JOURNAL OF BACTERIOLOGY, Apr. 1984, p. 141–147
0021-9193/84/040141-07$02.00/0

Vol. 158, No. 1

Functional Limits of the *araI*^{c} Promoter Suggest an Additional Regulatory Site for *araBAD* Expression

ARNOLD H. HORWITZ,† C. GARRETT MIYADA,‡ AND GARY WILCOX*

Department of Microbiology and Molecular Biology Institute, University of California, Los Angeles, California 90024

Received 31 May 1983/Accepted 10 January 1984

The *araBAD* promoter is defined, in part, by two types of *cis*-acting constitutive mutations, *araI*^{c} at position −35 and *araX*^{c} at position −10. Subcloning experiments demonstrated that the *araI*^{c} and *araI*^{c}*X*^{c} promoters require DNA sequence information out to position −53 to −56 for maximum constitutive expression. This is 8 to 10 base pairs more DNA than is generally thought to be necessary for RNA polymerase interaction. The −53 to −56 region is required for glucose repression, suggesting that an additional factor interacts in this region and is necessary for maximum expression.

The *araBAD* operon in *Escherichia coli* requires two positive regulatory proteins for maximum expression (8, 13). One is the 3′:5′ cyclic AMP (cAMP) receptor protein (CRP), which requires cAMP to stimulate *araBAD* expression. Like other systems that require this protein, the *araBAD* operon is sensitive to glucose repression (25). The second protein is the product of the closely linked, positive regulatory gene, *araC*, which is specific for the *ara* system. The requirement for *araC* protein is absolute, since strains that contain a deletion in the *araC* gene are unable to express *araBAD*. The results of DNA protection studies have demonstrated that the CRP-cAMP complex binds in vitro to the *araBAD* promoter from position −78 to −108 (+1 = start of *araBAD* transcription), *araC* protein (in the presence of L-arabinose) binds from −40 to −78, and RNA polymerase binds from +20 to −40 (14, 19). Based on the results of the protection studies as well as other studies on the *araBAD* operon, several models have been proposed to explain how these regulatory proteins interact with the controlling region DNA to promote *araBAD* expression (14, 19, 25). A common assumption of these models is that the CRP-cAMP complex interaction occurs at position −78 to −108 to stimulate maximum levels of *araBAD* expression. However, proof of this assumption awaits further experiments.

The *araBAD* promoter has been genetically defined, in part, by two classes of *cis*-acting constitutive mutations. One of these, designated *araI*^{c}, was isolated as a revertant of an *araC* deletion strain (7, 9). This mutation permits a low level of *araBAD* expression (5 to 10% of wild-type induced levels) in the absence of the *araC* gene product. Like the wild-type *araBAD* promoter, the *araI*^{c} promoter requires the CRP-cAMP complex for maximum expression and is sensitive to glucose repression (1, 5). The other type of mutation, designated *araX*^{c}, was isolated by mutagenizing an *araI*^{c} Δ*araC* strain and selecting for a higher-level constitutive mutant (5). The *araX*^{c} mutation also maps within the *araBAD* promoter and is separable from the *araI*^{c} mutation. The resulting *araI*^{c}*X*^{c} strain constitutively expresses *araBAD* at 25 to 30% of wild-type induced levels. In addition to providing a fivefold stimulation of constitutive expression over that of the *araI*^{c} mutation, the *araX*^{c} mutation confers some resistance to glucose repression.

We have described the cloning and DNA sequence analysis of the *araI*^{c} and *araX*^{c} mutations (3, 11). Fourteen independently isolated *araI*^{c} mutations were examined and found to contain the identical AT-to-GC transition at position −35 of the *araBAD* promoter. Three independently isolated *araX*^{c} mutations were found to contain the same GC-to-AT transition at position −10 of the promoter (11). The location of the mutations demonstrated the significance of the −35 and −10 regions of the *araBAD* promoter, but it did not explain how these mutations function or how they relate to the proposed models (14, 19, 25) for *araBAD* expression. For example, it is difficult to understand how expression of the *araI*^{c} promoter in the absence of *araC* protein could be modulated by the CRP-cAMP complex, which binds 35 to 65 base pairs (bp) upstream from the RNA polymerase interaction site.

The purpose of this study was to define the functional boundaries of the *araI*^{c} and *araI*^{c}*X*^{c} promoters. The plasmids bearing either the *araI*^{c} or *araI*^{c}*X*^{c} promoters *cis* to *araB* were uniformly truncated at position −47 by restriction at a *Bam*HI site. The resulting plasmids were greatly reduced in their ability to express *araB*. We attempted to reconstruct a functional *araI*^{c} and *araI*^{c}*X*^{c} promoter by cloning specific restriction fragments back into the *Bam*HI site. These experiments demonstrate that the *araI*^{c} and *araI*^{c}*X*^{c} promoters require DNA sequence information out to position −53 to −56 of the promoter. We also present results suggesting that glucose repression of both the *araI*^{c} and *araI*^{c}*X*^{c} promoters requires the region between −47 and −53 to −56. The implications of these results are discussed in relation to the proposed models for the wild-type promoter.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains used in this study are described in Table 1. Plasmids used in this study that have been described previously are shown in Table 1. Plasmids constructed in this study are described in the text and shown in Fig. 1 and 2 and in Table 2. Plasmids were introduced into the various bacterial strains by transformation as described by Cohen et al. (4). Plasmid DNA was prepared as described by Norgard et al. (18).

Media, chemicals, and enzymes. TYE broth has been previously described (2). Casamino Acids were obtained

* Corresponding author.
† Present address: International Genetic Engineering, Inc., Santa Monica, CA 90404.
‡ Present address: Beckman Institute of the City of Hope, Department of Molecular Genetics, Duarte, CA 91010.

TABLE 1. Bacterial strains and plasmids

Strain	Genotype	Origin or reference
Bacterial strains[a]		
SB1652	F⁻ Δ(*araBPOC-leuBCD*)*1119 leu-6 dau-5 rpsl*	12
LA310	F⁺ *araD*⁺ *A*⁺ Δ(*araB*)*809 araI*⁺ *araC*⁺ *lac gal pro thi hsdS*20	11
LA663	F⁻ Δ(*araABPOC*)*744* dau-5 rpsL	This paper
LA761	F⁺ Δ(*araABPOC*)*744 lac gal pro thi hsdS*20	This paper
Plasmids[b]		
pAH1	pBR322 Ω(*PstI,bla'*::B/r *araA'B*⁺*I*ᶜ110*X*ᶜ44 Δ*araC*766)	11
pAH5	pBR322 Ω(*PstI,bla'*::B/r *araA'B*⁺*I*ᶜ110 Δ*araC766*	11
pSB198	pBRH4 Ω(*Eco*RI::B/r *araA'B*⁺*I*⁺*C'*)	24
pAP-C	pPV33 Ω(*Eco*RI::B/r *araA'B*⁺*I*⁺ϕ[*araC'-tet*⁺])	24
pCGM1	pPV33 Ω(*Eco*RI::B/r *araA'B*⁺*I*⁺ Δ*araACT* ϕ[*araC'-tet*⁺])	17
PCGM3	pPV33 Ω(*Eco*RI::B/r *araA'B*⁺*I*⁺ Δ*araCRP* ϕ[*araC'-tet*⁺])	17

[a] Strains SB1652 and LA663 are derivatives of *E. coli* B/r. Strains LA310 and LA761 are derivatives of *E. coli* K-12 strain RR1.

[b] Genotype descriptions: Ω, Insertion of a restriction fragment in the designated vector; ϕ[*araC'-tet*⁺], fusion of the 3' end of a truncated *araC* gene to an intact *tet* gene; *PstI,bla'*::, insertion of a *PstI* restriction fragment at the *PstI* site in the *bla* gene of pBR322; *Eco*RI::, insertion of an *Eco*RI restriction fragment at the *Eco*RI site of the designated vector; B/r, DNA is from *E. coli* B/r. The designation Δ*araACT* refers to a 3-bp deletion in the *araC* activator protein binding site at position −57 to −59 of the *araBAD* promoter. The designation Δ*araCRP* refers to a 3-bp deletion in the CRP-cAMP binding site at position −85 to −87 of the *araBAD* promoter.

from Difco Laboratories (Detroit, Mich.). MacConkey agar base (Difco) was supplemented with 10 g of L-arabinose per liter (Sigma Chemical Co., St. Louis, Mo.). Tetracycline hydrochloride (Sigma) or ampicillin (Wyeth Laboratories, Philadelphia, Pa.) was added when required at a final concentration of 15 or 50 μg/ml, respectively.

Agarose (type II) and cAMP were obtained from Sigma. Phenol and ether were from Mallinckrodt Chemical Corp. (St. Louis, Mo.). Cesium chloride was from Kawecki-Berylco (Long Beach, Calif.). Acrylamide was from Miles Chemicals (Elkhart, Ind.). Ammonium peroxydisulfate and bisacrylamide were from Eastman Organic Chemicals (Rochester, N.Y.). [5,6-^{3}H]uridine (specific activity, 55 Ci/mmol) was from ICN (Irvine, Calif.).

All enzymes were either purified by published procedures or obtained from commercial sources.

Plasmid constructions. (i) Subcloning of *araI*ᶜ and *araI*ᶜ*X*ᶜ alleles. An *Eco*RI* digest of plasmid pAH5 containing the *araI*ᶜ allele and plasmid pAH1 (11) containing the *araI*ᶜ*X*ᶜ allele (Table 1) were ligated to pBR322 digested with *Eco*RI. An *Eco*RI digest of plasmid pSB198 containing the *araI*⁺ allele (Table 1) was ligated to pBR322 digested with *Eco*RI. The ligation mixtures were used to transform strain LA310 (Table 1) to Ara⁺ Tcʳ Apʳ. Plasmids were obtained for all three alleles which contained a 2.4-kilobase *Eco*RI* fragment in the orientation shown in Fig. 1a.

Plasmids containing a partial *araBAD* promoter with the *araI*⁺, *araI*ᶜ or *araI*ᶜ*X*ᶜ allele were constructed by digesting the plasmids shown in Fig. 1a with *Bam*HI and ligating under conditions which favor a self-closure reaction. Plasmids pAH104 (*araI*⁺), pAH112 (*araI*ᶜ), and pAH109 (*araI*ᶜ*X*ᶜ) (Fig. 1b) were generated by this procedure.

(ii) Constructing a hybrid *Salmonella typhimurium-E. coli* promoter. The plasmids pAH112 (*araI*ᶜ*B*) and pAH104 (*araI*⁺*B*) (Fig. 1b) were digested with *Bam*HI and ligated with a 300-bp *Sau*3A fragment containing the *S. typhimurium ara* controlling region sequence upstream from position −47. Plasmids were obtained which contained the 300-bp *Sau*3A fragment inserted in the *Bam*HI site in the orientation shown in Fig. 1c.

(iii) Cloning the synthetic *araC* and CRP binding-site deletions. The *araBAD* promoter containing Δ*araACT* or Δ*araCRP* from pCGM1 or pCGM3, respectively (Table 1), or the wild-type sequence from pAP-C (Table 1) was subcloned into the *Bam*HI site of the partial *araI*ᶜ promoter of plasmid pAH112 as follows. Plasmids pCGM1, pCGM3, and pAP-C were digested with *Bam*HI. An 800-bp *Bam*HI fragment which contains the *araBAD* promoter region upstream of position −47 (including Δ*araACT* and Δ*araCRP*), the promoter and amino terminus of the *araC* gene, and the first 350 bp of the *tet* gene under transcriptional control of the *araC* promoter (Table 1) was ligated to *Bam*HI-digested pAH112. Plasmids were obtained which contained the 800-bp *Bam*HI fragment inserted in the *Bam*HI site of pAH112 in the orientation which restores the complete *araBAD* promoter and the *tet* gene (Fig. 1d). The structure of the plasmids was confirmed by restriction analysis, by their complementation patterns with strain LA310, and by their Tcʳ phenotype.

(iv) Cloning of the 90-bp *Sau*3A fragment from the *S. typhimurium araC* gene. The plasmids pAH112 (*araI*ᶜ*B*) and pAH104 (*araI*⁺*B*) were digested with *Bam*HI and ligated to a *Sau*3A digest of a 1.5-kilobase *Taq*I fragment which contains the entire coding region of the *S. typhimurium araC* gene and includes the 90-bp *Sau*3A fragment of interest (see Fig. 3 and below). The ligated DNAs were transformed into strain SB1652 and plated on MacConkey agar containing arabinose and ampicillin. The pAH112 ligation mix gave rise to three types of colonies in SB1652. The majority (400) resembled the parent plasmid (+/− response in Table 2) and probably resulted from self-closure or insertion of restriction fragments which had no effect on *araI*ᶜ expression. Seven weak Ara⁺ colonies (+1 response in Table 2) and eight strong Ara⁺ colonies (+2 response in Table 2) were also observed. No weak or strong Ara⁺ colonies were observed out of 2,000 colonies with the pAH104 ligation mix. Several representatives of each colony type (weak Ara⁺ and strong Ara⁺) were purified, and plasmid DNA was prepared. Digestion of the plasmids isolated from the weak Ara⁺ clones with *Bam*HI + *Ava*I + *Eco*RI (see Fig. 1e) revealed that these plasmids contained an insert of approximately 90 bp which restored the *Bam*HI site in the *araBAD* promoter. A *Hin*fI digest (see Fig. 1e) confirmed that this plasmid contained a 90-bp *Bam*HI-sensitive insert in the *araBAD* promoter, and a *Sau*3A digest confirmed that this plasmid contained a 90-bp *Sau*3A fragment not present in pAH112. This plasmid was designated pAH125 and contained the 90-bp *Sau*3A fragment

TABLE 2. Genetic complementation tests and mRNA-DNA hybridization assays

Plasmid	*araBAD* promoter structure[a]		Complementation[b]		Hybridization (% input cpm)[c]
	−1 to −47	Beyond −47	SB1652	LA310	
pHM7	*araI*$^+$	Complete	−	+2	0.03
pAH104	*araI*$^+$	Δ	−	− (R)	<0.01
pAH131	*araI*$^+$	*Sau*3A no. 1	−	− (R)	<0.01
pHM1	*araI*^{c}	Complete	+2	+2	0.33
pAH112	*araI*^{c}	Δ	+/−	− (R)	0.07
pAH125	*araI*^{c}	*Sau*3A no. 1	+1	− (R)	0.13
pAH124	*araI*^{c}	*Sau*3A no. 2	+2	+/−	ND
pHM4	*araI*^{c}*X*^{c}	Complete	+2	+2	1.50
pAH109	*araI*^{c}*X*^{c}	Δ	+2	+1	0.23
pAH132	*araI*^{c}*X*^{c}	*Sau*3A no. 1	+2	+2	0.42
pAH129	*araI*^{c}*X*^{c}	*Sau*3A no. 2	+2	+2	1.09
pAH105	*araI*$^+$	Complete (*araC*$^+$)	+3	+3	4.50[d]

[a] Descriptions: complete, *araBAD* promoter sequences upstream of position −47 are present; Δ, *araBAD* promoter sequences upstream of position −47 have been deleted; *Sau*3A no. 1 and no. 2, 90-bp *Sau*3A fragment inserted into *Bam*HI site in orientations 1 or 2 as shown in Fig. 1e. All or part of the *araC* gene is deleted in all plasmids except pAH105.

[b] The L-arabinose phenotype was deduced from fermentation response at 24 h on MacConkey agar medium supplemented with L-arabinose: +3, small, intensely red colonies; +2, large red colonies; +1, small red center, white periphery; +/−, pink colony; −, white colony; R, recombination at 48 h.

[c] ^{3}H-labeled RNA was prepared as described in the text from strain LA663 containing the various plasmids. Hybridized counts per minute were determined by subtracting the counts per minute bound nonspecifically to M13mp2 plus strand DNA from the counts per minute bound to M13mp2::*araB* plus strand DNA. The difference was divided by the total input counts per minute. The deviation between individual experiments was less than 10%. ND, Not determined.

[d] Induced with 0.4% L-arabinose.

cloned in the *Bam*HI site of pAH112 in the orientation (designated as no. 1 in Fig. 1e) which restores the *araI*^{c} promoter out to position −53 (see below). The plasmids isolated from the strong Ara$^+$ clones also contained the 90-bp *Sau*3A fragment cloned in the *Bam*HI site of pAH112 but in the opposite orientation (designated as no. 2 in Fig. 1e), which resulted in the *Bam*HI site being restored in pBR322. This plasmid was designated pAH124. The 90-bp *Sau*3A fragment in orientation no. 1 (Fig. 1e) was subcloned adjacent to the partial *araI*$^+$ and *araI*^{c}*X*^{c} promoters by using the approach outlined in Fig. 2. The resulting plasmids were designated pAH131 (*araI*$^+$) and pAH132 (*araI*^{c}*X*^{c}).

Plasmid pAH129 was constructed by ligating *Bam*HI-digested pAH109 (*araI*^{c}*X*^{c}*B*) with the *Sau*3A digest of the 1.5-kilobase *Taq*I fragment described above. The ligated DNA was transformed into strain LA310. Plasmid DNA isolated from clones which gave a +2 response on MacConkey agar containing L-arabinose and ampicillin (see Table 2) was screened for insertion of the 90-bp *Sau*3A fragment as described above for pAH124 and pAH125.

Complementation assay for *araBAD* promoter activity. Plasmids were assayed for *araBAD* expression by transforming either strain LA310 or SB1652 (Table 1). Strain LA310 contains a small deletion in *araB* and is complemented by plasmids containing the *araB* gene and the intact *araBAD* promoter. Strain SB1652 contains deletion 1119 extending from *leuA* to *araB*, placing *araA* and *araD* under control of the leucine promoter. This strain expresses *araA* and *araD* at 30% of wild-type induced levels as a result of a constitutive *leu* promoter mutation, *leu-6* (12). Strain SB1652 is complemented by plasmids that contain the *araC* gene and the *araI*$^+$ allele *cis* to *araB* and by plasmids that are *araC*$^-$ but contain the *araI*^{c} or *araI*^{c}*X*^{c} alleles *cis* to *araB* (11). Transformed cells were plated on MacConkey agar supplemented with L-arabinose and ampicillin. The level of *araBAD* promoter activity was deduced from a fermentation response at 24 h as described in Table 2.

mRNA-DNA hybridization assay. Plasmid-containing cells were first grown to saturation in TYE broth containing ampicillin. A loopful of this culture was inoculated into 5 ml of M-9 medium (16) supplemented with 10 g of Casamino Acids per liter and 50 μg of tryptophan, 2 μg of thiamine hydrochloride, and 50 μg of ampicillin per ml and incubated at 37°C to saturation. This culture was then diluted 15 to 25 times into 50 ml of the same medium so that the starting turbidity was 20 to 30 Klett units (filter no. 42). The cultures were shaken at 300 rpm and grown at 37°C. All cultures grew at approximately the same rate. When the turbidity reached 120 Klett units (4×10^8 cells per ml), a 10-ml portion was removed, and 100 μCi of [^{3}H]uridine was added for 3 min at 30°C (300 rpm). When required, L-arabinose was added to a final concentration of 0.4% at a turbidity of 100 Klett units and the cells were grown to a turbidity of 120 Klett units before labeling was performed. RNA was isolated as described by Wilcox et al. (26). DNA-RNA hybridization was performed as described by Miyada et al. (17) except that 0.5 μg of the M13-mp2 and M13-mp2::*araB* plus-strand DNAs were used in the hybridizations.

RESULTS

Subcloning and expression of the *Bam*HI-*Eco*RI *ara*B fragment. The initial experiment to define the functional limits of the *araI*^{c} and *araI*^{c}*X*^{c} promoters utilized the *Bam*HI site at position −42 to −47 (see Fig. 3) to remove promoter sequences upstream from position −47. The resulting plasmids, shown in Fig. 1b, contain an intact *araB* gene and the first 47 bp of the promoter with either the *araI*$^+$ (pAH104), *araI*^{c} (pAH112), or *araI*^{c}*X*^{c} (pAH109) alleles, which were fused to pBR322 at the *Bam*HI site (see above). The constitutive activity of the partial *araI*^{c} and *araI*^{c}*X*^{c} promoters was determined by the complementation assay with strain SB1652 (Table 2). The *araI*^{c} promoter clearly required sequences upstream from position −47 in the *araBAD* promoter for maximal expression.

To quantitate the effect of removing the upstream se-

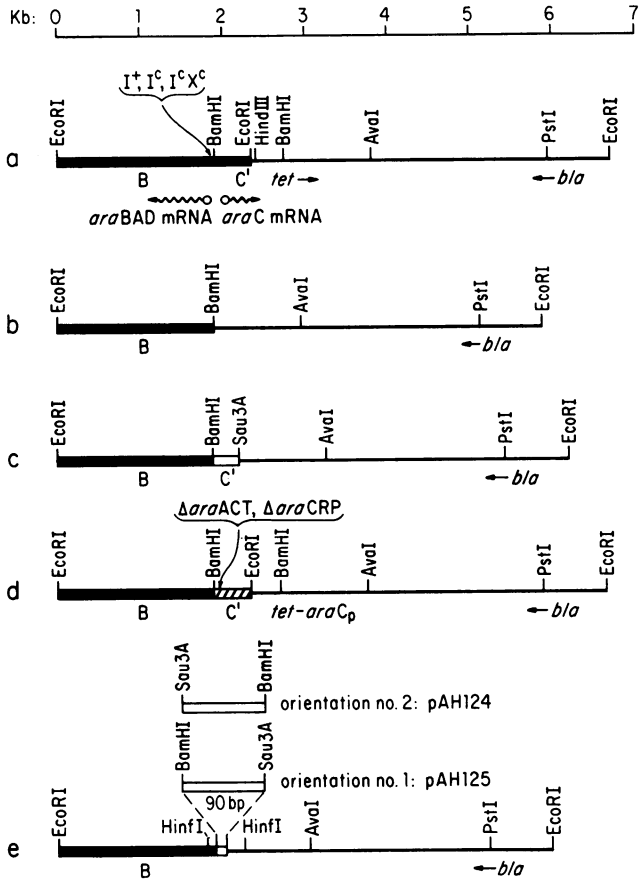

FIG. 1. Plasmids containing the $araI^+$, $araI^c$, and $araI^cX^c$ promoter alleles with various alterations upstream of a *Bam*HI site at position −47 of the *araBAD* promoter. (a) Plasmids constructed by an *Eco*RI* digestion of plasmids containing a large *PstI ara* insert. The *Eco*RI site was restored at both ends of the *Eco*RI* insert. (b) Plasmids obtained by digesting the plasmids in (a) with *Bam*HI and allowing self-closure to occur. (c) Plasmids constructed by subcloning a 300-bp *Bam*HI-*Sau*3A fragment containing the *S. typhimurium araBAD* promoter region into the *Bam*HI site of the plasmids in (b). (d) Plasmids obtained by subcloning an 800-bp *Bam*HI fragment containing the *araBAD* promoter with the wild-type sequence or with 3-bp deletions in the *araC* activator binding site (Δ*araACT*) or the CRP-cAMP binding site (Δ*araCRP*) into the *Bam*HI site of the plasmids in (b). (e) Plasmids obtained by subcloning a 90-bp *Sau*3A fragment from the *S. typhimurium araC* gene in both possible orientations into the *Bam*HI site of plasmid pAH112. The solid thin line represents pBR322 DNA. The solid thick line represents *E. coli ara* DNA. The open thick line represents *S. typhimurium ara* DNA. The hatched thick line represents *ara* DNA from the plasmids pCGM1, pCGM3, or pAP-C (see Table 1). Restriction sites are shown above the line. The approximate start points and the direction of transcription for *araBAD* and *araC* are indicated by horizontal wavy lines. The locations of the $araI^+$, $araI^c$, and $araI^cX^c$ alleles are shown above the line in (a). The locations of Δ*araACT* and Δ*araCRP* are shown above the line in (d). Abbreviations: B, *araB*; C', truncated *araC* gene; $tet\text{-}araC_p$, *tet* gene under transcriptional control of the *araC* promoter.

quences on $araI^c$ and $araI^cX^c$ promoter strengths, the constitutive (uninduced) level of transcription was determined by hybridization of single-stranded M13 mp2-*araB* DNA (17) with pulse-labeled mRNA isolated from strain LA663 carrying plasmids with either the complete or partial $araI^+$, $araI^c$, or $araI^cX^c$ promoters. As a control, the level of *araB* transcription under induced conditions was determined for strain LA663 carrying the plasmid pAH105, which contains complete *araC* and *araB* genes. The results of these experiments are shown in Table 2. The relative strengths of the complete $araI^c$ (pHM1) and $araI^cX^c$ (pHM4) promoters compared to each other and to the $araI^+$ promoter in the presence (pAH105) and absence (pHM7) of the *araC* activator corresponded to previously published data for these mutations as measured on the chromosome (5). Transcription from the partial $araI^c$ (pAH112) and $araI^cX^c$ (pAH109) promoters was reduced five- to six-fold compared to their respective complete promoters. Although this degree of reduction was apparently great enough to result in significantly reduced complementation of strain SB1652 by the plasmid carrying *araB cis* to the partial $araI^c$ promoter (pAH112), it was not great enough to have a significant effect on the ability of the plasmid carrying *araB cis* to the partial

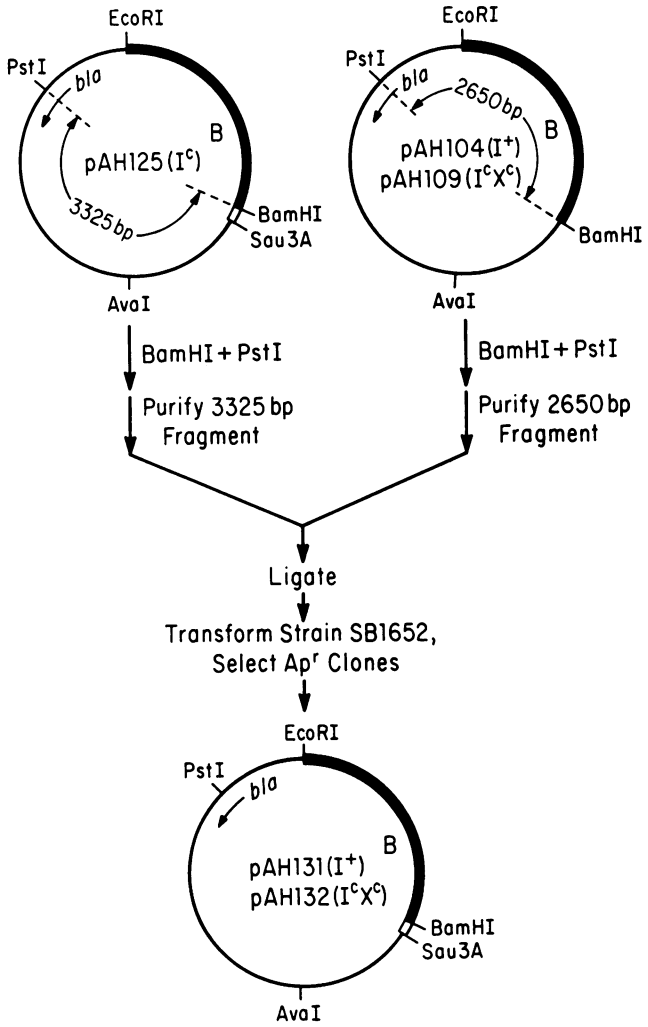

FIG. 2. Subcloning of the 90-bp *S. typhimurium araC* gene *Sau*3A fragment in orientation no. 1 (see Fig. 1e) adjacent to the *Bam*HI site of the partial $araI^+$ and $araI^cX^c$ promoters. The solid thin line represents pBR322 DNA. The solid thick line represents *E. coli ara* DNA. The open thick line represents *S. typhimurium ara* DNA. B, *araB*.

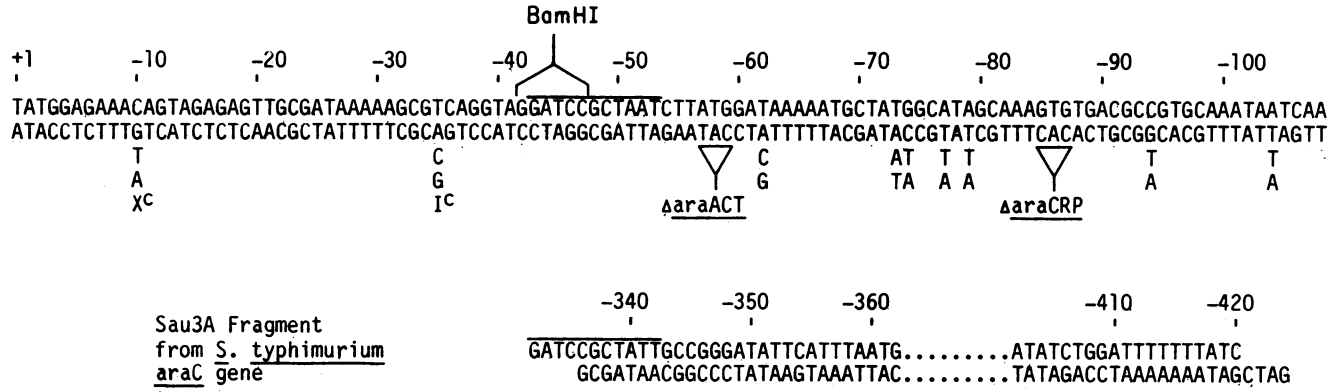

FIG. 3. The DNA sequence of the *araBAD* promoter. The locations of the *araI*^{c}, *araX*^{c}, Δ*ara*ACT and Δ*ara*CRP mutations are shown below the DNA sequence. The remaining bp below the *E. coli* sequence represent the base changes found upstream of the *Bam*HI site in the *araBAD* promoter for *S. typhimurium*. The DNA sequence of the ends of a 90-bp *Sau*3A fragment from the *araC* gene of *S. typhimurium* is also shown. A region of homology between the *S. typhimurium araC* gene sequence from −332 to −342 and the *araBAD* promoter from position −43 to −53 is overlined in both sequences. The location of the *Bam*HI site in the *araBAD* promoter is also shown.

araI^c*X*^{c} promoter (pAH109) to complement SB1652. This result was expected since the mRNA level of the partial *araI*^{c}*X*^{c} promoter is similar to that for the complete *araI*^{c} promoter.

Reconstruction of the *araI*^{c} promoter. A series of subcloning experiments was undertaken to define the minimum sequence required for restoration of a functional *araI*^{c} promoter.

(i) Hybrid *S. typhimurium-E. coli* promoter. Comparison of the *S. typhimurium* and *E. coli araBAD* promoter sequences revealed considerable sequence homology with no insertions or deletions (Fig. 3) (10). Since both promoters contain a *Bam*HI site at position −42 to −47, a hybrid *E. coli-S. typhimurium* promoter was constructed with this site as a junction point. Since there are 7-bp differences between the *E. coli* and *S. typhimurium* sequence from position −62 to −104 (Fig. 3), it was of interest to determine whether *E. coli araI*^{c} promoter function could be restored by placing this *S. typhimurium* promoter sequence upstream from the *Bam*HI site. Complementation assays with strain SB1652 revealed that the *S. typhimurium* sequence restored *araI*^{c} promoter function. The *araI*$^+$*B* plasmid (pAH104) which contained the same 300-bp *Sau*3A fragment inserted at the *Bam*HI site failed to complement SB1652, demonstrating that the restored *araI*^{c} promoter function resulted from replacement of necessary upstream sequences and not exogenous promoter activity in the fragment.

(ii) Synthetic *araC* and CRP binding-site deletions. Miyada et al. (17) used site-specific mutagenesis with synthetic oligonucleotides to place 3-bp deletions in the *araC* activator protein-binding site (Δ*araACT*) and the CRP-binding site (Δ*araCRP*) of the *araBAD* promoter (Fig. 3). The activities of the *araBAD* promoter containing wild-type, Δ*araCRP*, or Δ*araACT* regions *cis* to *araI*^{c}(Fig. 1d) were similar as determined by the complementation assay. We conclude that the regions deleted by Δ*araACT* and Δ*araCRP* are not essential for *araI*^{c} promoter function. The sequences necessary for *araI*^{c} promoter function probably reside between the endpoint of the *Bam*HI site at position −48 and the starting point of Δ*araACT* at position −57.

(iii) Cloning of the 90-bp *Sau*3A fragment from the *S. typhimurium araC* gene. A comparison of the *araBAD* promoter sequence with the initial translated region of the *araC* gene of *S. typhimurium* revealed that the *araBAD* promoter sequence from position −43 to −53 was almost exactly (10 of 11 bp) repeated at position −332 to −342 in the initial translated region of the *araC* gene (Fig. 3). This sequence is contained on a 90-bp *Sau*3A fragment which extends from position −332 to −422 in the *araC* gene of *S. typhimurium* (Fig. 3). Insertion of this 90-bp *Sau*3A fragment at the *Bam*HI site of pAH112 was described above and outlined in Fig. 1e. Plasmid pAH125, which contained the *Sau*3A fragment adjacent to the partial *araI*^{c} promoter in the orientation (no. 1 in Fig. 1e) that partially restores the promoter to position −53 weakly complemented strain SB1652 (Table 2). Plasmid pAH124, which contained the *Sau*3A fragment adjacent to the partial *araI*^{c} promoter in orientation no. 2 (Fig. 1e) strongly complemented strain SB1652 (Table 2). The constitutive level of transcription was determined for the plasmid containing the *Sau*3A fragment in orientation no. 1 (pAH125) in strain LA663 (Table 2). Plasmid pAH125 produced 1.9-fold higher levels of mRNA than the same plasmid without this insert (pAH112).

To demonstrate that the partial restoration of *araI*^{c} promoter function was due to stimulation of promoter activity and not to exogenous promoter activity from the 90-bp *Sau*3A fragment, this fragment was subcloned adjacent to the partial *araI*$^+$ and *araI*^{c}*X*^{c} promoters in orientation no. 1 (Fig. 2). The resulting plasmids, designated pAH132 (*araI*^{c}*X*^{c}) and pAH131 (*araI*$^+$), were transformed into both strain SB1652 and strain LA310 (Table 2). The plasmid containing the 90-bp *Sau*3A fragment fused to the *araI*$^+$ promoter (pAH131) failed to complement strain SB1652 or LA310 and had no detectable transcriptional activity as determined by mRNA-DNA hybridizations. The plasmid containing the 90-bp *Sau*3A fragment fused to the *araI*^{c}*X*^{c} promoter (pAH132) produced 1.9-fold higher levels of mRNA than the same plasmid without this insert (pAH109) (Table 2). In addition, unlike plasmid pAH109, plasmid pAH132 strongly complemented strain LA310, which must result from increased levels of constitutive expression of *araB* on this plasmid.

The 90-bp *Sau*3A fragment was cloned in orientation no. 2 adjacent to the partial *araI*^{c}*X*^{c} promoter. This plasmid (pAH129) resulted in strong Ara$^+$ colonies when transformed into strain LA310 and produced levels of mRNA

which were five times that of pAH109 and 70% of the levels produced by the complete *araI*^{c}*X*^{c} promoter (Table 2).

Effect of glucose repression on whole and partial *araI*^{c} promoters. The results of in vitro DNA protection studies have revealed that RNA polymerase protects the wild-type *araBAD* promoter from position +20 to −40 (14). If only RNA polymerase is required for maximum expression of the *araI*^{c} and *araI*^{c}*X*^{c} promoters, then the partial and complete *araI*^{c} and *araI*^{c}*X*^{c} promoters should be equally active. Our results indicate that maximal expression of both the *araI*^{c} and *araI*^{c}*X*^{c} promoters requires DNA sequence information out to at least position −53 to −56 (Table 2). An alternative hypothesis is that maximal expression of the *araI*^{c} and *araI*^{c}*X*^{c} promoters requires interaction of an additional factor in the region surrounding the *Bam*HI site and extending out to position −53 to −56. One candidate for such a factor is the CRP-cAMP complex (or an entity controlled by this complex). If this is the case, constitutive expression of the whole *araI*^{c} promoter should be sensitive to glucose repression. By contrast, expression of the partial *araI*^{c} promoter should be relatively insensitive to glucose repression since the putative factor binding site has been damaged. This hypothesis was tested by measuring the effect of glucose repression on *araB* mRNA levels in strain LA663 containing plasmids with the whole *araI*^{c} (pHM1) and *araI*^{c}*X*^{c} (pHM4) promoters and the partial *araI*^{c}*X*^{c} promoter (pAH109). The *araI*^{c}*X*^{c} promoter was used for the comparison of whole and partial promoter activities because it produces higher levels of mRNA (due to the presence of the *araX*^{c} mutation) and thus provides a greater degree of sensitivity. The results of these experiments (Table 3) demonstrated that expression of the whole *araI*^{c} and *araI*^{c}*X*^{c} promoters was highly sensitive to glucose repression. This effect, as shown with the *araI*^{c}*X*^{c} promoter, was partially reversed by addition of cAMP to the glucose-containing medium. By comparison, the partial *araI*^{c}*X*^{c} promoter was insensitive to glucose repression.

DISCUSSION

We defined the functional limits of the *araBAD* promoter containing the *araI*^{c} and *araX*^{c} mutations. These mutations, which are located at positions −35 (*araI*^{c}) and −10 (*araX*^{c}), allow expression in the absence of the *araC* activator protein. Both the *araI*^{c} and *araI*^{c}*X*^{c} promoters appear to require DNA sequence information out to position −53 to −56 for maximal constitutive expression for the following reasons. (i) Removal of *ara* promoter sequences upstream from a unique *Bam*HI site at positions −42 and −47 resulted in a five- to sixfold reduction in the levels of transcription from the partial *araI*^{c} and *araI*^{c}*X*^{c} promoters (Fig. 1a and b, Table 2). Inspection of the pBR322 sequence to which the remaining 47 bp of the promoter were fused revealed that this sequence partially (3 of 4 bp) restored the *ara* promoter out to position −51. (ii) Normal constitutive expression was restored to the partial *araI*^{c} and *araI*^{c}*X*^{c} promoters by inserting a *Bam*HI fragment containing the upstream *araBAD* promoter sequence with a synthetically produced 3-bp deletion located at position −57 to −59 in the *araC* binding site (see Fig. 3). We conclude that sequences beyond position −56 are not required for *araI*^{c} and *araI*^{c}*X*^{c} promoter function. (iii) The function of the partial *araI*^{c} and *araI*^{c}*X*^{c} promoters was stimulated 1.9-fold by insertion at the *Bam*HI site of a 90-bp *Sau*3A fragment from the coding region of the *S. typhimurium araC* gene, which completely restored the *ara* promoter sequence out to position −51 and partially restored the promoter out to −53 (see orientation no. 1 in Fig. 1e). This stimulatory effect was specific for the *araI*^{c} and *araI*^{c}*X*^{c} promoters since the same sequence placed adjacent to the *araI*$^+$ promoter had no effect on expression (Fig. 2, Table 2).

The functions of the *araI*^{c} and *araI*^{c}*X*^{c} promoters were restored to approximately 70% of normal levels by insertion at the *Bam*HI site of the 90-bp *Sau*3A fragment from *S. typhimurium araC* gene in orientation no. 2 (Fig. 1e, Table 2). This effect was specific for the *araI*^{c} and *araI*^{c}*X*^{c} promoters since the *araI*$^+$ promoter did not appear to be activated by insertion of this fragment at the *Bam*HI site. Inspection of the DNA sequence of this fragment in the region that had been inserted adjacent to the *Bam*HI site revealed that it is AT rich. Although this sequence shared no homology with the wild-type *araBAD* promoter sequence from position −48 to −56, it was homologous to the wild-type *araBAD* promoter sequence from positions −60 to −68 (see Fig. 3). The significance of this homology is not known. It is, however, interesting to note that the presence of AT-rich areas around position −55 in some strong constitutive promoters may be an important determining factor in the strength of these promoters (6).

The results of enzymatic, chemical, and photochemical probes of RNA polymerase contacts with various promoters suggest that a promoter requiring only RNA polymerase for expression extends out to position −45 (21). This conclusion is supported by the observation that conserved promoter sequences generally do not extend beyond position −45 and by the fact that all known promoter mutations in the RNA polymerase recognition region (−35 region) are clustered within 5 bp of position −35 (21, 22). Since the *araI*^{c} and *araI*^{c}*X*^{c} promoters require sequence information out to position −53 to −56 for maximal expression, we conclude that these promoters may require interaction of a factor in addition to RNA polymerase for maximal expression.

Although an entity controlled by the CRP-cAMP complex could be the factor, the CRP-cAMP complex itself is a likely candidate for the additional factor for several reasons. First, the *araBAD* promoter containing the *araI*^{c} mutation, and to a lesser extent the *araI*^{c}*X*^{c} mutation, were previously found to be sensitive to glucose repression (1, 5). We obtained the same result for the *araI*^{c} promoter on a plasmid and have demonstrated that the *araI*^{c}*X*^{c} promoter on a plasimd is equally sensitive to glucose repression (Table 3). Second, the expression of the *araBAD* operon in vitro requires the

TABLE 3. Effect of glucose on transcription from whole and partial *araI*^{c} and *araI*^{c}*X*^{c} promoters

Plasmid	*ara* genotype[a]	Carbon source[b]	Hybridized cpm (% of total input)[c]
pHM1	*araB*$^+$*I*^{c} *araC'*	Glycerol	0.10
		Glucose	0.025
pHM4	*araB*$^+$*I*^{c}*X*^{c} *araC'*	Glycerol	0.88
		Glucose	0.22
		Glucose + 5 mM cAMP	0.44
pAH109	*araB*$^+$*I*^{c}*X*^{c} Δ*araBp*	Glycerol	0.10
		Glucose	0.12

[a] Genotypes are described in more detail in Table 2 and in the text. Δ*araBp* indicates the *araBAD* promoter deleted upstream of position −47. *araC'* indicates the *araC* gene truncated at the 3' end.

[b] The carbon source was added at 0.4% (wt/vol) to M9-Casamino Acids medium supplemented with tryptophan.

[c] ^{3}H-labeled RNA was prepared as described in the text from strain LA663 containing the various plasmids. Hybridized counts per minute were determined as described in Table 2, footnote *c*.

CRP-cAMP complex (13). Third, although the complete *araI*^{c} and *araI*^{c}*X*^{c} promoters were sensitive to glucose repression, the partial *araI*^{c}*X*^{c} promoter was insensitive (Table 3). Thus, sequence information at or upstream of the *Bam*HI site at positions −42 to −47 of the *araI*^{c} and *araI*^{c}*X*^{c} promoters is required for glucose repression. Fourth, the CRP-cAMP binding site found by DNA protection studies to extend from −78 to −108 of the *araBAD* promoter (14, 19) does not appear to play a role in *araI*^{c} and *araI*^{c}*X*^{c} promoter function since the *araI*^{c} promoter *cis* to a 3-bp deletion within this region at position −85 to −87 (see Fig. 1 and 3) appears to function normally. We conclude that the glucose repression effect may be due to CRP interaction at a different site. Fifth, there have been reports of several CRP-dependent promoters in which CRP-cAMP interacts in the region from −30 to −53 (15, 20, 23). The *araBAD* promoter sequence from positions −27 to −55 has some degree of homology with the CRP binding sites of these promoters.

A common assumption of all models presented thus far to explain regulation of *araBAD* operon expression (14, 19, 25) is that the CRP-cAMP interaction must occur at position −78 to −108 for the operon to be expressed at maximum levels. We have found that this CRP-cAMP site is not required for *araI*^{c} promoter expression. We suggest that the CRP-cAMP interaction instead may occur in the −30 to −53 region of the *araI*^{c} promoter. Although the *araI*^{c} mutation may have created a new CRP-cAMP binding site, it is also possible that this site already exists in the −30 to −53 region of the wild-type promoter and functions in glucose repression of *araBAD*.

ACKNOWLEDGMENTS

We thank Maureen Gilmore-Hebert, Laurel Heffernan, and Dan Ray for critically reading the manuscript. We also thank Marilyn Day for typing and editing the manuscript.

This work was supported by National Science Foundation grant PCM78-26255. A.H.H. was supported by a postdoctoral traineeship from National Cancer Institute grant CA 09056-06. C.G.M. was supported by Public Health Service training grant 07185. G.W. was supported by an American Cancer Society Faculty Research Award.

LITERATURE CITED

1. **Bass, R., L. Heffernan, K. Sweadner, and E. Englesberg.** 1976. The site for catabolite deactivation in the L-arabinose BAD operon in *Escherichia coli* B/r. Arch. Microbiol. **110**:135–143.
2. **Cass, L. G., A. H. Horwitz, C. G. Miyada, L. Greenfield, and G. Wilcox.** 1980. The *araC* regulatory gene mRNA contains a leader sequence. Mol. Gen. Genet. **180**:219–226.
3. **Cass, L. G., A. H. Horwitz, and G. Wilcox.** 1981. The *araI*^{c} mutation in *Escherichia coli* B/r. J. Bacteriol. **146**:1098–1105.
4. **Cohen, S. N., A. C. Y. Chang, and L. Heu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Natl. Acad. Sci. U.S.A. **69**:2110–2119.
5. **Colome, J., G. Wilcox, and E. Englesberg.** 1977. Constitutive mutations in the controlling site region of the *araBAD* operon of *Escherichia coli* that decrease sensitivity to catabolite repression. J. Bacteriol. **129**:948–958.
6. **de Boer, H. A., L. J. Comstock, D. G. Yansura, and H. L. Heyneker.** 1982. Construction of a tandem *trp-lac* promoter and a hybrid *trp-lac* promoter for efficient and controlled expression of the human growth hormone gene in *Escherichia coli*, p. 462–481. *In* R. L. Rodriguez and M. J. Chamberlin (ed.), Promoters, structure and function. Praeger Publishers, New York.
7. **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.
8. **Englesberg, E., and G. Wilcox.** 1974. Regulation: positive control. Annu. Rev. Genet. **8**:219–242.
9. **Gielow, L., M. Largon, and E. Englesberg.** 1971. Initiator constitutive mutations of the L-arabinose operon (O*IBAD*) of *Escherichia coli* B/r. Genetics **69**:289–302.
10. **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.
11. **Horwitz, A. H., C. Morandi, and G. Wilcox.** 1980. Deoxyribonucleic acid sequence of *araBAD* promoter mutants of *Escherichia coli*. J. Bacteriol. **142**:659–667.
12. **Kessler, D. P., and E. Englesberg.** 1969. Arabinose-leucine deletion mutants of *Escherichia coli* B/r. J. Bacteriol. **98**:1159–1169.
13. **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.
14. **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. U.S.A. **78**:752–756.
15. **Le Grice, S. F. J., and H. Matzura.** 1981. Binding of RNA polymerase and the catabolite gene activator protein within the *cat* promoter in *Escherichia coli*. J. Mol. Biol. **150**:185–196.
16. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 431. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Miyada, C. G., X. Soberon, K. Itakura, and G. Wilcox.** 1982. The use of synthetic oligodeoxyribonucleotides to produce specific deletions in the *araBAD* promoter of *Escherichia coli* B/r. Gene **17**:167–177.
18. **Norgard, M., K. Emigholz, and K. Monihan.** 1979. Increased amplification of pBR322 plasmid deoxyribonucleic acid in *Escherichia coli* K-12 strains RR1 and χ1776 grown in the presence of high concentrations of nucleoside. J. Bacteriol. **138**:270–272.
19. **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. U.S.A. **77**:3346–3350.
20. **Queen, C., and M. Rosenberg.** 1981. A promoter of pBR322 activated by cAMP receptor protein. Nucleic Acids Res. **9**:3365–3377.
21. **Simpson, R. B.** 1982. Evidence for close contacts between RNA polymerase and promoter DNA, p. 164–180. *In* R. L. Rodriguez and M. J. Chamberlin (ed.), Promoters, structure and function. Praeger Publishers, New York.
22. **Soberon, X., J. J. Rossi, G. P. Larson, and K. Itakura.** 1982. A synthetic consensus sequence procaryotic promoter is functional, p. 407–431. *In* R. L. Rodriguez and M. J. Chamberlin (ed.), Promoters, structure and function, Praeger Publishers, New York.
23. **Taniguchi, T., M. O'Neill, and B. de Crombrugghe.** 1979. Interaction site of *Escherichia coli* cyclic AMP receptor protein on DNA of galactose operon promoters. Proc. Natl. Acad. Sci. U.S.A. **76**:5090–5094.
24. **West, Jr., R. W., E. L. Kline, A. H. Horwitz, G. Wilcox, and R. L. Rodriguez.** 1982. Cloning and analysis of inducible *Escherichia coli* promoters using promoter-probe plasmid vectors, p. 363–386. *In* R. L Rodriguez and M. J. Chamberlin (ed.), Promoters, structure and function. Praeger Publishers, New York.
25. **Wilcox, G., S. Al-Zarban, L. G. Cass, P. Clarke, L. Heffernan, A. H. Horwitz, and C. G. Miyada.** 1982. DNA sequence analysis of mutants in the *araBAD* and *araC* promoters, p. 183–194. *In* R. L. Rodriguez and M. J. Chamberlin (ed.), Promoters, structure and function. Praeger Publishers, New York.
26. **Wilcox, G., J. Singer, and L. Heffernan.** 1971. Deoxyribonucleic acid-ribonucleic acid hybridization studies on the L-arabinose operon of *Escherichia coli* B/r. J. Bacteriol. **108**:1–4.