAGCCTGATCTGCTCTGTACTGGGTTGCGTGT 3960
heGluGlyValAsnIleArgAsnPheValSerLeuIleCysSerValLeuGlyCysValP
TCTGCTTCATCGCGCTGATGGGCGCAAGCTCCTTCGAGCTGGCAGGTACCTTCATCGTCA 4020
heCysPheIleAlaLeuMetGlyAlaSerSerPheGluLeuAlaGlyThrPheIleValS
GCCTGATTATCCTGATGTTCTACGCTCGCAAAATGCACGAGCGCCAGAGCCACTCAATGG 4080
erLeuIleIleLeuMetPheTyrAlaArgLysMetHisGluArgGlnSerHisSerMetA
ATAACCACACCGGCGTCTAACGCACATTAATTAAAAGTATTTTCCGAGGCTCCTCCTTTCA 4140
spAsnHisThrAlaSerAsnAlaHis###
TTTTGTCCCATGTGTTGGGAGGGGCCTTTTTTACCTGGAGATATGACTATGAACGTTATT 4200
CadA: MetAsnValIle
SspI
GCAATATTGAATCACATGGGGGTTTATTTTAAAGAAGAACCCATCCGTGAACTTCATCGC 4260
AlaIleLeuAsnHisMetGlyValTyrPheLysGluGluProIleArgGluLeuHisArg
GCGCTTGAACGTCTGAACTTCCAGATTGTTTACCCGAACGACCGTGACGACTTATTAAAA 4320
AlaLeuGluArgLeuAsnPheGlnIleValTyrProAsnAspArgAspAspLeuLeuLys
CTGATCGAAAACAATGCGCGTCTGTGCGGCGTTATTTTTGACTGGGATAAATATAATCTC 4380
LeuIleGluAsnAsnAlaArgLeuCysGlyValIlePheAspTrpAspLysTyrAsnLeu
GAGCTGTGCGAAGAAATTAGCAAAATGAACGAGAACCTGCCGTTGTACGCGTTCGCTAAT 4440
GluLeuCysGluGluIleSerLysMetAsnGluAsnLeuProLeuTyrAlaPheAlaAsn
ACGTATTCCACTCTCGATGTAAGCCTGAATGACCTGCGTTTACAGATTAGCTTCTTTGAA 4500
ThrTyrSerThrLeuAspValSerLeuAsnAspLeuArgLeuGlnIleSerPheGlu
TATGCGCTGGGTGCTGCTGAAGATATTGCTAATAAGATCAAGCAGACCACTGACGAATAT 4560
TyrAlaLeuGlyAlaAlaGluAspIleAlaAsnLysIleLysGlnThrThrAspGluTyr
ATCAACACTATTCTGCCTCCGCTGACTAAAGCACTGTTTAAATATGTTCGTGAAGGTAAA 4620
IleAsnThrIleLeuProProLeuThrLysAlaLeuPheLysTyrValArgGluGlyLys

```
TATACTTTCTGTACTCCTGGTCACATGGGCGGTACTGCATTCCAGAAAAGCCCGGTAGGT   4680
TyrThrPheCysThrProGlyHisMetGlyGlyThrAlaPheGlnLysSerProValGly

AGCCTGTTCTATGATTTCTTTGGTCCGAATACCATGAAATCTGATATTTCCATTTCAGTA   4740
SerLeuPheTyrAspPheGlyProAsnThrMetLysSerAspIleSerIleSerVal

TCTGAACTGGGTTCTCTGCTGGATCACAGTGGTCCCACACAAAGAAGCAGAACAGTATATC   4800
SerGluLeuGlySerLeuLeuAspHisSerGlyProHisLysGluAlaGluGlnTyrIle

GCTCGCGTCTTTAACGCAGACCGCAGCTACATGGTGACCAACGGTACTTCCACTGCGAAC   4860
AlaArgValPheAsnAlaAspArgSerTyrMetValThrAsnGlyThrSerThrAlaAsn

AAAATTGTTGGTATGTACTCTGCTCCAGCAGGCAGCACCATTCTGATTGACCGTAACTGC   4920
LysIleValGlyMetTyrSerAlaProAlaGlySerThrIleLeuIleAspArgAsnCys

CACAAATCGCTGACCCACCTGATGATGATGAGCGATGTTACGCCAATCTATTTCCGCCCG   4980
HisLysSerLeuThrHisLeuMetMetMetSerAspValThrProIleTyrPheArgPro
                                    EcoRI
ACCCGTAACGCTTACGGTATTCTTGGTGGTATCCCACAGAGTGAATTC···            5028
ThrArgAsnAlaTyrGlyIleLeuGlyGlyIleProGlnSerGluPhe···
```

FIG. 2. DNA sequence of a portion of the chromosomal inserts in λ5G7 and λ21H11. The sequence begins at the leftmost *Sal*I site found in common between the phage shown in Fig. 1 and ends at the common *Eco*RI site. Selected restriction sites utilized in this work are shown above the sequence. *Cla*I^m is a *Cla*I site that is overlapped by a methylation site. The start site of transcription (+1) and the −10 and −35 sequences are indicated by double underlines, while underlined sequences correspond to the restriction site listed above the sequence. The predicted amino acids corresponding to three open reading frames (*cadC*, *cadB*, and *cadA*) are indicated below the sequence.

transductants resulting from crosses between E2088 or E2110 and JLS821 were Kan^r). Southern hybridization of chromosomal DNA showed that the Tn*10* insertion in E2088 was localized to the *Ssp*I$_{692}$-*Cla*I$_{1031}$ fragment and that of E2110 was found to be in the *Cla*I$_{1031}$-*Eco*RV$_{2018}$ fragment (data not shown). The mutations were given the allele numbers *cadC1*::Tn*10* and *cadC2*::Tn*10* corresponding to the mutations in E2088 and E2110, respectively. The levels of β-galactosidase in these mutants (E2088 and E2110) when grown in liquid culture are shown in Table 3. Since these insertions were well upstream of P*cad*, it suggested that an element upstream of P*cad* was required for induction. This result was consistent with those from the promoter probe cloning experiments in which only plasmids carrying P*cad* plus a large upstream region gave maximal pH inductions (Fig. 3). A model explaining these results is presented in the discussion.

The *cadC* gene product, CadC, acts in *trans* to activate P*cad*. The results described above suggested that an upstream genetic element was required for P*cad* induction. This element could either encode a *trans*-acting factor or be a site needed for optimal promoter activity (enhancer). A likely candidate for a *trans*-acting factor would be the putative *cadC* gene product. The different possibilities were distinguished by testing for the ability of the upstream region to complement the *cadC1*::Tn*10* allele in *trans*. Plasmid pPH2200 contains an *Ssp*I$_{692}$-*Ssp*I$_{2443}$ fragment extending from 177 bp upstream to 33 bp downstream of the *cadC* open reading frame inserted downstream of the *lac* promoter in plasmid pSR-P*lac*. Plasmid pSR-P*lac* also contains transcription termination signals on each side of the *lac* promoter. Plasmid pPH2249 differs from pPH2200 by the presence of a frameshift mutation at the *Cla*I site at nucleotide 1031 (53 amino acids from the amino terminus of CadC). This mutation was introduced by utilizing the Klenow fragment of DNA polymerase I and nucleotides to fill in the *Cla*I site and performing intramolecular ligation.

Plasmids pPH2200 and pPH2249 were introduced into E2229 and E2230, both of which contained the P*cad-lacZ* chromosomal fusion from JLS8602 and an F′ containing *lacI*^{q} (encoding *lac* repressor). In addition, E2230 contained the *cadC1*::Tn*10* allele. These strains, with and without pPH2200 and pPH2249, were assayed for β-galactosidase after growth at pH 5.8 and 7.6 (Table 4). Several conclusions were drawn from this data. First, plasmid pPH2200 encodes a product that functions in *trans* to activate P*cad*. Second, that product is most likely a protein, since introduction of a frameshift mutation within the *cadC* open reading frame destroyed its ability to complement. Third, since the presence of pPH2200 increased expression in the wild-type background, CadC must normally be limiting for P*cad* activation. Although this data is for strains grown without the P*lac* inducer, isopropyl-β-D-thiogalactopyranoside (IPTG), its addition had no substantial effect on expression levels detected from the chromosomal P*cad-lacZ* fusion (data not shown). Thus, the 177 bp between the *Ssp*I$_{692}$ site and the start of the *cadC* coding region may contain a promoter capable of expressing *cadC* under these conditions. At present, we cannot explain the difference in β-galactosidase values shown for strains E2229 and E2230 carrying pPH2249 at pH 7.6 (72 versus 4 units). The only difference in these

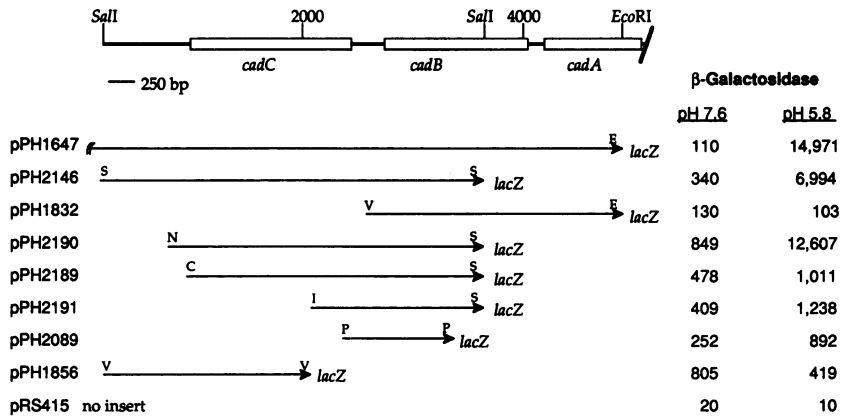

FIG. 3. Restriction fragments inserted upstream of *lacZYA* in the promoter probe plasmids pRS415 and pRS528 and the levels of β-galactosidase produced from each in strain JLS821. Cells were grown and assayed as described in Materials and Methods. The location of the endpoints of each insert are denoted by a letter corresponding to a restriction site shown in the sequence in Fig. 2. Restriction site abbreviations: C, *Csp*451; E, *Eco*RI; I, *Nsi*I; N, *Nco*I; P, *Pvu*II; S, *Sal*I; V, *Eco*RV.

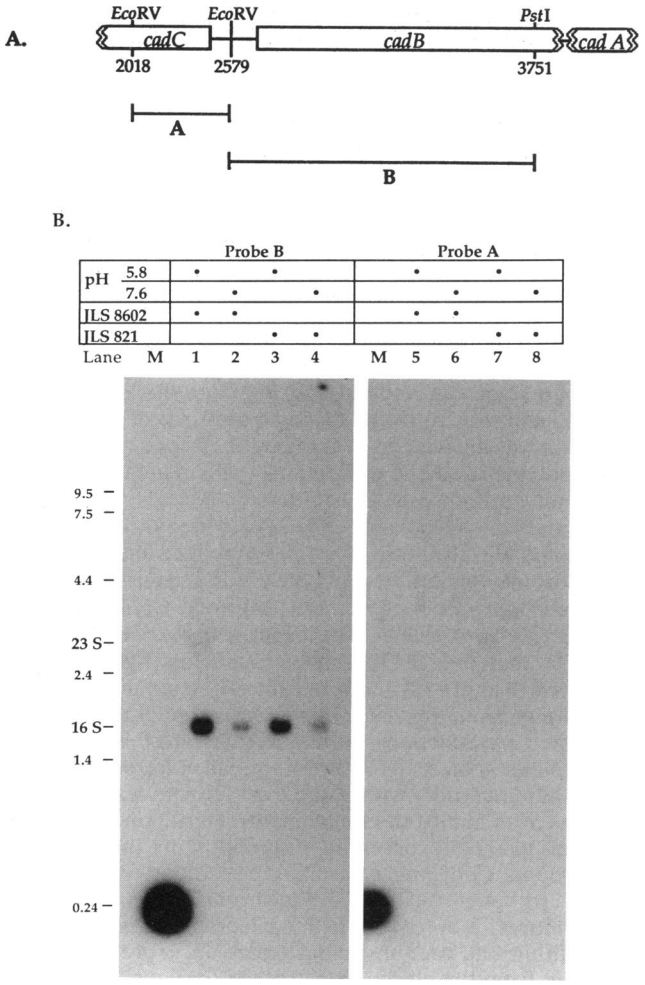

FIG. 4. Northern blot of RNA isolated from cells grown at pH 5.8 and 7.6. (A) Location of fragments cloned into pBS to generate single-stranded antisense RNA probes. (B) The matrix above the gel indicates which probe was used, at what pH the cells were grown, and source of the RNA. The lane marked M contains RNA markers used as standards. To the left of the gels are the sizes (in kilobases) and locations of the markers and 16S and 23S RNA, as determined by staining the filter following hybridization. Following electrophoresis, the filter was cut in half to generate the two filters shown.

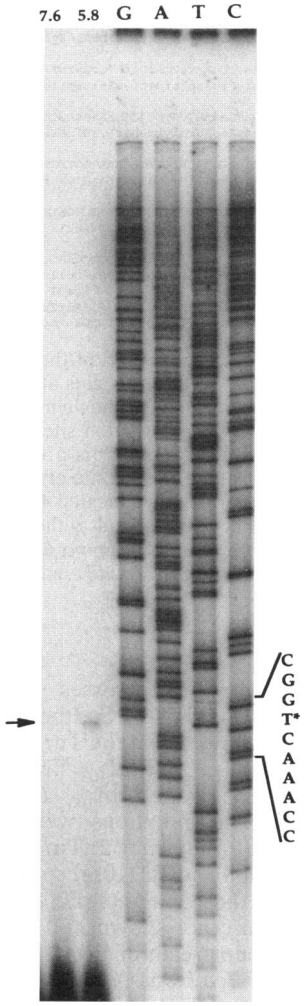

FIG. 5. Primer extension analysis. The two leftmost lanes contain RNA isolated from JLS8602 grown at pH 5.8 and 7.6 that was extended with an antisense primer corresponding to nucleotides 2758 to 2778. The right lanes contain a DNA sequencing ladder generated by using the same oligonucleotide and plasmid pPH33. The asterisk to the right of the gel indicates the endpoint of the extended product. The arrow indicates the location of the extended product.

strains is the presence of a wild-type copy of *cadC* in E2229. One explanation is that the truncated version of CadC from pPH2249 may be partially active when present with wild-type CadC. This and other hypotheses remain to be tested.

TABLE 2. Effects of −10 and −35 changes on P*cad* activity[a]

Strain	β-Galactosidase activity[b] at pH:	
	7.6	5.8
JLS821/pRS415	19	12
JLS821/pPH2089 (wild type)	312	1,041
JLS821/pPH2134 (−10 mutation)	292	220
JLS821/pPH2143 (−35 mutation)	253	189
JLS821 *cadC1*::Tn*10*/pPH2089	210	164

[a] Cells were grown and assayed as described in Materials and Methods.
[b] Reported in Miller units (22).

TABLE 3. Effects of Tn*10* mutations on P*cad* activity during aerobic and oxygen-limiting growth[a]

Strain	β-Galactosidase activity[b]			
	Aerobic		O_2-limiting	
	pH 7.6	pH 6.0	pH 7.6	pH 6.0
JLS8602	1	192	<1	1,928
JLS8602 *cadC1*::Tn*10*	3	4	2	7
JLS8602 *cadC2*::Tn*10*	3	6	1	6

[a] Cells were grown and assayed as described in Materials and Methods.
[b] Reported in Miller units.

TABLE 4. Complementation by the *cadC* plasmid, pPH2200[a]

Strain	β-Galactosidase activity[b] at pH:	
	7.6	5.8
E2229	1	605
E2229/pPH2200	151	1,963
E2229/pPH2249	72	597
E2230	2	3
E2230/pPH2200	188	1,740
E2230/pPH2249	4	5

[a] Cells were grown and assayed as described in Materials and Methods.
[b] Reported in Miller units.

Sequences upstream of P*cad* interact with a factor necessary for its activation. Plasmid pPH33 is a pBS derivative containing a 1.1-kbp *Pvu*II fragment flanking P*cad*. This fragment is in an orientation such that P*cad* and P*lac* from the vector direct transcription in opposite directions. Introduction of pPH33 into JLS8602 completely abolished expression from the chromosomal *cadA-lacZ* fusion at pH 5.8, suggesting that the plasmid is titrating out a factor necessary for P*cad* activity (Fig. 6). Additional plasmids containing portions of this 1.1-kbp region were constructed and tested for their ability to compete. Figure 6 shows that an *Rsa*I fragment containing the entire intergenic region between *cadC* and *cadB* also inhibited expression, but not to the same degree as pPH33. Splitting either the *Pvu*II or *Rsa*I fragment at the *Eco*RV site 121 bp upstream from the transcription start site, however, diminished the effect, suggesting that a site overlapping the *Eco*RV site titrates out a positive activator of P*cad*. None of these plasmids affected expression from the chromosomal *cadA-lacZ* fusion at pH 7.6 (data not shown). Although the limiting factor being titrated in these experiments has not yet been identified, one candidate would be CadC.

Computer-assisted analysis of CadC. The amino acid sequence of CadC was compared with sequences in the NBRF protein data base. A region of about 100 amino acids at the amino terminus of CadC was found to be similar to the carboxy termini of several bacterial proteins known to be transcriptional activators (31). Included in this class, among others, is PhoP from *Bacillus subtilis* (27), PhoB (15), OmpR (8) from *E. coli*, and VirG from *Agrobacterium tumefaciens*

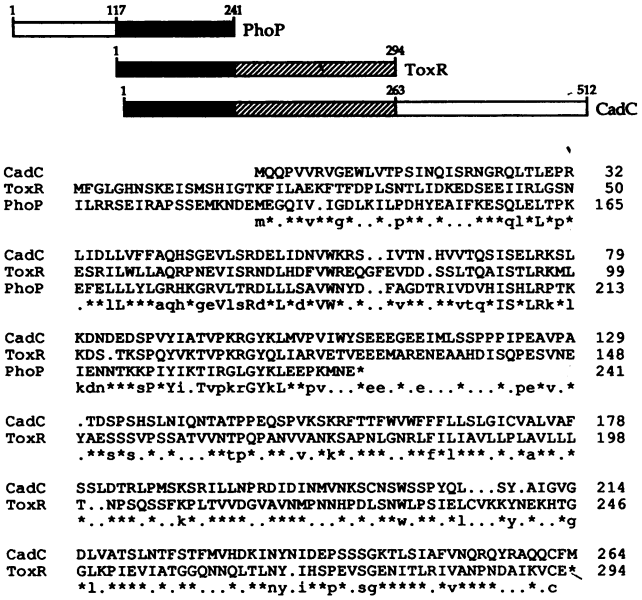

FIG. 7. Comparison between PhoP, ToxR, and CadC. Homology between the PhoP, ToxR, and CadC proteins is depicted schematically above an alignment of the primary sequences. Numbers indicate the residue number for each protein. The comparisons are summarized below the sequences as follows: (i) asterisks correspond to residues that are conserved between CadC and at least one of the other two proteins, (ii) lowercase letters correspond to residues that are identical between CadC and at least one of the other proteins, and (iii) capital letters correspond to residues that are identical in all three proteins. The sequences were aligned by performing pairwise comparisons with the GAP program of the University of Wisconsin GCG software and then making minor alterations on the basis of inspection. The carboxy ends of PhoP and ToxR are residues 241 and 294, respectively.

(20). This region of OmpR has been shown to be both sufficient and necessary for the specific interaction between OmpR and the *ompC* and *ompF* promoters (24, 32). Another protein containing this DNA binding domain is the ToxR transcriptional activator from *V. cholerae* (23). Unlike the other members of this class, however, the DNA binding

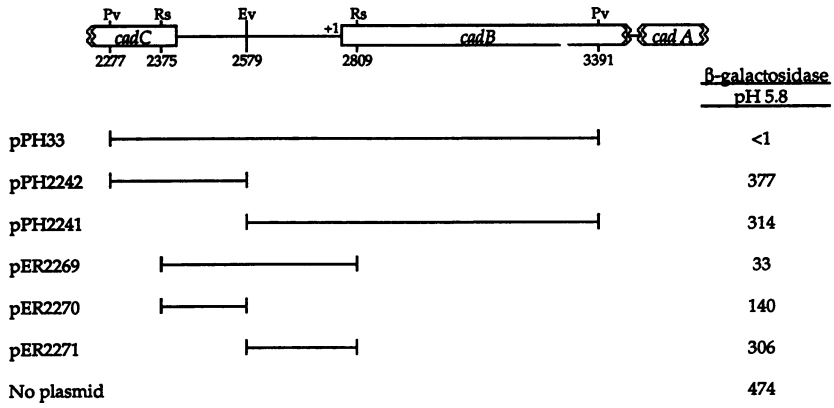

FIG. 6. Titration of a cellular factor by multicopy P*cad*-containing sequences. The solid lines represent the fragments cloned into pBS. *Pvu*II (Pv), *Eco*RV (Ev), and *Rsa*I (Rs) restriction sites are indicated. The resulting plasmids were introduced into JLS8602, and the cultures were grown and assayed for β-galactosidase as described in Materials and Methods.

domain in ToxR, like in CadC, is located at its amino terminus. A comparison between CadC, PhoP, and ToxR is shown in Fig. 7.

DISCUSSION

The Mud insertion defining the *exa-1* locus in strain JLS8602 was found to be in the gene encoding lysine decarboxylase (*cadA*). This gene has previously been shown to be regulated at the transcriptional level by external acid pH as well as by the presence of lysine and oxygen (2). Recently, the complete DNA sequence of *cadA* has been determined (3). Regulation of *cadA* expression provides a potentially useful model for understanding signalling and transcriptional mechanisms in bacteria that are influenced by external pH. The approach described in this work was to genetically and physically identify elements that influence *cadA* expression by cloning the region upstream from *cadA*, determining its DNA sequence, and examining its role in *cadA* transcription. Isolation of Tn*10* insertion mutants with altered expression of a *cadA-lacZ* transcriptional fusion were also identified. Our interpretation of the data presented here supports the following conclusions. (i) A promoter, which we have named P*cad*, is responsible for pH-regulated expression of *cadA*. (ii) There is an open reading frame situated between P*cad* and *cadA* called *cadB*. (iii) Upstream of P*cad* is an open reading frame, *cadC*, which encodes a protein necessary for pH-induced expression of P*cad*. (iv) A region upstream of the P*cad* start site interacts with a factor necessary for P*cad* induction. (v) CadC is limiting in the cell for P*cad* induction.

Plasmids pPH2189, pPH2191, and pPH2089 all contain P*cad-lacZ* fusions but only gave modest (two- to threefold) inductions of *lacZ* at pH 5.8 (Fig. 3). This observation is somewhat inconsistent with our interpretation of the Northern hybridization and primer extension data which physically located the P*cad* start site between *cadC* and *cadB*. We hypothesize that CadC levels were limiting in cells carrying a single copy of *cadC* such that multiple copies of the P*cad-lacZ* fusion were not fully induced. Cells that carried both *cadC* and P*cad* on a plasmid would be predicted to have higher levels of CadC and thus a more pronounced induction of P*cad*. The exception to this was the result with plasmid pPH2189. This plasmid contains a complete copy of *cadC*, yet the level of expression resembled those lacking *cadC*. Since the endpoint of the fragment in pPH2189 is just 18 bp upstream of the *cadC* start codon (at the *Csp*451 site), we presume that this plasmid does not express *cadC* and that an element(s) upstream of the *Csp*451 site (a promoter?) is required for *cadC* expression. If CadC expression is required for P*cad* activation, one would expect detection of a *cadC*-specific mRNA. Upon longer exposure of the Northern blots using the *Eco*RV probe shown in Fig. 4, we did observe weak hybridization; however, we do not know whether this was *cadC* specific or due to some cross-reactivity with other mRNAs.

The mechanism by which CadC functions to activate P*cad* is not clear; however the observation that its amino terminus is similar to the carboxyl DNA binding domain of a group of transcriptional activators (OmpR-like proteins) suggests that a CadC-DNA interaction is involved (24, 31, 32). The OmpR class of proteins comprise a subclass of the response elements of bacterial two-component regulatory systems responsible for the cell's ability to respond to several different environmental stimuli (31). Comprising the second component are the membrane-bound proteins which are capable of both sensing the external environment and interacting with the response element to effect a change inside the cell. The ToxR protein of *V. cholerae*, which is involved in the regulation of several genes associated with the infectious process, is also a member of this class and has been implicated in the ability of the organism to sense several features of its environment including osmolarity, pH, amino acids, and temperature (21). ToxR, however, differs from the OmpR class of response proteins in at least two fundamental ways: it is membrane bound, and like CadC, the conserved DNA binding domain is at the amino terminus (23).

The similarity in organization between CadC and ToxR leads to speculation about whether functional similarities between the two exist. Specifically, is CadC a DNA-binding protein? Does CadC, like ToxR, have both cytoplasmic and periplasmic domains? Although a direct interaction between CadC and P*cad* has not been demonstrated, the genetic evidence presented here along with the similarity between CadC and other DNA-binding proteins strongly suggests that such an interaction is responsible for pH-induced transcriptional response. The physical location of CadC has not been determined; however, computer analysis of the CadC primary structure predicts a 26-amino-acid sequence (residues 155 through 180) that resembles membrane-spanning regions of other proteins (9) (data not shown). This putative transmembrane segment aligns with the predicted transmembrane segment from ToxR (Fig. 7) (23).

The similarity between CadC and other transcriptional activators, the requirement of CadC for P*cad* function, the prediction of a transmembrane segment, and the analogy with ToxR supports the following outline of a working model as to how transcription from P*cad* is pH regulated. The basic premise of the model is that like ToxR, the amino domain of CadC is a DNA binding domain and resides in the cytoplasm with one or more domains localized in the outer face of the cytoplasmic membrane or periplasmic space. This outer domain senses, either directly or indirectly, fluctuations in periplasmic pH resulting in an unspecified alteration of CadC. Such a change would result in a signal being transmitted to the cytoplasmic DNA binding domain, which in turn would interact with either P*cad* directly or another DNA target controlling the expression of a required factor. It should be emphasized that the basic features, of this model, namely, the cellular location of CadC and the ability to bind DNA, have not yet been experimentally tested.

Implicit in this model is the notion that the activity, as opposed to the expression, of CadC is altered at different pHs. Since the promoter for *cadC* has not been rigorously identified, this hypothesis also awaits direct testing. However, two lines of evidence shed light on this issue. First, the fact that pPH2200 expressed *cadC* independently of IPTG addition suggests that there is a promoter between the $SspI_{692}$ site and the start of *cadC*. Second, a fragment containing this region was inserted upstream of *lacZ* and found to contain promoter activity that was not induced at low pH (pPH1856 [Fig. 3]). In fact, activity was slightly higher with this plasmid at pH 7.6 than at pH 5.8. Together, this suggests that transcription of *cadC* is not induced by low external pH.

The hypothesis that CadC is limiting in the cell, even at pH 5.8, was invoked to explain differences in the induced levels of P*cad-lacZ* plasmids shown in Fig. 3. Another piece of evidence supporting this notion comes from the data presented in Table 4, in which expression of *cadC* from pPH2200 was shown to increase expression from the chromosomal *cadA-lacZ* fusion. The observation that the P*cad*

promoter region, when on a multicopy plasmid (but not fused to *lacZ*), titrates out a factor necessary for expression of the *cadA-lacZ* chromosomal fusion (Fig. 6) also suggests that a positive element is limiting, although we do not yet know whether CadC is the limiting factor.

The competition experiment suggests that the *Eco*RV$_{2579}$ site is part of the site involved in recognizing a limiting factor. G. Bennett's laboratory has carried out similar experiments and obtained the same result (3). The finding that the 1.1-kbp *Pvu*II fragment competes more effectively than the 0.4-kbp *Rsa*I fragment suggests that either the context that the site is in has an effect on the ability to recognize the putative factor or that sequences beyond the *Rsa*I sites are directly involved in the interaction. We do not have additional data to support either one of these hypotheses.

The *cadB* gene and its putative gene product, CadB, have been examined by G. Bennett and colleagues. They proposed that the protein is involved in the exchange of lysine and cadaverine between the cytoplasm and external environment (3).

Induction of decarboxylase genes, including *cadA*, by low external pH is further enhanced under anaerobic conditions. For *cadA*, this could be due to either a promoter other than P*cad* being utilized during anaerobiosis or a higher level of induction of P*cad*. Although the data here does not distinguish between these two possibilities, it is clear from Table 3 that *cadC* is required for expression under oxygen-limiting conditions. In addition, we observed using indicator broth under anaerobic conditions that the *cadC1*::Tn*10* and *cadC2*::Tn*10* alleles in JLS821 resulted in a lysine decarboxylase-deficient phenotype. Physical localization of the 5′ ends of mRNA produced under anaerobic conditions is required to determine whether P*cad* is utilized under these conditions.

Our data suggests that *cadA* is cotranscribed with *cadB* and that *cadC* is transcribed separately. We suggest that the *cadB*, *cadA*, and P*cad* elements define, at least in part, an operon to be referred to as the *cad* operon. The size of the operon is unknown, since the region downstream of *cadA* has not been thoroughly analyzed. A heat shock gene coding for a lysine synthetase, *lysU*, maps downstream from *cadA* (7, 33) but whether it is also transcribed by P*cad* is not known. The isolation of *cadC1*::Tn*10* and *cadC2*::Tn*10* mutations should allow one to determine whether *lysU* expression is affected by P*cad* activity.

ACKNOWLEDGMENTS

We wish to thank the following individuals: G. Bennett for sharing unpublished data; C. Ford for technical assistance; R. Simons and C. Gross for supplying strains; D. Friedman, D. Court, and C.-S. C. Tomich for helpful discussions.

J. Slonczewski was supported by National Science Foundation grant DMB9003751.

REFERENCES

1. **Aliabadi, Z., Y. K. Park, J. L. Slonczewski, and J. W. Foster.** 1988. Novel regulatory loci controlling oxygen- and pH-regulated gene expression in *Salmonella typhimurium*. J. Bacteriol. **170**:842–851.

1a. **Amersham, Inc.** 1988. cDNA synthesis system plus, p. 20. Amersham, Inc., Arlington Heights, Ill.

2. **Auger, E. A., K. E. Redding, T. Plumb, L. C. Childs, S.-Y. Meng, and G. N. Bennett.** 1989. Construction of *lac* fusions to the inducible arginine- and lysine decarboxylase genes of *Escherichia coli* K12. Mol. Microbiol. **3**:609–620.

3. **Bennett, G. (Rice University).** 1991. Personal communication.

4. **Bingham, R. J., K. S. Hall, and J. L. Slonczewski.** 1990. Alkaline induction of a novel gene locus, *alx*, in *Escherichia coli*. J. Bacteriol. **172**:2184–2186.

5. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.

6. **Booth, I. R.** 1985. Regulation of cytoplasmic pH in bacteria. Microbiol. Rev. **49**:359–378.

7. **Clark, R. L., and F. C. Neidhardt.** 1990. Roles of the two lysyl-tRNA synthetases of *Escherichia coli*: analysis of nucleotide sequences and mutant behavior. J. Bacteriol. **172**:3237–3243.

8. **Comeau, D. E., K. Ikenaka, K. Tsung, and M. Inouye.** 1985. Primary characterization of the protein products of the *Escherichia coli ompB* locus: structure and regulation of synthesis of the OmpR and EnvZ proteins. J. Bacteriol. **164**:578–584.

9. **Engelman, D. M., T. A. Steitz, and A. Goldman.** 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. **15**:321–353.

10. **Frischauf, A., H. Lehrach, A. Poustka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170**:827–842.

11. **Gale, E. F., and H. M. R. Epps.** 1942. The effect of the pH of the medium during growth on the enzymic activities of bacteria (*Escherichia coli* and *Micrococcus lysodiekticus*) and the biological significance of the changes produced. Biochem. J. **36**: 600–619.

12. **Harley, C. B., and R. P. Reynolds.** 1987. Analysis of *E. coli* promoter sequences. Nucleic Acids Res. **15**:2343–2361.

13. **Hill, C. W., and B. W. Harnish.** 1981. Inversions between ribosomal RNA genes of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **78**:7069–7072.

14. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.

15. **Makino, K., H. Shinagawa, M. Amemura, and A. Nakata.** 1986. Nucleotide sequence of the *pho*B gene, the positive regulatory gene for the phosphate regulon of *Escherichia coli* K12. J. Mol. Biol. **190**:37–44.

16. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

17. **Marotti, K. R., and C.-S. C. Tomich.** 1989. Simple and efficient oligonucleotide-directed mutagenesis using one primer and circular plasmid DNA template. Gene Anal. Tech. **6**:67–70.

18. **Mat-Jan, F., F. Y. Alam, and D. P. Clark.** 1989. Mutants of *Escherichia coli* deficient in the fermentative lactate dehydrogenase. J. Bacteriol. **171**:342–348.

19. **Mekalanos, J. J.** 1985. Cholera toxin: genetic analysis, regulation, and role in pathogenesis. *In* W. Goebel (ed.), Genetic approaches to microbial pathogenicity. Curr. Top. Microbiol. Immunol. **118**:97–118.

20. **Melchers, L. S., D. V. Thompson, K. B. Idler, R. A. Schilperoort, and P. J. J. Hooykaas.** 1986. Nucleotide sequence of the virulence gene *virG* of the *Agrobacterium tumefaciens* octopine Ti plasmid: significant homology between *virG* and the regulatory genes *ompR*, *phoB* and *dye* of *E. coli*. Nucleic Acids Res. **14**:9933–9942.

21. **Miller, J. F., J. J. Mekalanos, and S. Falkow.** 1989. Coordinate regulation and sensory transduction in the control of bacterial virulence. Science **243**:916–922.

22. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

23. **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.

24. **Mizuno, T., M. Kato, Y. L. Jo, and S. Mizushima.** 1988. Interaction of OmpR, a positive regulator, with the osmoregulated *ompC* and *ompF* genes of *Escherichia coli*. Studies with wild-type and mutant OmpR proteins. J. Biol. Chem. **263**:1008–1012.

25. **Olson, E. R., D. S. Dunyak, L. M. Jurss, and R. A. Poorman.**

1991. Identification and characterization of *dppA*, an *Escherichia coli* gene encoding a periplasmic dipeptide transport protein. J. Bacteriol. **173**:234–244.
26. **Padan, E., D. Zilberstein, and S. Schuldiner.** 1981. pH homeostasis in bacteria. Biochim. Biophys. Acta **650**:151–166.
27. **Seki, T., H. Yoshikawa, H. Takahashi, and H. Saito.** 1987. Cloning and nucleotide sequence of *phoP*, the regulatory gene for alkaline phosphatase and phosphodiesterase in *Bacillus subtilis*. J. Bacteriol. **169**:2913–2916.
28. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53**:85–96.
30. **Slonczewski, J. L., T. N. Gonzalez, F. M. Bartholomew, and N. J. Holt.** 1987. Mu d-directed *lacZ* fusions regulated by low pH in *Escherichia coli*. J. Bacteriol. **169**:3001–3006.
31. **Stock, J. B., A. J. Ninfa, and A. M. Stock.** 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. **53**:450–490.
32. **Tsung, K., R. E. Brissette, and M. Inouye.** 1989. Identification of the DNA-binding domain of the OmpR protein required for transcriptional activation of the *ompF* and *ompC* genes of *Escherichia coli* by *in vivo* DNA footprinting. J. Biol. Chem. **264**:10104–10109.
33. **VanBogelen, R. A., V. Vaughn, and F. C. Neidhardt.** 1983. Gene for heat-inducible lysyl-tRNA synthetase (*lysU*) maps near *cadA* in *Escherichia coli*. J. Bacteriol. **153**:1066–1068.
34. **Way, J. C., M. A. Davis, D. Morisato, D. E. Roberts, and N. Kleckner.** 1984. New Tn*10* derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. Gene **32**:369–379.