J. Mol. Biol. (1983) **171**, 369–381

The *araE* Low Affinity L-Arabinose Transport Promoter

Cloning, Sequence, Transcription Start Site and DNA Binding Sites of Regulatory Proteins

CAROL STONER AND ROBERT SCHLEIF

Department of Biochemistry
Brandeis University, Waltham, Mass. 02254, U.S.A.

(Received 26 January 1983, and in revised form 8 August 1983)

The promoter for the gene encoding the low affinity L-arabinose uptake protein in *Escherichia coli* was studied. The promoter was cloned, sequenced, its transcription start site determined by S_1 nuclease mapping, the proteins required for *in vitro* transcription were determined, and the regulatory protein binding sites located by DNase footprinting. The *araE* promoter shows no evidence of an operator site upstream from the CRP binding site, but otherwise it is similar to the *araBAD* promoter.

1. Introduction

The L-arabinose metabolic operon *araCBAD* in *Escherichia coli* is under the positive and negative control of *araC* regulatory protein (Englesberg *et al.*, 1965; Greenblatt & Schleif, 1971; Lee *et al.*, 1974). Previous studies have revealed the DNA sequence of the *araCBAD* regulatory region (Smith & Schleif, 1978) and the DNA binding sites of the regulatory proteins (Ogden *et al.*, 1980; Lee *et al.*, 1981). However, these studies have not shown the mechanism by which the promoter of the *araBAD* genes, p_{BAD}, is regulated. One approach to a better understanding of the mechanisms involved in this regulation is to compare p_{BAD} to other arabinose-induced promoters, including p_E and p_{FG}, which are themselves positively and negatively regulated by *araC* protein (Kolodrubetz & Schleif, 1981). One such promoter has been cloned and sequenced (Kosiba & Schleif, 1982), although the identity of the genes it drives has not been definitively proven.

This paper describes the cloning, sequencing, S_1 nuclease mapping, *in vitro* transcription and DNase I footprinting of the promoter for the genes involved in the low affinity uptake of arabinose, p_E. S_1 mapping of DNA/mRNA hybrids formed between *araE* DNA and total cellular RNA isolated from cells grown in the presence of inducer L-arabinose was used to determine the direction of transcription and the *in vivo* transcription start site of the *araE* message. *In vitro* transcription experiments were performed to confirm the transcription start site found *in vivo* and to determine the proteins necessary for transcription of *araE*.

0022-2836/83/360369-13 $03.00/0

DNase I footprinting was used to determine the DNA binding sites for cyclic AMP receptor protein (CRP) and *araC* protein, The binding sites for these proteins are located in the same order as they are on p_{BAD}, except for the apparent absence of a site analogous to *araO* (Ogden *et al.*, 1980). Thus the promoters p_E and p_{BAD} are probably induced by similar mechanisms, but may be repressed by different mechanisms.

2. Materials and Methods

(a) *Media, strains, chemicals and DNA manipulations*

Routine procedures and DNA preparations and manipulations were as described by Schleif & Wensink (1981) and Maxam & Gilbert (1980). All strains are derivatives of *E. coli* K-12 and are listed in Table 1.

Restriction endonucleases, S_1 nuclease, and *E. coli* RNA polymerase were purchased from New England BioLabs. *araC* protein was prepared according to the procedure described by Steffen & Schleif (1977) and Schleif & Favreau (1982). Protein concentrations in reactions refer to final concentrations of total protein (active + inactive). Based on DNA binding abilities (Hendrickson & Schleif, unpublished results), the *araC* protein was greater than 10% active. CRP was made according to the procedure described by Eilen *et al.* (1978). Its concentration and specific activity were 330 μg/ml and 4400 units/mg, respectively. CRP was stored in 1 M-KCl and the final salt concentration in all the samples was adjusted to remain constant at 50 mM. Cyclic AMP, L-arabinose and D-fucose (when present) were added to final concentrations of 200 μM, 100 mM and 100 mM, respectively.

(b) *Construction and verification of plasmids containing the* araE *regulatory region*

The construction and isolation of 2 plasmids, pRFS1599 and pRFS1595, containing the *araE* regulatory region was performed according to the procedure described by Kosiba & Schleif (1982). A Clarke–Carbon plasmid (Clarke & Carbon, 1976), pBEK75 containing the *araE* gene (Kolodrubetz, 1980) was digested with restriction endonucleases, *Alu*I, *Fnu*DII or *Hae*III. End-labeled *Eco*RI linkers were ligated onto the DNA fragments and the

TABLE 1

List of plasmids and strains

Strain	Genotype	Comments
pRFS1599	DLS24 : : *araE* 370 bp regulatory region ligated into pBR325, CM^s tet^r	From R. Schleif & A. Favreau
pRFS1595	DLS24 : : *araE* 1250 bp regulatory region ligated into pBR325, CM^s tet^r	From R. Schleif & A. Favreau
pBEK75	pBEK75 = DJK131 contains the *araE* region	Kolodrubetz (1980)
DLS24	*endoI hsdM hsdA* B_1^-	Steffen (1976)
RFS1563	pMB9 with *araCBAD* 440 regulatory region inserted into *Eco*RI site	Smith & Schleif (1978)
SH121	F′ *araB53 thr*$^+$ *leu*$^+$/Δ*araC1022* Δ*lac74*	From Schleif (1972)
SH121/84	F′ *araB53 thr*$^+$ *leu*$^+$/Δ*araC1022* Δ*lac74* Δ*cya strA*(Sm^r)	From Brickman *et al.* (1973)
SH218	F^- *leu* Δ*lac74 araF* : : Mu-*lac*$^+$ *amp*$^+$	From S. Hahn
SH246	F^- *leu* Δ*lac74 araF* : : Mu-*lac amp*$^+$ Δ*cya*	From S. Hahn

fragments were then ligated to *Eco*RI-cut pKB260-Kan vector plasmid and transformed into an Ara^+ strain, 294 (Kosiba & Schleif, 1982). Candidates were selected that showed arabinose-dependent tetracycline resistance.

The DNA fragment isolated by these experiments ought to contain the promoter for the *araE* genes. A further experiment was performed to strengthen this conclusion. DNA isolated from strains containing Mu-d*lac* inserted into the chromosomal *araE* gene (Kolodrubetz & Schleif, 1981) was digested with a variety of restriction enzymes, separated according to size on agarose gels, and probed with the putative p_E fragment (Southern transfer). The size of the fragment homologous to p_E following digestion with *Eco*RI, *Hin*dIII or *Bam*HI was altered compared to wild type (Hendrickson, unpublished results). The smallest of these fragments was about 2 kb†, thereby locating the p_E sequence within 2 kb of the *araE* gene.

(c) *DNA preparations*

The 1250 bp and 370 bp *araE* fragments were isolated by *Eco*RI restriction endonuclease digests of the respective plasmids, electrophoresis on 5% polyacrylamide preparative gels and elution from crushed gel slices. A 203 bp *Hae*III fragment containing the *lac* promoter (Majors, 1975) was used as a template in *in vitro* transcription experiments for a size standard.

(d) *End-labeling DNA fragments*

The 1250 bp or 370 bp DNA fragments were labeled at their 5′ protruding ends with phage T4 polynucleotide kinase, then digested with restriction endonuclease *Hin*fI or *Sau*3A to remove one of the 5′ labeled ends, electrophoresed on 6% polyacrylamide gels and eluted.

(e) *DNA sequence determination*

Both the sense and antisense strands of *araE* DNA fragments were sequenced. The sequencing reactions used were the G, A > C, C+T and C reactions of Maxam & Gilbert (1980). The DNA fragments were electrophoresed on 35 cm × 85 cm × 0·05 cm polyacrylamide gels as described by Sanger & Coulson (1978).

(f) *Preparation of total cellular RNA + DNA*

The procedure for the isolation of total cellular RNA is based on that described by Salser *et al.* (1967). The cells were grown in M10 glycerol medium except for the Δ*cya* strain, which was grown in M10 glucose. Cells were harvested at an A_{550} of 0·6 to 0·7. All cells were harvested by pouring 100 ml of culture into a centrifuge bottle containing 100 g of ice, 12·5 mg of chloramphenicol and 0·02 mol NaN_3. The cells were then centrifuged at 4000 ***g*** for 10 min, resuspended in 4 ml of 10 mM-Tris · HCl (pH 7·3), 10 mM-KCl, 5 mM-$MgCl_2$ and the RNA prepared as described by Salser *et al.* (1967).

(g) *S_1 nuclease mapping*

DNA/RNA hybridizations were performed essentially as described by Berk & Sharp (1977). In each reaction, 50 ng of end-labeled *araE* 1250 bp *Sau*3A or *araE* 370 bp *Hin*fI DNA fragments (see Fig. 1) containing 50,000 cts/min were added to 25 μg of total cellular RNA. Assuming *araE* messenger possesses a 2 min half-life and is translated with efficiency comparable to β-galactosidase messenger, these reactions contain a 10-fold molar excess of *araE* DNA to *araE* RNA. The nucleic acids were precipitated with ethanol and

† Abbreviations used: kb, 10^3 bases; bp, base-pairs.

resuspended in 20 μl of 80% (v/v) freshly deionized formamide, 0·04 M-PIPES · HCl (piperazine-*N*,*N*′-bis[2-ethane sulfonic acid], pH 6·4), 0·4 M-NaCl and 0·001 M-EDTA. The mixture was heated at 70°C for 10 min, then immediately incubated at 44°C, the temperature of maximum hybrid formation, for 3 h. The samples then were diluted with 400 μl of 5% (v/v) glycerol, 0·25 M-NaCl, 0·03 M-sodium acetate (pH 4·5), 1 mM-$ZnSO_4$ and digested for 30 min at 37°C with 20 to 30 Vogt units of S_1 nuclease (Vogt, 1973). Digestion was terminated with the addition of 100 μl of 2·5 M-ammonium acetate, 50 mM-EDTA, 10 μg of tRNA carrier, and 500 μl of isopropanol. The samples were precipitated 3 times, resuspended in 3 μl of loading buffer (80% (v/v) deionized formamide, 50 mM-Tris-borate (pH 8·3), 1 mM-EDTA, 0·1% (w/v) xylene cyanol, 0·1% (w/v) bromophenol blue: Maxam & Gilbert, 1980), vortexed vigorously, heated to 90°C for 3 min, and loaded immediately on 8% polyacrylamide denaturing gels (Sanger & Coulson, 1978). A G sequencing reaction of the same fragment was run concurrently as a size standard. The gels were autoradiographed with preflashed Kodak X-Omat AR film and an intensifying screen at −70°C.

(h) In vitro *transcription*

In vitro transcription reactions were done according to the procedure described by Hirsh & Schleif (1977), with the following modifications. The final transcription buffer was 50 mM-KCl, 10 mM-Tris · HCl (pH 8·0), 10 mM-$MgCl_2$, 1 mM-EDTA, 1 mM-dithioerythritol. The final concentrations of nucleotide triphosphates were 300 μM-ATP, 300 μM-GTP, 100 μM-UTP, 10 μM-CTP (including 0·65 μM-[α-^{32}P]CTP (760 Ci/mmol)). The template DNA concentration was 40 ng per 25-μl reaction sample. Some of these samples were digested with restriction endonucleases, *Hha*I, *Bst*EII, *Hin*fI and *Sau*3A for 30 min in the transcription buffer, then heated to 65°C to denature the enzymes before the transcription reaction. The order of addition of proteins was CRP, RNA polymerase, and *araC* protein. Transcription size standards were *araCBAD* 440 DNA, 246 bases and 47 bases, *araCBAD* 440 DNA/*Hha*I, 176 bases, *araCBAD* DNA *Bst*EII, 56 bases, and *lac*UV5 DNA, 63 and 56 bases.

(i) *DNase I footprinting*

DNase I footprinting experiments were performed according to the procedure described by Galas & Schmitz (1978) with the following modifications. DNase I digestion buffer was 20 mM-Tris · HCl (pH 7·2), 2·5 mM-$MgCl_2$, 0·3 mM-$CaCl_2$, 0·1 mM-EDTA, 50 mM-KCl and 25 μg bovine serum albumin/ml. Both the sense strand [^{32}P]*araE* 1250 bp/*Sau*3A or [^{32}P]*araE* 370 bp/*Sau*3A DNA and the antisense strand [^{32}P]*araE* 370 bp/*Hin*fI DNA were footprinted. The DNA fragments were heated to 65°C for 10 min before use. Other additions and conditions were similar to the transcription experiments. DNase I was added to a concentration of 40 to 80 ng/ml and the DNase I reaction was performed at room temperature for 30 s, then stopped by the addition of 4 M-NH_4Cl, 10 μg yeast tRNA. The samples were precipitated with ethanol 3 times, lyophilized and resuspended in 3 μl of loading buffer as described by Maxam & Gilbert (1980) and electrophoresed on 8% denaturing polyacrylamide gels (Sanger & Coulson, 1978). A G sequencing reaction of the same fragment was run concurrently as a size standard.

3. Results

(a) *Isolation and properties of plasmids containing the* araE *regulatory region*

A plasmid from the Clarke–Carbon (1976) *E. coli* library, which carries the *araE* gene (Kolodrubetz, 1980), was the source of p_E. DNA fragments from this plasmid were ligated into a plasmid in which expression of tetracycline resistance requires

insertion of an active promoter (Kosiba & Schleif, 1982) and then arabinose-dependent, tetracycline-resistant transformed cells were selected. Plasmid pRFS1595 containing a 1250 bp *araE* insert showed arabinose-dependent drug resistance at tetracycline concentrations from 4 to 10 μg/ml. Plasmid pRFS1599 containing a 370 bp *araE* insert, a subfragment of the 1250 bp fragment (Fig. 1), showed less arabinose-dependent drug resistance. These colonies grew slowly on YT tet or minimal glycerol tet plates in the absence of arabinose but when arabinose was added, the size of the colony was much larger.

(b) *Sequencing the* araE *regulatory region*

A restriction map of the *araE* 370 bp fragment and the sequencing strategy used are shown in Figure 1 and the DNA sequence of p_E and adjacent regions is shown in Figure 2.

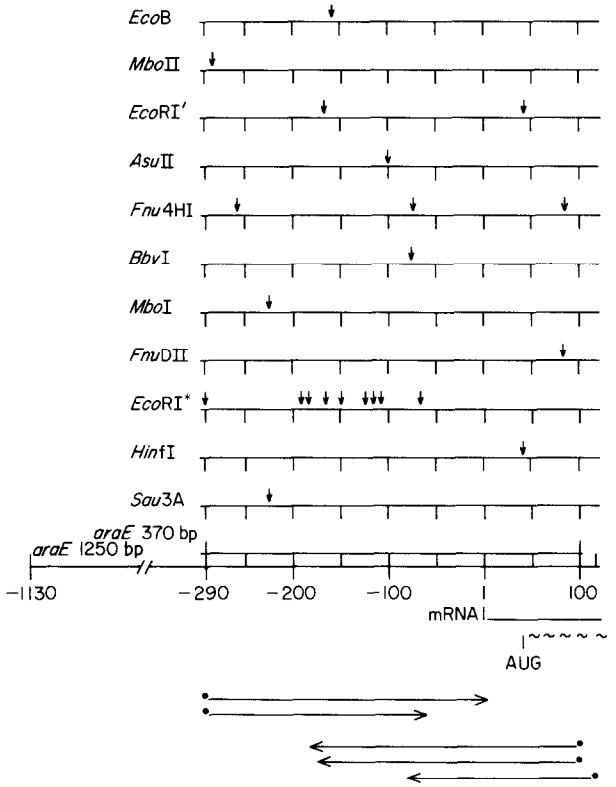

FIG. 1. Restriction map and sequencing strategy of the *araE* regulatory region. The overlapping region between the *araE* 370 bp and *araE* 1250 bp DNA fragments containing the *araE* regulatory region are indicated. Nucleotides are numbered with +1 being the start site of transcription. The DNA restriction fragments sequenced are indicated.

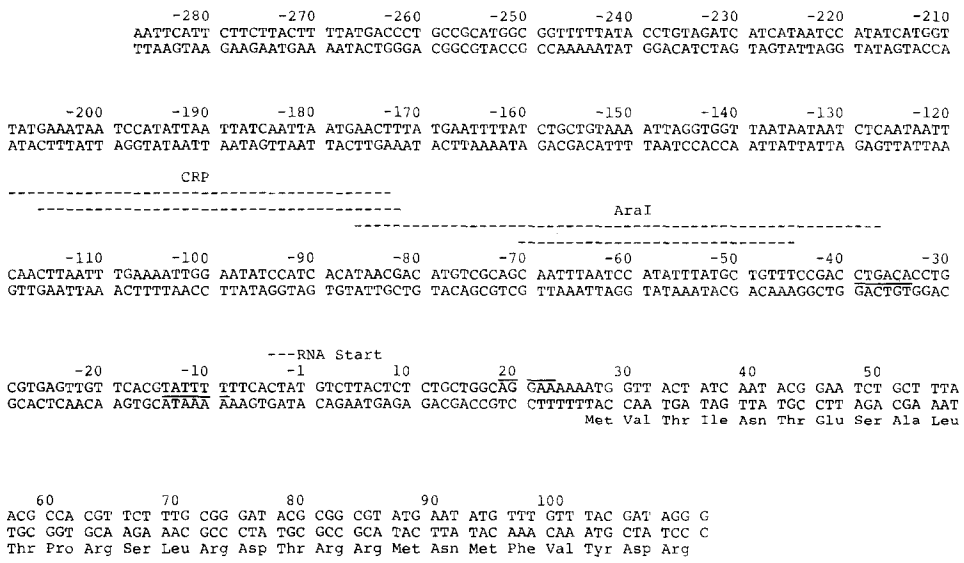

FIG. 2. Nucleotide sequence of the *araE* regulatory region and the amino acid sequence of its open reading frame. The potential Pribnow box at −10 and the −35 RNA polymerase recognition sequence are underlined. The first AUG codon for fMet and an open reading frame for 28 amino acids is drawn with the amino acid sequence below the nucleotide sequence. The potential ribosome binding site 8 bp upstream from fMet is shown with a line above it. Sequence hyphens have been omitted for clarity.

(c) *S_1 nuclease mapping of the* in vivo araE *mRNA transcription start site*

The location of the *in vivo araE* mRNA transcription start site and the direction of transcription were determined by S_1 nuclease digestion of mRNA/DNA hybrids. *araE* 1250 bp DNA was radioactively labeled at its 5′ termini and digested with restriction endonuclease *Sau*3A to give a labeled fragment that would hybridize to mRNA transcribed rightward, as shown in Figure 1. These DNA fragments were hybridized with *in vivo* RNA extracted from several different *E. coli* strains grown either in the absence or presence of arabinose for varying lengths of time or grown in the presence of anti-inducer,

FIG. 3. Hybridization of *araE* DNA to RNA from different strains grown under different conditions. End-labeled *araE* 1250 bp DNA digested with *Sau*3A was denatured and hybridized with RNA isolated from cells that were: (a) Lanes 2 to 9, 20, 21, *araE*$^+$ *araF*$^+$; (b) lane 10, Δ*cya araE*$^+$ *araF*$^+$; (c) lanes 11 to 17, *araE*$^+$ *araF*$^+$; (d) lane 18, 19, Δ*cya araE*$^+$ *araF*$^-$. Lane 1, G sequencing reaction size standard. Lane 2, RNA from cells −arabinose. Lane 3, RNA from cells +arabinose on ice with NaN_3 and chloramphenicol, control. Lane 4, RNA from cells +arabinose for 2 min. Lane 5, RNA from cells +arabinose for 6 min. Lane 6, RNA from cells +arabinose for 18 min. Lane 7, RNA from cells +arabinose for 50 min. Lane 8, RNA from cells +arabinose for 150 min. Lane 9, RNA from cells +fucose for 6 min. Lane 10, RNA from Δ*cya* cells +arabinose for 6 min. Lane 11, RNA from cells −arabinose. Lane 12, RNA from cells +arabinose for 2 min. Lane 13, RNA from cells +arabinose for 6 min. Lane 14, RNA from cells +arabinose for 18 min. Lane 15, RNA from cells +arabinose for 50 min. Lane 16, RNA from cells +arabinose for 150 min. Lane 17, RNA from cells + fucose for 6 min. Lane 18, RNA from Δ*cya* cells −arabinose. Lane 19, RNA from Δ*cya cells* +arabinose for 6 min. Lane 20, RNA alone, control. Lane 21, DNA alone, control.

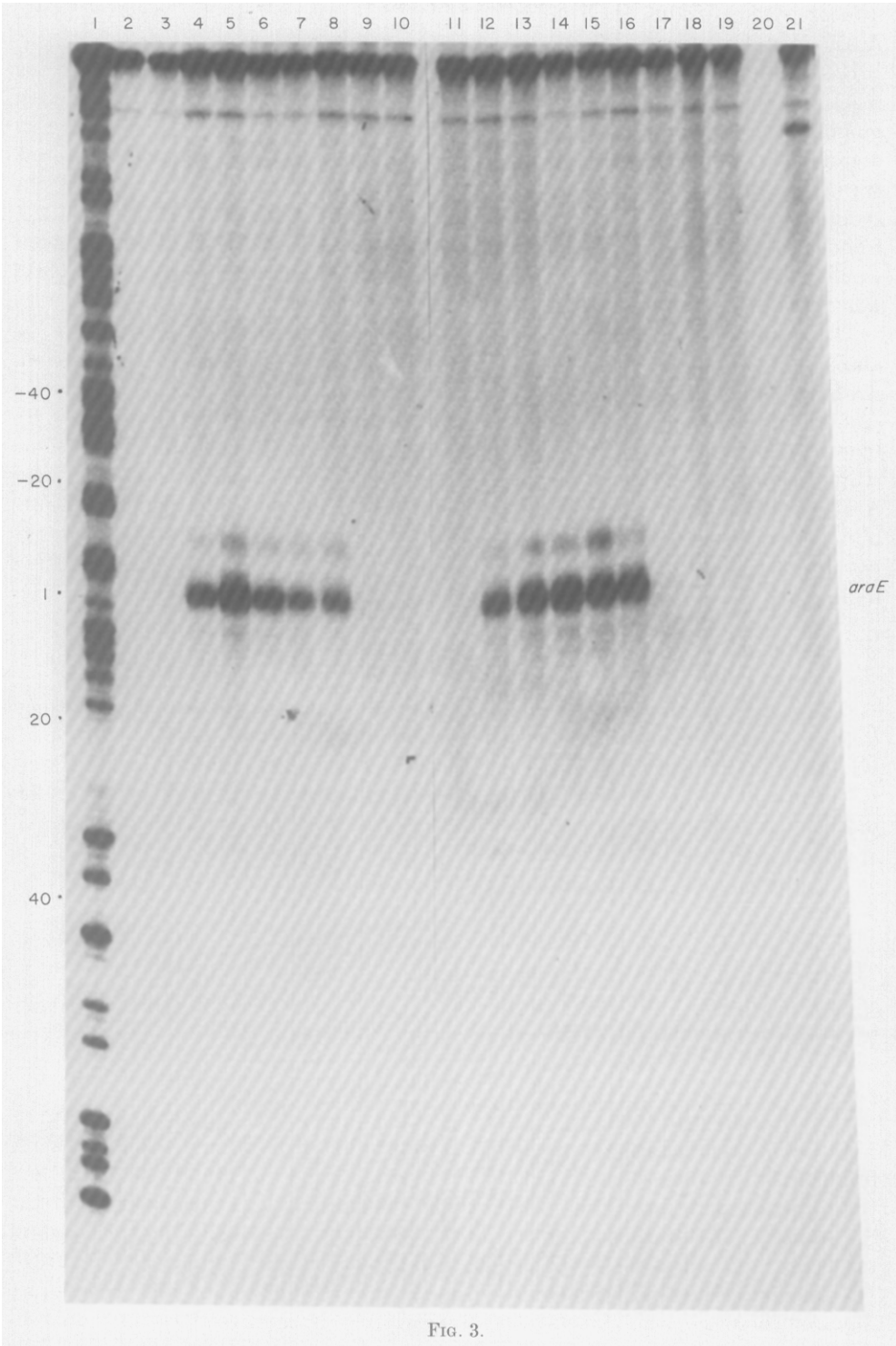

FIG. 3.

D-fucose. The strains used were $AraE^+F^+$, $AraE^+F^+\Delta Cya$, $AraE^+F^-$ and $AraE^+F^-\Delta Cya$.

Based on the mobility of the *araE* hybrid DNA relative to the G sequencing reaction (Fig. 3) of the same DNA and correcting for the 1·5 base difference in mobility between a chemically modified DNA molecule and an S_1 nuclease-digested molecule (Hentschel *et al.*, 1980), the most probable start site for *araE* transcription is the G nucleotide indicated as +1 in Figure 2. The other DNA strand, [^{32}P]*araE* 370 DNA was also digested with *Hin*fI and hybridized to RNA from cells grown in the absence or presence of arabinose. This DNA fragment would have hybridized with mRNA transcribed leftward in Figure 1. No hybrid was detected using this DNA fragment.

The results of the *araE* induction experiments are shown in Figure 3. In the absence of arabinose, the basal level of transcript was too low to be detected. Arabinose induced *araE* transcription within two minutes and the maximum amount of transcript occurred in cells grown in the presence of arabinose for six minutes. Thereafter, the level of transcription somewhat decreased. The same strain grown in the presence of anti-inducer fucose was not induced. An $AraE^+F^+\Delta Cya$ strain was not induced by the addition of arabinose. The results obtained with RNA isolated from an $AraE^+F^-$ strain were similar, although the peak of transcription occurred later, at 18 minutes, and declined more slowly than in an $AraE^+F^+$ strain.

(d) *DNase I footprinting of the* araE *regulatory region*

DNase footprinting experiments on the *araE* regulatory DNA were performed to locate the DNA binding sites for the regulatory proteins, CRP · cAMP, *araC* protein · L-arabinose, and *araC* protein · D-fucose. The results of DNase I footprinting on the sense strand of DNA are shown in Figure 4. CRP · cAMP protected from nucleotide position −83 to −119. *araC* protein in the presence of arabinose or fucose protected from −37 to −86. When CRP and *araC* · arabinose were both present, the region of protection included both sites, from −37 to −119.

DNase I footprinting of the antisense strand of *araE* DNA is shown in Figure 5. CRP · cAMP alone protected from −116 to −82; *araC* protein in the presence of arabinose or fucose protected from −75 to −34. CRP and *araC* protein in the presence of arabinose protected from −116 to −37.

(e) In vitro *transcription of the* araE *promoter*

In vitro transcription experiments using *araE* 370 bp and 1250 bp DNA fragments were done to determine the size of the *in vitro* run-off transcript and the direction of transcription. The results confirmed the *in vivo* findings of the S_1 mapping. Transcription of the *araE* 370 DNA fragment yielded a run-off transcript approximately 80 bases long, which is 35 bases shorter when *araE* 370 DNA is first cleaved with restriction endonuclease *Hin*f I but not when cleaved by *Sau*3A. Therefore, the direction of transcription is rightward, as indicated in

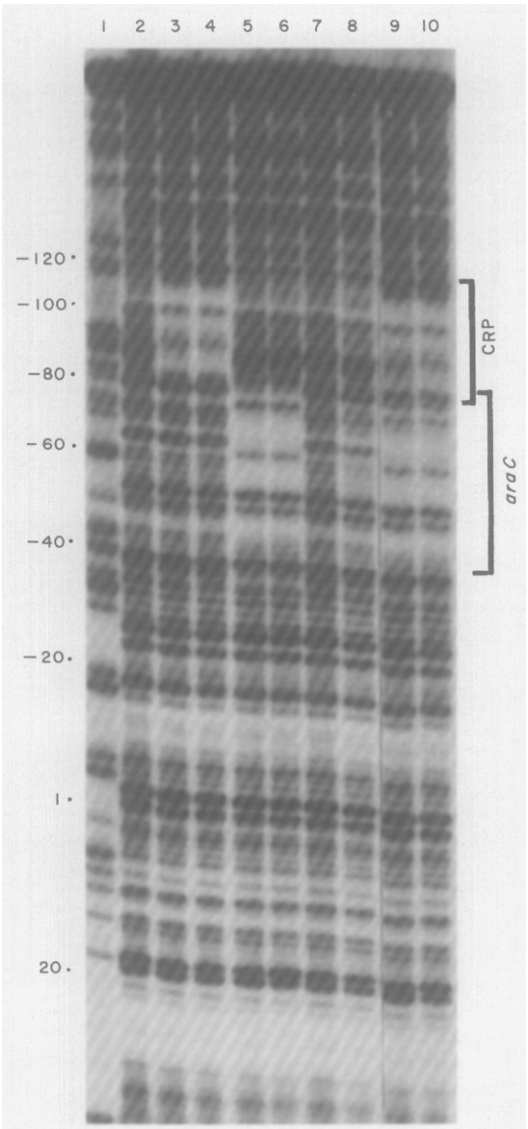

FIG. 4. DNase I footprinting of the sense strand of *araE* DNA. DNase I footprinting of the "sense" strand of *araE* DNA to determine DNA binding sites of CRP · cAMP, *araC* · arabinose, *araC* · fucose. *araE* 1250 bp DNA was end-labeled, digested with *Sau*3A, electrophoresed on a 6% polyacrylamide gel and eluted from the gel. This 343 bp fragment was used for DNase I footprinting and a G sequencing reaction. Samples in lanes 2 to 9, 11 and 12 were digested with DNase I at 40 ng/ml final concn. The sample in lane 10 was digested with DNase I at 80 ng/ml. Lane 1, G sequencing reaction size standard. Lane 2, no protein, control, DNase I = 40 ng/ml final concn. Lane 3, 44 nM-CRP · cAMP. Lane 4, 220 nM-CRP · cAMP. Lane 5, 165 nM-*araC* · arabinose. Lane 6, 660 nM-*araC* · arabinose. Lane 7, 165 nM-*araC* · fucose. Lane 8, 660 nM-*araC* · fucose. Lane 9, 44 mM-CRP · cAMP + 165 nM-*araC* · arabinose. Lane 10, 220 nM-CRP · cAMP + 165 nM-*araC* · arabinose.

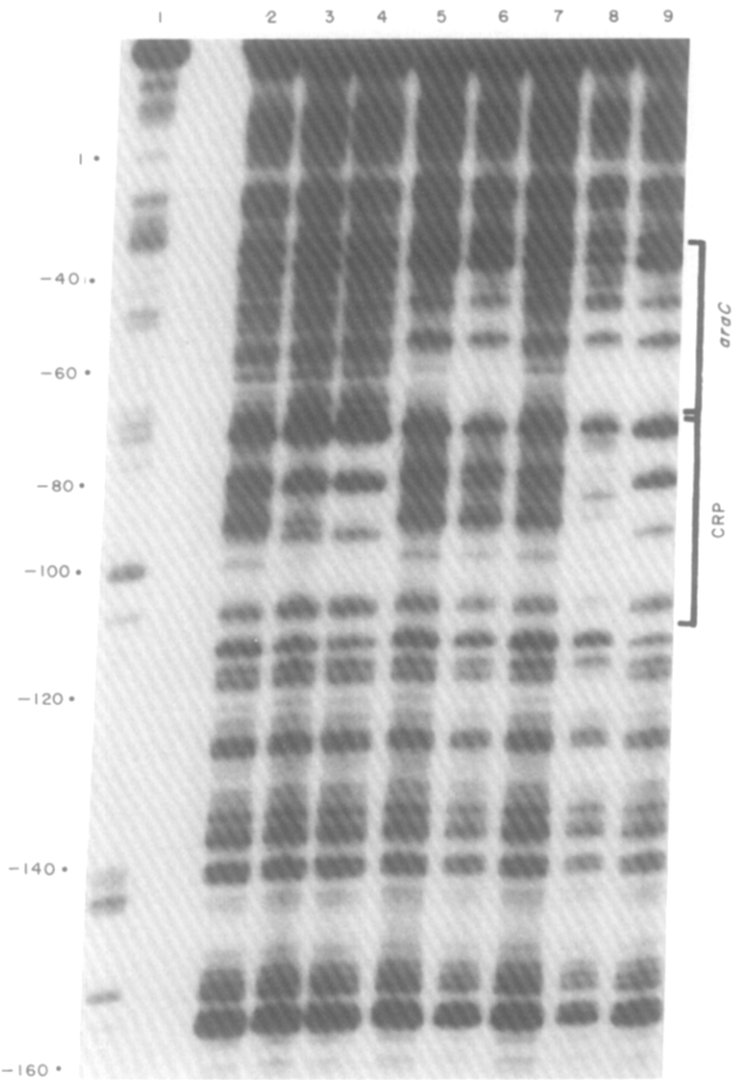

FIG. 5. DNase I footprinting of the "antisense" strand of *araE* DNA to determine DNA binding sites of CRP·cAMP, *araC*·arabinose, *araC*·fucose. *araE* 370 bp DNA was end-labeled, digested with *Hin*fI, electrophoresed on a 6% polyacrylamide gel and eluted from the gel. This 335 bp DNA fragment was used in DNase I footprinting and in a G sequencing reaction. Lane 1, G sequencing reaction, size standard. Lane 2, no protein, control, DNase I = 80 ng/ml final concn. Lane 3, 44 nM-CRP·cAMP. Lane 4, 220 nM-CRP·cAMP. Lane 5, 165 nM-*araC*·arabinose. Lane 6, 495 nM-*araC*·arabinose. Lane 7, 165 nM-*araC*·fucose. Lane 8, 495 nM-*araC*·fucose. Lane 9, 220 nM-CRP·cAMP + 495 nM-*araC*·arabinose. Samples in lanes 3 to 9 were digested with DNase I (40 ng/ml). The sample in lane 2 was digested with DNase I (80 ng/ml).

Figure 1. This was confirmed using the *araE* 1250 bp fragment as template either undigested, digested with *Hin*fI or digested with *Sau*3A restriction endonucleases for the template DNA.

In vitro experiments were also performed to determine the proteins required for transcription of *araE*. When *araC* protein and arabinose were included with RNA polymerase, transcription increased threefold above the basal level of transcription with RNA polymerase alone. When *araC* protein and D-fucose were included with RNA polymerase, the basal level of transcription was essentially unchanged. The addition of CRP to RNA polymerase did not enhance transcription in the absence of *araC* protein and arabinose, but did stimulate the transcription fivefold above the basal level in the presence of *araC* protein, arabinose, and RNA polymerase.

4. Discussion

Study of the promoter for the *araE* gene was facilitated by its cloning. DNA from a plasmid known to carry the intact *araE* gene was the source of the fragment containing the p_E promoter, and a plasmid permitting direct selection of recombinants containing inserted fragments with promoter activity revealed arabinose-dependent, tetracycline-resistant transformants.

RNA extracted from cells, hybridized to radioactive DNA, and digested with the single-strand specific nuclease S_1 permitted high-resolution mapping of the transcription start location and determination of the direction of transcription. The start site has been determined to within about ± 2 base-pairs, a level of precision adequate for interpretation of the other experiments presented in this paper.

The DNA sequence of the site to which RNA polymerase must bind in order to produce the transcripts observed by S_1 mapping or by *in vitro* transcription possesses high homology to consensus RNA polymerase binding sequences (Siebenlist *et al.*, 1980), both in the -10 region (4 out of 6 bp) and the -35 region (5 out of 6 bp). Therefore, it is surprising that this promoter requires *araC* protein and CRP for significant activity. These auxiliary proteins may compensate for the greater than usual distance between the -10 and -35 regions, 19 bp as compared to the 17 bp observed in most promoters (von Hippel & Bear, 1982).

The CRP binding sequence agrees well with the proposed CRP consensus sequences (Ebright, 1982; Queen & Rosenberg, 1981). The sequence of the *araC* protein binding site in *araE* shows homology to the *araC* binding sites in the *araCBAD* regulatory region and will be further discussed in a later paper. After the established genetic nomenclature on p_{BAD} (Englesberg *et al.*, 1965,1969; Gielow *et al.*, 1971) and its biological identification on p_{BAD} (Ogden *et al.*, 1980) we name the *araC* binding site on p_E the $araI_E$ site.

A portion of the DNA sequenced along with p_E is transcribed *in vivo* and could code for the amino terminus of the *araE* transport protein. The methionine codon at position $+27$ is preceded by a typical ribosome binding site (Shine & Dalgarno, 1974), indicated in Figure 2, and an open reading frame continues for 28 amino acid residues to the end of the sequenced fragment. The encoded amino acid

sequence possesses the hydrophobic and charge distributions characteristic of membrane proteins (Yamada *et al.*, 1982) as would be expected for a protein involved in active transport.

The DNase footprinting showed that the cyclic AMP receptor protein and *araC* protein bind adjacent to but upstream from the site to which RNA polymerase initiates transcription. Since these positions are the same as those found in p_{BAD} (Ogden *et al.*, 1980; Miyada *et al.*, 1980; Lee *et al.*, 1981), both promoters likely are induced by the same mechanism. As on p_{BAD}, CRP is not adjacent to RNA polymerase and therefore must exert its stimulatory action indirectly on RNA polymerase through the intervening DNA, *via* the intervening *araC* protein, or directly on RNA polymerase by means of rather tightly bent DNA so the two proteins may contact one another.

The apparent absence of a site analogous to the *araO* site on p_{BAD} raises questions concerning repression of p_E. This promoter does appear to be repressible, since the C^+ allele is dominant to the C^c allele (Kolodrubetz & Schleif, 1981). However, on p_{BAD} the *araO* site appears to be involved with repression (Englesberg *et al.*, 1969; Miyada *et al.*, 1980). Possibly this means that repression on p_E is generated by *araC* binding to $araI_E$.

We thank Anne Favreau for cloning p_E and William Hendrickson for performing the Southern transfer experiment for locating p_E and for comments on the manuscript. This work was supported by United States Public Health Service research grant GM-18277. The paper is publication no. 1460 from the Department of Biochemistry, Brandeis University, Waltham, Mass. 02254, U.S.A.

REFERENCES

Berk, A. J. & Sharp, P. A. (1977). *Cell*, **12**, 721–732.
Brickman, E., Soll, L. & Beckwith, J. (1973). *J. Bacteriol.* **116**, 582–587.
Clarke, L. & Carbon, J. (1976). *Cell*, **9**, 91–99.
Ebright, R. (1982). In *Molecular Structure and Biological Function* (Griffen, J. & Duax, W., eds), Elsevier North-Holland, N.Y. In the press.
Eilen, E., Pampeno, C. & Krakow, J. S. (1978). *Biochemistry*, **17**, 2469–2473.
Englesberg, E., Irr, L., Power, J. & Lee, N. (1965). *J. Bacteriol.* **90**, 946–957.
Englesberg, E., Sheppard, D., Squires, C. & Meronk, F. (1969). *J. Mol. Biol.* **43**, 281–298.
Galas, D. & Schmitz, A. (1978). *Nucl. Acids Res.* **5**, 3157–3170.
Gielow, L., Largen, N. & Englesberg, E. (1971). *Genetics*, **69**, 289–302.
Greenblatt, J. & Schleif, R. (1971). *Nature New Biol.* **233**, 166–170.
Hentschel, C., Irminger, J., Bucher, P. & Birnstiel, M. L. (1980). *Nature (London)*, **285**, 147–151.
Hirsh, J. & Schleif, R. (1977). *Cell*, **11**, 545–566.
Kolodrubetz, D. (1980). Ph.D. thesis, Brandeis University.
Kolodrubetz, D. & Schleif, R. (1981). *J. Mol. Biol.* **151**, 215–227.
Kosiba, B. E. & Schleif, R. (1982). *J. Mol. Biol.* **156**, 53–66.
Lee, N., Wilcox, G., Gielow, W., Arnold, J., Cleary, P. & Englesberg, E. (1974). *Proc. Nat. Acad. Sci., U.S.A.* **71**, 634–638.
Lee., N. L., Gielow, W. O. & Wallace, R. G. (1981). *Proc. Nat. Acad. Sci., U.S.A.* **78**, 752–756.
Majors, J. (1975). *Proc. Nat. Acad. Sci., U.S.A.* **72**, 4396–4398.
Maxam, A. M. & Gilbert, W. (1980). In *Methods in Enzymology* (Grossman, L. & Moldave, K., eds). vol. 65, pp. 499–560, Academic Press, New York.

Miyada, C., Horowitz, A., Cass, L., Timko, J. & Wilcox, G. (1980). *Nucl. Acids Res.* **8**, 5267–5274.

Ogden, S., Haggerty, D., Stoner, C. M., Kolodrubetz, D. & Schleif, R. (1980). *Proc. Nat. Acad. Sci., U.S.A.* **77**, 3346–3350.

Queen, C. & Rosenberg, M. (1981). *Nucl. Acids Res.* **9**, 3365–3377.

Salser, W., Gesteland, R. F. & Bolle, A. (1967). *Nature (London)*, **215**, 588–591.

Sanger, F. & Coulson, A. R. (1978). *FEBS Letters*, **87**, 107–110.

Schleif, R. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3479–3484.

Schleif, R. & Favreau, A. (1982). *Biochemistry*, **21**, 778–789.

Schleif, R. & Wensink, P. C. (1981). *Practical Methods in Molecular Biology*, Springer-Verlag New York Inc., New York.

Shine, J. & Dalgarno, L. (1974). *Proc. Nat. Acad. Sci., U.S.A.* **71**, 1342–1346.

Siebenlist, U., Simpson, R. B. & Gilbert, W. (1980). *Cell*, **20**, 269–281.

Smith, B. R. & Schleif, R. (1978). *J. Biol. Chem.* **253**, 6931–6933.

Steffen, D. (1976). Ph.D. thesis, Brandeis University.

Steffen, D. & Schleif, R. (1977). *Mol. Gen. Genet.* **157**, 333–339.

Vogt, V. M. (1973). *Eur. J. Biochem.* **33**, 192–200.

von Hippel, P. & Bear, D. (1982). In *Proc. Symp. Promoters: Structure and Function.* In the press.

Yamada, M., Ebina, Y., Miyata, T., Nakazawa, T. & Nakazawa, A. (1982). *Proc. Nat. Acad. Sci., U.S.A.* **79**, 2827–2831.

Edited by M. Gottesman