JOURNAL OF BACTERIOLOGY, June 1986, p. 892–900
0021-9193/86/060892-09$02.00/0

Vol. 166, No. 3

Mutations in the *araC* Regulatory Gene of *Escherichia coli* B/r That Affect Repressor and Activator Functions of AraC Protein

LAURA G. CASS AND GARY WILCOX*

Department of Microbiology, University of California, Los Angeles, California 90024

Received 12 November 1985/Accepted 7 March 1986

Mutations in the *araC* gene of *Escherichia coli* B/r were isolated which alter both activation of the *araBAD* operon expression and autoregulation. The mutations were isolated on an *araC*-containing plasmid by hydroxylamine mutagenesis of plasmid DNA. The mutant phenotype selected was the inability to autoregulate. The DNA sequence of 16 mutants was determined and found to consist of seven different missense mutations located within the distal third of the *araC* gene. Enzyme activities revealed that each *araC* mutation had altered both autoregulatory and activator functions of AraC protein. The mutational analysis presented in this paper suggests that both autoregulatory and activator functions are localized to the same determinants of the AraC protein and that the amino acid sequence within the carboxy-terminal region of AraC protein is important for site-specific DNA binding.

The AraC protein regulates the expression of four transcriptional units involved in L-arabinose utilization in *Escherichia coli* (13, 26). In the presence of L-arabinose, AraC protein activates transcription of the *araBAD* operon, which encodes the three L-arabinose-catabolizing enzymes, and of the *araE* and *araF* genes, which encode L-arabinose transport proteins (13, 26). In the absence of L-arabinose, AraC protein represses *araBAD* transcription to a low basal level (13, 26). In vitro and in vivo analyses have demonstrated that expression of the *araC* gene, which encodes the AraC protein, is autoregulated (3, 27). In the presence or absence of L-arabinose, AraC protein represses its own transcription, resulting in a constant steady-state level of AraC protein in the cell (3, 16). Thus, there are three regulatory properties associated with AraC protein: activation of *araBAD* and the transport genes, repression of *araBAD*, and autoregulation of *araC*.

The *araC* gene is located adjacent to the *araBAD* operon but is transcribed in the opposite direction (46). The transcription initiation sites are separated by 147 base pairs (bp) (44), and the DNA sequence of this regulatory region is known (15, 39). DNA protection studies and in vivo deletion analysis of this region have defined three AraC protein-binding sites (10, 27, 36). AraC protein binds: (i) at *araI* where, in the activator conformation, it stimulates initiation of transcription of *araBAD*; (ii) at *araO1* which overlaps the RNA polymerase-binding site of the *araC* promoter preventing initiation of transcription; and (iii) at *araO2* located within the transcribed but untranslated region of the *araC* gene which is required for repression of *araBAD*. The *E. coli* *araC* gene encodes a 292-amino acid polypeptide as predicted by DNA sequence analysis (33). AraC protein from *E. coli* has been purified and shown to be a dimer consisting of two identical subunits (47, 48).

We are interested in understanding the structure-function relationships involved in the regulatory properties of AraC protein. A mutational analysis in which these properties are altered will provide a means to define the site(s) on the AraC protein involved in the three regulatory functions. Two of the regulatory functions, activation of *araBAD* and autoregulation of *araC*, are easily measured in vivo utilizing an *araC-lacZ* protein fusion strain. In this paper we describe the isolation and characterization of nonself-regulatory *araC* mutants. Each mutation affects both autoregulatory and activator properties of AraC protein. Our results suggest that these functions are localized to the same determinants on AraC protien. This is consistent with the observation that the AraC-binding sites, *araI* and *araO1*, show significant DNA sequence homology (19). The mutations isolated in this study are located in the carboxy-terminal coding region of the *araC* gene. We believe these mutations affect, either directly or indirectly, the ability of AraC protein to bind at *araI* and *araO1* and thus impair induction of *araBAD* expression and repression of *araC* expression.

MATERIALS AND METHODS

Media. The media and antibiotics used in this study have been previously described (6) with the following addition: growth medium for enzyme assays consisted of M9 (31) salts containing 0.4% (vol/vol) glycerol, 0.1% (wt/vol) casein hydrolysate, 40 μg of L-proline per ml, 40 μg of L-leucine per ml, and 10 μg of thiamine per ml.

Chemicals and enzymes. [α-^{32}P]dATP (400 Ci/mmol) was purchased from Amersham Corp. (Arlington Heights, Ill.). ^{125}I-labeled protein A (2 to 10 μCi/μg) was a product of New England Nuclear Corp. (Boston, Mass.). Trypsin was from Worthington Diagnostics (Freehold, N.J.). All restriction enzymes and other enzymes used in DNA manipulations were purchased from New England BioLabs, Inc. (Beverly, Mass.), Boehringer Mannheim Biochemicals (Indianapolis, Ind.), or Bethesda Research Laboratories, Inc. (Gaithersburg, Md.) and used as recommended by the manufacturer. Deoxynucleotide triphosphates and dideoxynucleotide triphosphates were obtained from P-L Biochemicals, Inc. (Milwaukee, Wis.). Blocked dimers and monomers were purchased from Bachemgentec.

Bacterial strains and plasmids. The bacterial strains and plasmids used or constructed in this study are listed in Table 1. Plasmid DNA was isolated by the alkaline lysis method described by Maniatis et al. (29).

Plasmid constructions. (i) pLGC500. pBR322 DNA was linearized with *Eco*RI and dephosphorylated with calf intestinal phosphatase. A 2-kilobase *Bam*HI-*Eco*RI *araA' araB*$^+$ fragment isolated from pAH105 (21) and a 3.3-kilobase

* Corresponding author.

TABLE 1. Bacterial strains and plasmids

Strain[a] or plasmid[b]	Genotype[c]	Source or reference
Bacterial strains		
UP1000	F⁻ wild type	14
LA920	*araC*::Mu d(Ap^r*lac*) Δ*leu* Δ(*lac pro*)	34
LA839	LA920 Δ(*araB popC*)*500*	This work
LA840	LA839 Φ(*araC-lacZ*)(Hyb)	This work
BD817	F⁻ *met nadB7 ung-1 hsdR sup-2 sup-3 recA56*	9
Plasmids		
pBR322	*bla*$^+$ *tet*$^+$	41
pLJS1	pSC101 [*Eco*RI::*araA' araB*$^+$ Φ(*araC-lacZ*)(Hyb)] *lacY'*	34
pLGC500	pBR322 (EcoRI::*araA'araB'lacZ'*)	This work
pLGC1000	pBR322 (*Pst*I-*Pvu*I *bla'*::LT2 *araB'* B/r *araC*$^+$)	This work
pLGC183	pLGC1000 (*araC183*)	This work
pLGC210	pLGC1000 (*araC210*)	This work
pLGC227	pLGC1000 (*araC227*)	This work
pLGC231	pLGC1000 (*araC231*)	This work
pLGC235	pLGC1000 (*araC235*)	This work
pLGC264	pLGC1000 (*araC264*)	This work
pLGC268	pLGC1000 (*araC268*)	This work

[a] All strains are derivatives of *E. coli* B/r UP1000 except BD817 which is an *E. coli* K-12 strain.

[b] Plasmids carrying *araC* mutations constructed in this work are numbered with respect to the position of the resulting amino acid substitution in the AraC protein. Strains carrying these plasmids will be referred to as mutants numbered the same as the corresponding plasmid, but with the number followed by the single-letter abbreviation of the amino acid substitution.

[c] Genetic nomenclature is as described by Bachmann (1) and Novick et al. (35). LT2, DNA from *S. typhimurium* LT2; B/r, DNA from *E. coli* B/r; all other *ara* DNA is from *E. coli* B/r.

*Bam*HI-*Eco*RI *lacZ'* fragment isolated from pMC1403 (4) were ligated with *Eco*RI-digested pBR322. The *Bam*HI site of pMC1403 is located within codon 8 of the *lacZ* structural gene. A plasmid, designated pLGC500, was obtained containing both fragments inserted into pBR322 in the orientation diagramed in Fig. 1.

(ii) pLGC1000. pBR322 was digested with *Pst*I and *Pvu*I, removing a 126-bp fragment within the *bla* gene. A 670-bp *Pst*I-*Bam*HI *araB'* fragment isolated from pMH2 (20) and a *Bam*HI-*Pvu*I *araC*$^+$ fragment isolated from pAH105 were ligated with *Pst*I-*Pvu*I-restricted pBR322. The ligation mix was used to transform the Δ(*araOC*)*719* deletion strain LA6 (15), and transformants were selected on MacConkey-L-arabinose-tetracycline (MAT) medium. Ara$^+$ tetracycline-resistant (Tet^r) transformants containing the plasmid designated pLGC1000 were obtained. Plasmid pLGC1000 contains a small portion of the *araB* gene, all of the *ara* control region, and all of the *araC* gene.

Construction of *araC-lacZ* protein fusion strain LA840. The in vivo recombination procedure used to isolate strain LA840 is shown in Fig. 1. Strain LA920, containing a stabilized *araC-lacZ* operon fusion created by Mu d(Ap^r *lac*) (5), was transformed with plasmid pLGC500. The transformants were Lac$^+$ Ara$^-$ as there was no *araC* gene product to repress the expression of the *araC-lacZ* fusion. Plasmid pLGC500 shares homology with the operon fusion within *araB* and part of *araA* and *lacZ*. It does not contain the *ara* control region, *araC*, or Mu d DNA preceding the *lacZ* gene that is present in LA920. A white Lac$^-$ homogenote in which the control site deletion on pLGC500 recombined onto the chromosome was detected on MacConkey-lactose-tetracycline (MLT) medium. Repeated growth of the cells in the absence of tetracycline resulted in Lac$^-$ Ara$^-$ Tet^s colonies owing to the loss of the plasmid. The Lac$^-$ Ara$^-$ Tet^s strain was designated LA839. Strain LA839 was transformed with plasmid pLJS1, and transformants were selected on MLT medium. Plasmid pLJS1 contains an *araC-lacZ* protein fusion in which codon 6 of the *araC* gene is fused with codon 8 of the *lacZ* gene. These transformants were white on MLT because there is no functional *lacY* gene on the plasmid. A red, Lac$^+$ homogenote in which the protein fusion recombined onto the chromosome replacing the deletion in strain LA839 was isolated and purified. The plasmid was eliminated, and the resulting Lac$^+$ Ara$^-$ strain was designated LA840. AraC protein produced by an *araC*$^+$ plasmid in *trans* to the protein fusion activates *araBAD* expression and represses the *araC-lacZ* fusion. To confirm the presence of the protein fusion in strain LA840, a Southern blot hybridization analysis of LA920 and LA840 chromosomal DNA was performed, using as a probe a DNA fragment spanning the *araC-lacZ* protein fusion. The probe hybridized to an LA840 chromosomal fragment of the same size as the probe but to a larger LA920 chromosomal fragment containing *araC* and Mu d DNA (data not shown).

Mutant isolation. pLGC1000 DNA (2.5 μg) was incubated in 250 μl of 1 M hydroxylamine (HA)–0.1 M Tris hydrochloride (pH 6.0)–1 mM EDTA at 70°C for 30 min. The DNA was then precipitated with ethanol, suspended in 10 mM Tris hydrochloride (pH 7.5)–1 mM EDTA–2% acetone, and used to transform strain BD817. This *E. coli* strain is deficient in uracil-DNA glycosylase (*ung*) which functions in the base excision repair of premutational lesions resulting from deamination of cytosine (9). HA-mutagenized DNA cycled through strain BD817 should contain an increased frequency of transition mutations. Transformants were selected on tryptone-yeast extract-tetracycline agar, the plates were scraped of growth, and plasmid DNA was isolated. This DNA was used to transform strain LA840, and transformants were selected on MLT medium. Transformants harboring a plasmid that contains the wild-type *araC* gene were Lac$^-$ because the AraC protein expressed by the plasmid represses the expression of the chromosomal *araC-lacZ* protein fusion. Conversely, transformants containing a plasmid with a mutation in the *araC* gene were detected as Lac$^+$ colonies that appeared after overnight incubation at 37°C. Lac$^+$ isolates were purified, and plasmid DNA was isolated and used to retransform strain LA840. The Lac$^+$ phenotype of the transformants confirmed that the mutation had occurred on the plasmid. Plasmids containing the putative *araC* mutations were purified for further characterization.

Enzyme assays. Cultures (5 ml) of strains LA920 or UP1000 containing the desired *araC* plasmid were grown in supplemented M9 salts medium containing tetracycline (15 μg/ml) at 37°C for 16 h. These cells were used to inoculate 50 ml of the same medium, and growth was monitored with a Klett-Summerson colorimeter (filter no. 42). Cultures were inoculated to 10 units (3.3×10^7 cells per ml), and at 70 units (2.3×10^8 cells per ml) L-arabinose was added to a final concentration of 0.4% (wt/vol) where indicated. Growth was continued for one doubling of LA920 cultures or to 115 units (3.8×10^8 cells per ml) for UP1000 cultures, at which time 50 μg of chloramphenicol per ml was added, and the samples were placed on ice. Cell extracts were prepared, and L-arabinose isomerase was assayed as previously described (12). β-Galactosidase activity was measured as described by Miller (31).

DNA sequence analysis. The dideoxy method was used for DNA sequence determination (38). A series of *araC*-specific oligodeoxyribonucleotide primers were designed. Four

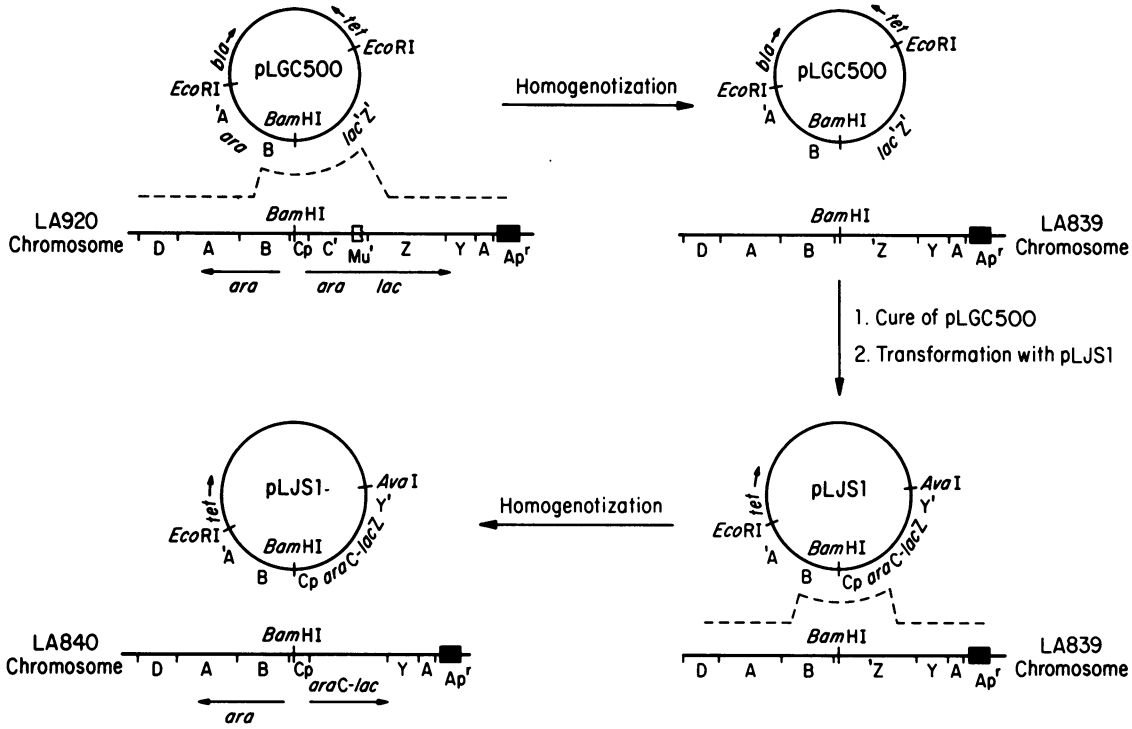

FIG. 1. Construction of *araC-lacZ* protein fusion strain LA840 as described in Materials and Methods.

primers were manually synthesized on a solid support by the phosphotriester method (24), and five were synthesized with an Applied Biosystems model 380A DNA synthesizer instrument. Seven of the primers spaced approximately 160 bp apart are complementary to the coding strand of *araC*. These primers allow sequencing from within the mRNA leader region through the *araC* coding region to a region beyond the carboxy terminus. Two primers are identical to the coding strand and correspond to regions in the carboxy-terminal third of the *araC* gene, allowing sequencing of both DNA strands in this region of the gene. CsCl-purified plasmid DNA was sequenced directly by using the *araC*-specific primers by the method of Wallace et al. (43).

Immunoblot analysis. Crude AraC antiserum was prepared from New Zealand White rabbits after injection of purified *Salmonella typhimurium* AraC protein. *S. typhimurium* AraC protein (>90% pure) and *E. coli* B/r AraC protein (>80% pure) were gifts from P. Clarke (this laboratory) and R. Linzmeier (this laboratory), respectively. Growth conditions and preparation of cell extracts were those used for enzyme assays with the following modification: cell extract buffer (buffer B) consisted of 0.15 M potassium acetate, 0.01 M Tris acetate (pH 7.4), 1 mM EDTA, 0.4% (wt/vol) L-arabinose, and 50 μg of phenylmethylsulfonyl fluoride per ml. Purified protein or cell extracts were separated by sodium dodecyl sulfate-acrylamide gel electrophoresis followed by electrotransfer to nitrocellulose (42). The portion of nitrocellulose containing the molecular weight standards was removed and stained with Coomassie blue. The nitrocellulose strip was soaked in 0.01 M Tris hydrochloride (pH 8.0)–2 mM EDTA–0.05 M NaCl for 15 min, stained in 25% isopropanol–10% acetic acid–0.04% Coomassie blue for 15 s, and destained in several changes of 25% isopropanol–10% acetic acid (Ken Yamaguci, M.A. thesis, University of California, Los Angeles, 1985). Immunological detection of AraC-specific proteins was carried out as previously described (42). Scanning densitometry of the resulting autoradiogram yielded the relative quantity of protein in each band.

Gel electrophoresis-DNA binding assay. A 167-bp *Bam*HI-*Bst*EII fragment containing both the *araI* and *araO1* sites was 3' end labeled by incubation with *E. coli* DNA polymerase I Klenow fragment and [α-^{32}P]dATP (29). The DNA binding assays were performed as described by Brunelle et al. (2) with the following modifications: binding buffer contained 150 mM KCl and 1 mg of sonicated salmon sperm DNA per ml. A constant amount of ^{32}P-labeled DNA fragment (500 pg) was incubated with a specified amount of cell extract containing AraC protein. Cell extracts were those used for immunoblot analysis. The protein-DNA complexes were separated from free DNA by gel electrophoresis as described previously (2).

Trypsin digestion. Purified *E. coli* B/r AraC protein or cell extract was incubated with 10% (wt/wt protein) trypsin in 10 mM Tris hydrochloride (pH 7.4)–1 mM dithiothreitol–5% glycerol at room temperature. Samples were removed at various time points, and digestion was stopped by the addition of an equal volume of 40% trichloroacetic acid. After 20 min at room temperature, the suspension was centrifuged. The resulting protein pellet was washed with 200 μl of acetone, air dried, and suspended in buffer B.

RESULTS

Isolation of *araC* nonself-regulatory mutants. The *araC-lacZ* protein fusion strain LA840 (Fig. 1) was used in the isolation of *araC* mutants that are unable to autoregulate. β-Galactosidase is synthesized in LA840 at a constitutive level under the control of the *araC* promoter, and the strain is Lac$^+$ on MLT medium. As there is no intact *araC* gene on the fusion chromosome there is neither repression of *araC-lacZ* nor activation of *araBAD* expression. These *araC* functions are complemented when strain LA840 is transformed with the *araC*$^+$ plasmid pLGC1000. The AraC pro-

TABLE 2. Sequence changes in AraC mutants

Plasmid	Codon change	Amino acid location[a]	Amino acid substitution
pLGC183	TGT→TAT	183	Cys→Tyr
pLGC210	CGT→TGT	210	Arg→Cys
pLGC227	CGC→CAC	227	Arg→His
pLGC231	CGT→CAT	231	Arg→His
pLGC235	GCG→ACG	235	Ala→Thr
pLGC264	GTA→ATA	264	Val→Ile
pLGC268	TGC→TAC	268	Cys→Tyr

[a] Amino acid numbering is that used previously (33) with Met as amino acid 1.

tein represses the *araC-lacZ* fusion, resulting in a Lac$^-$ phenotype on MLT. Strain LA840 was transformed with HA-mutagenized pLGC1000, and nonself-regulated mutants were detected as Lac$^+$ colonies on MLT medium. Purified Lac$^+$ colonies were tested for the ability to utilize L-arabinose on MAT medium. To eliminate mutants that were a result of a nonsense mutation, deletion, or inversion in the *araC* gene, only Lac$^+$ colonies having at least a weak Ara$^+$ phenotype were studied further. Under the HA mutagenesis conditions used, Lac$^+$ colonies appeared at an approximate frequency of 1 in 300. A total of 50% of the Lac$^+$ colonies had a detectable Ara$^+$ phenotype on MAT medium. Purified plasmids from 16 of these colonies were further characterized. All of the Lac$^+$ colonies resulted from a mutation occurring on the plasmid as shown by retransformation of strain LA840.

DNA sequence analysis of *araC* mutations. DNA sequence analysis of plasmids containing the *araC* mutations was performed with *araC*-specific synthetic primers. The DNA sequence of the *araC* gene in 16 *araC* mutants was determined, revealing seven unique point mutations. The results are summarized in Table 2. Each mutation resulted from a G/C-to-A/T transition. The DNA sequence of at least one strand of the entire *araC* gene was determined to confirm the presence of a single point mutation in *araC*. The seven mutations are located within the region coding for amino acids 183 to 268 of AraC protein. Three of the mutants isolated contained the identical mutation at nucleotide position 693, the most internal mutation, and resulted in an amino acid substitution at position 183 (Cys→Tyr). Eight identical mutations were found at nucleotide position 837 and generated an Arg→His substitution at amino acid 231. These multiple mutants were isolated from a single transformation of strain LA840 with HA-treated pLGC1000 DNA. It is possible that they are siblings as the plasmid DNA was first cycled through strain BD817 before transformation of strain LA840.

Characterization of *araC* mutations in vivo. Each *araC* mutation was characterized with respect to its effect on the regulatory properties of AraC protein. The ability of the plasmid-encoded AraC protein to activate *araBAD* expression and to repress its own gene transcription was measured in the *araC-lacZ* operon fusion strain LA920. The levels of L-arabinose isomerase, the product of the *araA* gene, and of β-galactosidase were assayed. Strain LA920 rather than strain LA840 was used to characterize the *araC* mutants because the unrepressed level of β-galactosidase synthesized from the operon fusion was higher than that from the protein fusion. Small changes in β-galactosidase levels would thus be more significant. Table 3 shows the results of these assays. Each mutation in *araC* affects the autoregulatory function, the characteristic that was selected in the

TABLE 3. Effect of *araC* mutations on activator and autorepressor function of AraC protein[a]

Plasmid	Autoregulatory function		Activator function	
	β-Galactosidase sp act[b]	% Wild-type activity	L-Arabinose isomerase sp act[c]	% Wild-type activity
pLGC1000	152 ± 5	100	36.6 ± 3.2	100
pLGC264	791 ± 36	70	27.7 ± 3.0	76
pLGC268	863 ± 85	67	28.7 ± 3.5	78
pLGC235	1,663 ± 103	30	12.1 ± 3.4	33
pLGC227	2,018 ± 164	14	18.0 ± 1.2	49
pLGC210	2,079 ± 127	11	2.4 ± 1.9	7
pLGC183	2,091 ± 160	10	18.9 ± 4.3	52
pLGC231	2,260 ± 68	2	0.7 ± 0.3	2
PBR322	2,314 ± 130	0	<0.2	0

[a] Strain LA920 was transformed with the *araC*-containing plasmids, and enzyme levels were measured in crude extracts. Results expressed as mean ± average error of four experiments.

[b] β-Galactosidase activity was measured in cells grown in the absence of L-arabinose. β-Galactosidase specific activity is expressed as nanomoles of *O*-nitrophenol formed per minute per milligram of protein at 28°C.

[c] L-Arabinose isomerase activity was measured in cells induced with 0.4% L-arabinose. L-Arabinose isomerase specific activity is expressed as micromoles of L-ribulose formed per minuter per milligram of protein at 30°C.

mutant isolation, and in addition affects activator function. The mutants retain various amounts of autoregulatory and activator function, autoregulation ranging from 2 to 70% of that of *araC*$^+$ cells and activation ranging from 2 to 76% of that of *araC*$^+$ cells. Five of the seven *araC* mutations affect the ability to activate and to repress to a similar extent. This parallel loss of regulatory functions is shown graphically in Fig. 2. The *araC* mutations on plasmids pLGC227 and pLGC183 affect autoregulation to a greater extent than activation, both retaining only approximately 10% of wild-type autoregulatory function but 50% activator function.

An *araC* mutation which alters a site in the AraC protein necessary for subunit association might result in a loss of ability to form AraC dimers. To test this possibility, we examined the *araC* mutants for the ability to produce *trans*-dominant polypeptides in the presence of an *araC*$^+$ allele. If

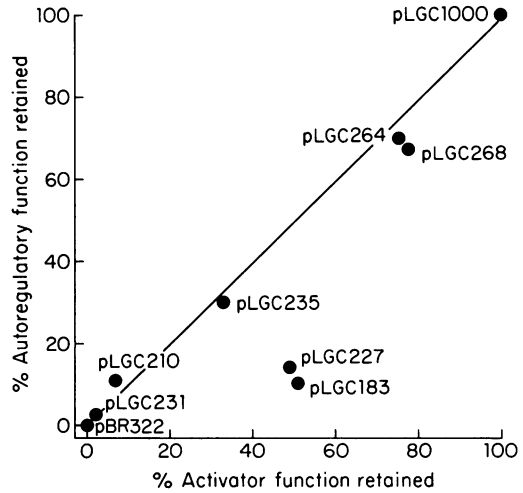

FIG. 2. Comparison of the amount of wild-type AraC activator and autoregulatory function retained in each *araC* mutant. Data are taken from Table 3.

the mutant AraC monomers are able to form dimers, a mutant AraC polypeptide produced on a multicopy plasmid might reduce the activity of a chromosome-encoded wild-type AraC protein. If unable to function, the resulting mixed dimers consisting of wild-type and mutant monomers and the reduced number of wild-type dimers would result in a decrease in *araBAD* expression. A mutant AraC monomer that is impaired in its ability to form dimers would not decrease the level of wild-type dimers, and thus the level of *araBAD* expression would remain the same. Mutants that have a *trans*-dominant phenotype have been isolated in the *lacI* (32), *trpR* (25), and *tetR* (23) genes. To determine whether the *araC* mutations have a *trans*-dominant effect on *araBAD* expression, we measured the level of L-arabinose isomerase in the wild-type *araC*$^+$ strain UP1000 transformed with the mutated *araC* plasmids (Table 4). All of the *araC* mutations resulted in a decreased level of activation by the *araC*$^+$ chromosomal allele. The level of activation ranged from 30 to 70% of the haploid UP1000 activity. The mutants 210C and 231H which showed the most severe negative complementation also had the most severe impairment of activator function. Since the mutants range in their impairment of activator function (Table 3), the level of nonfunctional mixed dimers in the presence of the *araC*$^+$ allele should also vary. The *trans* dominance of all of the mutants suggest that the mutated AraC polypeptides are able to efficiently associate to form dimers.

Mutations in *araC* designated *araC*^{c} which result in higher levels of *araBAD* expression in the absence of L-arabinose have previously been isolated and mapped (11). The *araC*^{c} mutants were isolated by their resistance to D-fucose, an analog of L-arabinose. The postulated phenotype is an altered *araC* gene product which, at least partially, assumes an activator conformation in the absence of the inducer L-arabinose interaction. The mutations have been mapped (11) and correspond to the 5′ end of *araC* (46) or the amino-terminal half of the AraC protein; nine of these mutations cluster within the coding region for the amino acids 6 to 21 (R. G. Wallace, Ph.D thesis, University of California, Santa Barbara, 1982). The carboxy-terminal location of the *araC* mutations isolated in the present study indicates that they may not affect the inducer-binding site of AraC protein. However, experiments by Casadaban (3) have shown that 20% of the *araC*^{c} mutants examined are unable to autoregulate. Therefore, the nonself-regulatory mutants isolated in this study were tested for an AraC^c phenotype. Strain LA920 cells transformed with the plasmids carrying the *araC* mutations all showed a D-fucose sensitivity when plated on MacConkey-arabinose-fucose medium. The uninduced levels of L-arabinose isomerase were all less than that of the *araC*$^+$ cell (pLGC1000) showing no constitutive expression (data not shown). These results indicate that the *araC* mutants isolated in this study do not have an AraC^c phenotype.

TABLE 4. *Trans* dominance of plasmid-encoded *araC* mutations as measured by activation of *araBAD* in an *araC*$^+$ strain[a]

Plasmid	L-Arabinose isomerase sp act	% Wild-type haploid activity
pBR322	29.3 ± 4.1	100
pLGC1000	29.9 ± 5.0	102
pLGC264	19.9 ± 4.0	68
pLGC268	16.7 ± 2.2	57
pLGC235	13.3 ± 3.2	45
pLGC227	18.5 ± 1.4	63
pLGC210	10.0 ± 3.1	34
pLGC183	14.7 ± 3.4	54
pLGC231	10.6 ± 3.0	47

[a] L-Arabinose isomerase was measured in cell extracts of strain UP1000 containing the appropriate plasmid. Cultures were induced with 0.4% L-arabinose. Enzyme activity is as defined in Table 3, footnote *c*. Results are expressed as mean ± average error of three experiments.

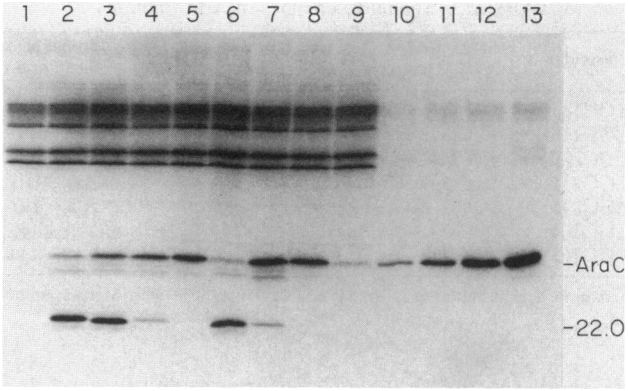

FIG. 3. Autoradiograph of immunoblot analysis of wild-type and mutant AraC protein. Lanes 1 to 9 contain cell extracts prepared from uninduced cultures of strain LA920 containing the desired *araC* plasmid. A 100-μg sample of total protein was loaded in each lane. Lanes: 1, pBR322; 2, pLGC231; 3, pLGC183; 4, pLGC227; 5, pLGC264; 6, pLGC235; 7, pLGC210; 8, pLGC268; 9, pLGC1000. Lanes 10 to 13, A dilution series of purified *E. coli* B/r AraC protein: 10, 200 ng; 11, 500 ng; 12, 1 μg; 13, 2 μg. Protein size is in kilodaltons.

Immunoblot analysis of AraC mutants. The AraC phenotype selected in this analysis was the inability to repress expression of the *araC-lacZ* fusion strain LA840, resulting in increased levels of β-galactosidase. The altered AraC protein, if unable to bind to the DNA-binding site *araO1*, would not be able to repress its own synthesis on the *araC* plasmid. This would result in an increase in the intracellular AraC protein concentration over that of the *araC*$^+$ cell. Alternatively or in addition, an increase in *araC-lacZ* expression could be due to an altered AraC protein that is unstable and is broken down at a higher rate than the wild-type AraC protein.

The intracellular level of the mutant AraC proteins was examined by immunoblot analysis (Fig. 3). A dilution series of purified *E. coli* B/r AraC protein is shown. The intensity of the purified AraC protein band in each lane was linear with respect to the amount of AraC protein loaded, indicating that the immunoblot is quantitative in this range of protein concentrations. All of the mutants had an AraC cross-reactive band which was the same size as the wild-type AraC protein. In addition, five of the seven mutants had an AraC cross-reactive breakdown product with an apparent molecular weight of 22,000. Other minor breakdown products were also present. In mutants 183Y, 231H, and 235T, the 22,000-molecular-weight product was the major cross-reactive band. The relative amounts of intact AraC protein are shown in Table 5. Six of the seven *araC* mutations resulted in an overproduction of intact AraC protein ranging from 3.7- to 13-fold. Only mutation 235T resulted in a level of AraC protein less than the wild-type level. Since more intact AraC protein is present in six of the seven mutants, it is unlikely that the decreased regulatory activity was caused by instability of the mutant AraC protein. If we assume that the intact AraC protein is active, then it is more likely that these

TABLE 5. Relative amount and DNA-binding activity of mutant AraC proteins in crude extracts

Plasmid	Relative amt intact AraC protein[a]	Relative DNA-binding activity (% wild-type activity)[b]
pLGC1000	1.0	100
pLGC235	0.7	4
pLGC231	3.7	<1
pLGC183	6.7	1
pLGC227	8.4	<1
pLGC268	11.2	5
pLGC264	11.6	5
pLGC210	13.1	<1
pBR322	0	0

[a] Scanning densitometry of the autoradiogram in Fig. 3 was used to determine areas of protein bands. Two non-AraC-specific high-molecular-weight protein bands present in all extracts were normalized for equivalent amounts of protein loaded and then used to correct the amount of AraC-specific protein. The densitometric tracing of a nonspecific band present in the pBR322 extract shown in Fig. 3 overlaps in position with the tracing of the intact AraC band present in other extracts. It was therefore necessary to subtract the area of this pBR322 band from that of the intact AraC band.

[b] Scanning densitometry of autoradiograms such as the example shown in Fig. 4 was used to determine the relative amount of protein bound to ^{32}P-labeled DNA. The dilution series of wild-type extract was used to generate a standard curve.

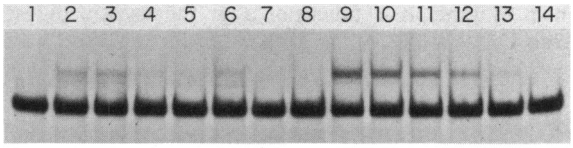

FIG. 4. Autoradiograph of gel electrophoresis-DNA binding assay of wild-type and mutant AraC protein. Cell extracts were prepared from uninduced cultures of strain LA920 containing the desired *araC* plasmid. The direction of electrophoresis is from top to bottom. Lanes 1 to 9, A constant amount of ^{32}P-labeled DNA fragment (500 pg) was incubated with equal amounts of AraC protein based on the relative amount of mutant AraC protein in each extract (Table 5). The amount of mutant extract per binding assay that yielded an equal amount of AraC protein was relative to 20 μg of wild-type extract (pLGC1000). When necessary, extract containing no AraC protein (pBR322) was added to the binding assay to give a constant amount of total protein (26 μg). Amount of extract (micrograms of protein) (*araC* plasmid): lane 1, 26 μg (pBR322); lane 2, 1.78 μg (pLGC268); lane 3, 1.72 μg (pLGC264); lane 4, 2.98 μg (pLGC183); lane 5, 2.38 μg (pLGC227); lane 6, 26 μg (pLGC235); lane 7, 1.53 μg (pLGC210); lane 8, 5.4 μg (pLGC231); lane 9, 20 μg (pLGC1000). Lanes 10 to 14, Decreasing amounts of wild-type extract (pLGC1000): 10, 10 μg; 11, 5 μg; 12, 1 μg; 13, 0.2 μg; 14, no extract.

mutant AraC proteins have less affinity for *araO1* DNA than wild-type AraC protein.

DNA-binding activity of mutant AraC proteins. A simple and quantitative gel electrophoresis-DNA binding assay for AraC protein has been developed (2, 18). Using this assay, we determined the relative DNA-binding affinities of wild-type and mutant AraC protein in crude extracts. Since each *araC* mutation affects both autoregulatory and activator functions of AraC protein, a DNA fragment containing both *araO1* and *araI* was used in the binding assay. The relative binding affinities are summarized in Table 5, and a specific example is shown in Fig. 4. Three of the *araC* mutations, 231H, 227H, and 210C, reduced the affinity of AraC protein for *ara* DNA by more than 100-fold. The other four mutations resulted in DNA-binding affinities ranging from 1 to 5% of that of wild-type AraC protein. These results directly indicate that the seven *araC* mutations have altered the ability of AraC protein to bind to the DNA sites *araO1* and *araI*.

Immunoblot analysis of AraC protein trypsin digestion products. Extensive trypsin digestion of purified AraC protein from *S. typhimurium* results in a trypsin-resistant core protein. Trypsin-generated core protein has been purified and shown to be a dimer of the amino-terminal half of AraC protein extending to Arg-178 or Arg-180 (P. Clarke and G. Wilcox, personal communication). Thus, the major sites necessary for subunit association are located within this core protein. Since the predicted amino acid sequences of *E. coli* and *S. typhimurium* AraC protein are 92% homologous and the proteins are functionally interchangeable (7), then they probably have a similar tertiary structure. A similar structural conformation of *E. coli* AraC protein should result in a similar susceptibility to proteolytic cleavage. *E. coli* AraC protein was trypsin digested, and the products were subjected to immunoblot analysis. Extensive trypsin digestion of purified *E. coli* AraC protein produced a trypsin-resistant protein having an apparent molecular weight of 21,000 (Fig. 5). The mobility of the *S. typhimurium* tryptic core product is the same as that of the *E. coli* trypsin-resistant core in Fig. 5. The *araC* mutations isolated in this study generate amino acid substitutions in AraC protein located downstream from the tryptic core terminus, Arg-180. The presence of a trypsin-generated core similar in size to that in *S. typhimurium* suggests that the *araC* mutations do not directly affect the functional site necessary for dimer formation.

An amino acid substitution might affect the structure of AraC protein such that the susceptibility to trypsin and the trypsin-generated core would be altered. Mutation 183Y is the most internal mutation within *araC* generating an amino acid substitution at position 183, very close to the *S. typhimurium* trypsin-resistant core terminus. A cell extract of mutant 183Y was digested with trypsin (Fig. 5). Intact AraC protein and the 22,000-molecular-weight breakdown product were further digested by trypsin to the tryptic core protein of 21,000 molecular weight. The amino acid substi-

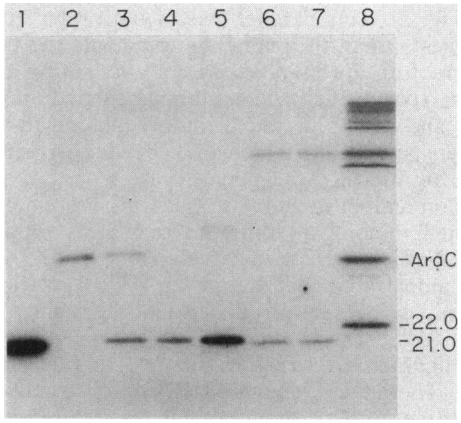

FIG. 5. Autoradiograph of immunoblot analysis of AraC protein trypsin digestion products. Lane 1 250 ng of purified trypsin-resistant core protein from *S. typhimurium*. Lanes 2 to 5, Purified *E. coli* B/r protein incubated with trypsin for different time points as described in Materials and Methods. A 666-ng sample was loaded in lanes 2 to 4, and 1 μg was loaded in lane 5. Lane 2, 0 min; lane 3, 10 min; lane 4, 60 min; lane 5, 16 h. Lanes 6 to 8, Cell extract of strain LA920 containing plasmid pLGC183 incubated with trypsin. A 98-μg sample was loaded in each lane. Lane 6, 16 h; lane 7, 60 min; lane 8, 0 min. Protein sizes are in kilodaltons.

tution in mutant 183Y does not alter the trypsin digestion pattern.

DISCUSSION

The AraC protein is a complex regulator, binding to different regulatory sites of the *ara* DNA to perform its different functions involved in L-arabinose utilization. We isolated and characterized mutations in the *araC* gene of *E. coli* B/r. The mutant phenotype we selected was for the inability to autoregulate. An *araC-lacZ* fusion strain was transformed with HA-mutagenized *araC* plasmid DNA, and transformants were screened for increased *lacZ* expression indicating a loss of autoregulation of the *araC* promoter. The *araC* mutants were characterized in vivo with respect to the level of autoregulation of *araC-lacZ* fusion expression and to the level of activation of *araBAD* operon expression. Enzyme activities revealed that each *araC* mutation had altered both the autoregulatory and activator functions of AraC protein (Table 3). This was not necessarily a predicted phenotype since only loss of autoregulation was selected. In the isolation of the nonself-regulatory mutants, only mutants which showed at least a very weak Ara^+ phenotype were further characterized to eliminate deletions, inversions, or nonsense mutations. The decreased regulatory activity of six of the seven different mutants was not simply due to a reduced amount of intact AraC protein caused by rapid turnover of the altered protein structure. These mutations cause an overproduction of intact AraC protein (Fig. 3; Table 5). We isolated one mutation, on pLGC235, which resulted in an intracellular concentration of intact AraC protein corresponding to 70% of that of wild type. However, the total amount of AraC cross-reactive material in this mutant was greater than the amount of AraC cross-reactive material in the wild-type strain. The *trans*-dominant phenotype of all of the mutations (Table 4), and their location in AraC protein outside of the proposed dimer-forming domain is consistent with the hypothesis that these *araC* mutations do not substantially affect the sites necessary for subunit interaction. Five of the seven mutations caused a linear loss of both autoregulatory and activator properties of AraC protein (Fig. 2). This parallel loss of both regulatory functions suggests that the same determinants of the AraC protein structure are involved in both activation and autoregulation. Two of the *araC* mutations carried on plasmids pLGC183 and pLGC227 impair autoregulation to a greater extent than activation, which may be due in part to our screening for a weak Ara^+ phenotype.

Activation and autoregulation by AraC protein are mediated through binding to different regulatory sites on the DNA, *araI* and *araO1*, respectively. In vitro DNA binding assays demonstrated that the seven *araC* mutations each resulted in a significantly reduced binding affinity of AraC protein for these DNA sites (Fig. 4; Table 5). The relative DNA-binding activities ranged from <1 to 5% of the wild-type level. There are three regions of conserved bases within the respective AraC protein-binding sites *araI* and *araO1* (19). A single domain on the AraC protein involved in both of the site-specific bindings is consistent with the DNA sequence homology, assuming there is no duplication of a given domain in the protein. The five mutations which result in a parallel loss of function could affect either directly or indirectly similar contacts between the protein and DNA. The DNA sequences of *araI* and *araO1* are not identical, allowing for some unique contacts between AraC protein and the DNA sites. Mutations 183Y and 227H may reflect these differences, affecting binding at *araO1* more severely than at *araI*.

The seven different missense mutations isolated in this study were located within the approximate distal third of the *araC* coding region (Table 2). One other mutation in this same region has been described which results in a mutant AraC protein (Asp to Asn at position 228) with a two- to threefold-weaker affinity for *araI* (2). A comparison of the *araC* gene from *E. coli* and two other enteric bacteria, *S. typhimurium* (7) and *Erwinia carotovora* (28), reveals a high degree of sequence homology within the carboxy region of the *araC* gene (28). The mutations isolated in this study code for substitutions in conserved amino acids of all three organisms, with one exception. Mutation 268Y generates a substitution of an amino acid conserved in two of three organisms. The different amino acid is in *Erwinia carotovora* which shows the least overall sequence homology of the AraC protein. Thus, a variety of amino acid substitutions within the highly conserved carboxy region of AraC protein lead to a lack of autoregulatory and activator function, underlying the importance of this region of AraC protein in regulatory events.

We believe the mutations isolated in this study affect either directly or indirectly the site-specific binding of AraC protein to the DNA. Several other regulatory proteins, *E. coli* cyclic AMP receptor protein, λ CI, and λ Cro, all contain a conserved amino acid sequence that forms a characteristic three-dimensional structure responsible for DNA binding (30, 37, 40, 45). These proteins bind DNA as a dimer and use α-helices to contact two adjacent major grooves along one face of the double helix. Recently, the carobxy terminus of AraC protein has been aligned with the α-helix E-turn-α-helix F region within the DNA-binding domain of cyclic AMP receptor protein (17). The pattern of identical or conserved amino acids suggests that this region forms similar protein structures.

The amino acid substitutions of mutants 268Y and 264I are located in the region aligned with the α-helix (F) of cyclic AMP receptor protein which is proposed to make direct contacts with the major groove of DNA (45). These two mutants do not alter AraC protein susceptibility to proteases (Fig. 3), suggesting that they do not cause a large change in protein conformation. Mutant 264I introduces a Val-264→Ile substitution, increasing the hydrophobicity and size of the side chain. This residue in cyclic AMP receptor protein is also hydrophobic (Ile) and is believed to be partially buried facing the hydrophobic core of the protein (45). These mutations may introduce unfavorable steric contacts.

Mutants 235T, 231H, and 183Y all introduce a larger amino acid side chain and show an increased susceptibility to proteases. These mutations lie outside of the alignment with α-helices E and F (17). These mutations most likely affect DNA binding indirectly, altering the protein conformation such that the DNA-binding helices are no longer in the proper alignment with the DNA.

A dimer of AraC protein has been proposed to contact three adjacent major groove regions of the *araI* DNA site (19). The larger DNA site requires that if the helix-turn-helix motif is involved in DNA binding by AraC protein, then some additional structure must be involved in AraC-DNA contacts. Determination of the three-dimensional structure of AraC protein is necessary before the mechanism of DNA binding is precisely defined. The mutational analysis presented in this paper indicates, however, that the amino acid sequence within the carboxy region of AraC protein is important for site-specific DNA binding.

The Western blot analysis shown in Fig. 3 gave us a means to quantitate the minimum number of AraC monomers inside the *E. coli* cell. Assuming a dry weight of 2.84×10^{-13} g per cell (22), a plasmid copy number of 18 (8), and 2.3 chromosomes per cell, the number of AraC monomers per *araC* gene is estimated to be 24. This estimation is in close agreement with the previously reported estimations of 100 monomers per cell (3) and 40 monomers per cell (G. Wilcox, Ph.D thesis). Assuming the same value of 2.3 chromosomes per cell, these previous numbers would yield 43 monomers per *araC* gene and 17 monomers per *araC* gene, respectively.

ACKNOWLEDGMENTS

We would like to thank Lori Stoltzfus and Laurel Heffernan for reading the manuscript. We thank P. Clarke for discussing with us unpublished data.

This work was supported by Public Health Service grant GM30896 from the National Institutes of Health.

LITERATURE CITED

1. **Bachmann, B. J.** 1983. Linkage map of *Escherichia coli* K-12, edition 7. Microbiol. Rev. **47**:180–230.
2. **Brunelle, A., W. Hendrickson, and R. Schleif.** 1985. Altered DNA contacts made by a mutant AraC protein. Nucleic Acids Res. **13**:5019–5026.
3. **Casadaban, M. J.** 1976. Regulation of the regulatory gene for the arabinose pathway, *araC*. J. Mol. Biol. **104**:557–566.
4. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143**:971–980.
5. **Casadaban, M. J., and S. N. Cohen.** 1979. Lactose genes fused to exogenous promoters in one step using a Mu-*lac* bacteriophage: in vivo probe for transcriptional control sequences. Proc. Natl. Acad. Sci. USA **76**:4530–4533.
6. **Cass, L. G., A. H. Horwitz, C. G. Miyada, L. Greenfield, and G. Wilcox.** 1980. The *araC* gene mRNA contains a leader sequence. Mol. Gen. Genet. **180**:219–226.
7. **Clarke, P., H.-C. Lin, and G. Wilcox.** 1982. The nucleotide sequence of the *araC* regulatory gene in *Salmonella typhimurium* LT2. Gene **18**:157–163.
8. **Covarrubias, L., L. Cervantes, A. Covarrubias, X. Soberon, I. Vichido, A. Blanco, Y. M. Kupersztoch-Portnoy, and F. Bolivar.** 1981. Construction and characterization of new cloning vehicles. V. Mobilization and coding properties of pBR322 and several deletion derivatives including pBR327 and pBR328. Gene **13**:25–35.
9. **Duncan, B. K., P. A. Rockstroh, and H. R. Warner.** 1978. *Escherichia coli* K-12 mutants deficient in uracil-DNA glycosylase. J. Bacteriol. **134**:1039–1045.
10. **Dunn, T. M., S. Hahn, S. Ogden, and R. Schleif.** 1984. An operator at −280 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc. Natl. Acad. Sci. USA **81**:5017–5020.
11. **Englesberg, E.** 1971. Regulation in the L-arabinose system, p. 257–296. *In* H. Vogel (ed.), Metabolic pathways, vol. 5. Academic Press, Inc., New York.
12. **Englesberg, E., D. Sheppard, C. Squires, and F. Meronk, Jr.** 1969. An analysis of "revertants" of a deletion mutant in the *C* gene of the L-arabinose gene complex in *Escherichia coli* B/r: isolation of initiator constitutive mutants (I^c). J. Mol. Biol. **43**:281–298.
13. **Englesberg, E., and G. Wilcox.** 1974. Regulation: positive control. Annu. Rev. Genet. **8**:219–242.
14. **Gielow, L., M. Largen, and E. Englesberg.** 1971. Initiator constitutive mutants of the L-arabinose operon (OIBAD) of *Escherichia coli* B/r. Genetics **69**:289–302.
15. **Greenfield, L., T. Boone, and G. Wilcox.** 1978. DNA sequence of the *araBAD* promoter in *Escherichia coli* B/r. Proc. Natl. Acad. Sci. USA **75**:4724–4728.
16. **Hahn, S., and R. Schleif.** 1983. In vivo regulation of the *Escherichia coli araC* promoter. J. Bacteriol. **155**:593–600.
17. **Heinemann, U.** 1984. Molekulare Basis der Genkontrolle. Nachr. Chem. Tech. Lab. **32**:712–716.
18. **Hendrickson, W., and R. Schleif.** 1984. Regulation of the *Escherichia coli* L-arabinose operon studied by gel electrophoresis DNA binding assay. J. Mol. Biol. **174**:611–628.
19. **Hendrickson, W., and R. Schleif.** 1985. A dimer of AraC protein contacts three adjacent major groove regions of the *araI* DNA site. Proc. Natl. Acad. Sci. USA **82**:3129–3133.
20. **Horwitz, A. H., L. Heffernan, C. Morandi, J.-H. Lee, J. Timko, and G. Wilcox.** 1981. DNA sequence of the *araBAD-araC* controlling region in *Salmonella typhimurium* LT2. Gene **14**:309–319.
21. **Horwitz, A. H., C. M. Miyada, and G. Wilcox.** 1984. Functional limits of the $araI^c$ promoter suggest an additional regulatory site for *araBAD* expression. J. Bacteriol. **158**:141–147.
22. **Ingraham, J. L., O. Maaloe, and F. C. Neidhardt.** 1983. Growth of the bacterial cell, p. 3. Sinaver Associates Inc., Sunderland, Mass.
23. **Isackson, P., and K. P. Bertrand.** 1985. Dominant negative mutations in the Tn*10 tet* repressor: evidence for use of the conserved helix-turn-helix motif in DNA binding. Proc. Natl. Acad. Sci. USA **82**:6226–6230.
24. **Ito, H., Y. Ike, S. Ikuta, and K. Itakura.** 1982. Solid phase synthesis of polynucleotides. VI. Further studies on polystyrene copolymers for the solid support. Nucleic Acids Res. **10**:1755–1769.
25. **Kelly, R. L., and C. Yanofsky.** 1985. Mutational studies with the *trp* repressor of *Escherichia coli* support the helix-turn-helix model of repressor recognition of operator DNA. Proc. Natl. Acad. Sci. USA **82**:483–487.
26. **Lee, N.** 1978. Molecular aspects of *ara* regulation, p. 389–409. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Lee, N. L., W. O. Gielow, and R. G. Wallace.** 1981. Mechanism of *araC* autoregulation and the domains of two overlapping promoters, P_C and P_{BAD} in the L-arabinose regulatory region of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **78**:752–756.
28. **Lei, S.-P., H.-C. Lin, L. Heffernan, and G. Wilcox.** 1985. The *araB* gene and the nucleotide sequence of the *araC* gene of *Erwinia carotovora*. J. Bacteriol. **164**:717–722.
29. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual, p. 89–91, 113. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **McKay, D. B., I. T. Weber, and T. A. Steitz.** 1982. Structure of catabolite gene activator protein at 2.9-Å resolution. J. Mol. Biol. **257**:9518–9524.
31. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 352–355, 431. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Miller, J. H.** 1978. The *lacI* gene: its role in *lac* operon control and its use as a genetic system, p. 31–88. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Miyada, C. G., A. H. Horwitz, L. G. Cass, J. Timko, and G. Wilcox.** 1980. DNA sequence of the *araC* regulatory gene from *Escherichia coli* B/r. Nucleic Acids Res. **8**:5267–5274.
34. **Miyada, C. G., L. Stoltzfus, and G. Wilcox.** 1984. Regulation of the *araC* gene of *Escherichia coli*: catabolite repression, autoregulation, and effect on *araBAD* expression. Proc. Natl. Acad. Sci. USA **81**:4120–4124.
35. **Novick, R. P., R. C. Clowes, S. N. Cohen, R. Curtiss III, N. Datta, and S. Falkow.** 1976. Uniform nomenclature for bacterial plasmids: a proposal. Bacteriol. Rev. **40**:168–189.
36. **Ogden, S., D. Haggerty, C. M. Stoner, D. Kolodrubetz, and R. Schleif.** 1980. The *Escherichia coli* L-arabinose operon: binding sites of the regulatory proteins and a mechanism of positive and negative regulation. Proc. Natl. Acad. Sci. USA **77**:3346–3350.
37. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53**:293–321.

38. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
39. **Smith, B., and R. Schleif.** 1978. Nucleotide sequence of the L-arabinose regulatory region of *Escherichia coli* K-12. J. Biol. Chem. **253**:6931–6933.
40. **Steitz, T. A., D. H. Olendorf, D. B. McKay, W. F. Anderson, and B. W. Matthews.** 1982. Structural similarity in the DNA-binding domains of catabolite gene activator and *cro* repressor proteins. Proc. Natl. Acad. Sci. USA **79**:3097–3100.
41. **Sutcliffe, J. G.** 1978. pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. Nucleic Acids Res. **5**:2721–2728.
42. **Tweten, R. K., and R. J. Collier.** 1983. Molecular cloning and expression of gene fragments from corynebacteriophage β encoding enzymatically active peptides of diphtheria toxin. J. Bacteriol. **156**:680–685.
43. **Wallace, R. B., M. J. Johnson, S. V. Suggs, K. Miyoshi, R. Bhatt, and K. Itakura.** 1981. A set of synthetic oligodeoxyribonucleotide primers for DNA sequencing in the plasmid vector pBR322. Gene **16**:21–26.
44. **Wallace, R. G., N. Lee, and A. V. Fowler.** 1980. The *araC* gene of *Escherichia coli*: transcriptional and translational start-points and complete nucleotide sequence. Gene **12**:179–190.
45. **Weber, I. T., and T. A. Steitz.** 1984. Model of specific complex between catabolite gene activator protein and B-DNA suggested by electrostatic complementarity. Proc. Natl. Acad. Sci. USA **81**:3973–3977.
46. **Wilcox, G., J. Boulter, and N. Lee.** 1974. Direction of transcription of the regulatory gene *araC* in *Escherichia coli* B/r. Proc. Natl. Acad. Sci. USA **71**:3635–3639.
47. **Wilcox, G., K. J. Clemetson, D. V. Santi, and E. Englesberg.** 1971. Purification of the *araC* protein. Proc. Natl. Acad. Sci. USA **68**:2145–2148.
48. **Wilcox, G., and P. Meuris.** 1976. Stabilization and size of *araC* protein. Mol. Gen. Genet. **145**:97–100.