

# Nucleotide sequence of the *argR* gene of *Escherichia coli* K-12 and isolation of its product, the arginine repressor

(repressor protein/regulatory gene/arginine regulon)

DONGBIN LIM\*, JOEL D. OPPENHEIM\*, THOMAS ECKHARDT†, AND WERNER K. MAAS\*‡

\*Department of Microbiology, New York University Medical Center, New York, NY 10016; and †Smith Kline & French Laboratories, P.O. Box 7929, Philadelphia, PA 19101

Communicated by Bernard D. Davis, June 15, 1987 (received for review April 6, 1987)

**ABSTRACT** In *Escherichia coli*, the arginine repressor, the product of the *argR* gene, in conjunction with L-arginine controls the synthesis of the enzymes of arginine biosynthesis. We describe the nucleotide sequence of the *argR* gene, including its control region, and show that formation of the repressor is autoregulated. The *argR* control region contains two promoters, one of which overlaps the operator site and, as with other *arg* genes, consists of two adjacent palindromic sequences ("ARG boxes"). The arginine repressor protein and an arginine repressor- $\beta$ -galactosidase fusion protein were purified, and the amino acid sequence of the N-terminal end of the repressor protein portion of the fusion protein was determined. Antibodies prepared against the fusion protein react with the repressor. The repressor is precipitable by L-arginine, which facilitates its purification. The native repressor is a hexamer with a molecular weight of 98,000; its monomeric subunit has a molecular weight of 16,500. To verify its properties postulated from genetic studies, we show that in the presence of L-arginine, repressor inhibits transcription of *argF* and binds to the ARG boxes of *argF* and *argR*.

The control of enzyme formation in the pathway of arginine biosynthesis in *Escherichia coli* by the end product has been studied extensively (for a review see ref. 1). The formation of the eight enzymes of arginine biosynthesis is controlled in conjunction with arginine by the arginine repressor (ArgR), the product of the regulatory gene *argR* (2, 3). Control was shown to be negative (4, 5), and the term regulon was coined (5) to describe systems where a single repressor controls several unlinked genes. A DNA segment containing the *argR* gene has been cloned (6). In sequencing studies of the genes for the biosynthetic enzymes, an 18-base-pair-long palindromic region, referred to as the "ARG box," was found in the promoter region of the genes. From mutant studies it is thought to be the binding site for ArgR (1). In all of the genes there are two adjacent ARG boxes.

In the present paper we describe the nucleotide sequence of the *argR* gene, including its control region. By putting the *argR* gene and an *argR-lacZ* gene fusion under the control of strong promoters, clones were obtained that hyperproduce ArgR and an ArgR- $\beta$ -galactosidase fusion protein. Both types of protein were purified to homogeneity. ArgR could be precipitated from crude extracts by L-arginine, which facilitated its purification. Besides this unusual behavior, ArgR has other properties that distinguish it from known repressors: unlike the tryptophan repressor and the methionine repressor, both of which are dimers (7-9), it is a hexamer. Unlike the tryptophan repressor, which binds to a single palindromic sequence (10), it binds to two adjacent ARG boxes. These differences lend significance to ArgR for studies of repressor-target interactions.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

## MATERIALS AND METHODS

**Strains, Plasmids, and Materials.** The plasmids used are described in Table 1. The host strains for plasmids were: *E. coli* N5151, a biotin-requiring heat-inducible  $\lambda$  lysogen derived from strain SA500 (13) and obtained from A. Shatzman (15); *E. coli* JM101 with the genotype  $\Delta lac pro, supE thi/F'trad36, proAB^+, lacI^qZ\Delta M15$ , obtained from Pharmacia; *E. coli* EC146 ( $\lambda AZ-7$ ), a prototrophic  $\lambda$  lysogen used for scoring ArgR phenotypes, with the genotype *argD, argR, argA::lacZ* (6). Restriction endonucleases, BAL-31 nuclease, Moloney murine leukemia virus reverse transcriptase, DNA polymerase Klenow fragment, RNA polymerase, radiochemicals, and isopropyl  $\beta$ -D-thiogalactoside were from commercial sources. Oligonucleotides were synthesized by B. Goldschmidt (Department of Biochemistry, New York University School of Medicine).

**Plasmid Constructions.** The *argR* gene was subcloned from pEC101 (6) as an 880-bp *Sph* I-*Acc* I fragment into a polylinker fragment in a derivative of pUC9 (14). From the resulting plasmid, pER3, a 1.2-kilobase (kb) *Sph* I-*Sal* I fragment containing *argR* can be recovered. Plasmid pDB101 was constructed by combining a 1.7-kb *Eco*RI-*Hpa* I fragment of pKC30 with a 360-bp flush-ended *Sph* I-*Pvu* II fragment of pER3 and with a 6.2-kb filled-in *Bgl* II-*Eco*RI fragment of pORF2. Plasmid pDB102 was constructed by joining a 400-bp filled-in *Bam*HI-*Eco*RI fragment from pDR540 to the same *Sph* I-*Pvu* II and *Bgl* II-*Eco*RI fragments as in pDB101. Plasmid pDB100 was constructed by joining a 1.2-kb flush-ended *Sph* I-*Sal* I fragment of pER3 to a 4.1-kb fragment of pDR540, which has a filled-in *Bam*HI site and a cohesive *Sal* I end. The series of deletions derived from pDB100 (pDB122 to pDB190) were obtained by digestion with BAL-31 nuclease starting from the *Bam*HI site. After subsequent digestion with *Sal* I, the deleted fragments were isolated by agarose gel electrophoresis and ligated to a filled-in *Bam*HI-*Sal* I fragment of pDR540. Plasmid pDB200 was constructed by deleting 2 bp in the unique *Eco*RV site of pER3 (Fig. 1). To do this, pER3 was linearized with *Eco*RV and treated with 5 units of Klenow fragment in the presence of 250  $\mu$ M dGTP. After phenol extraction and ethanol precipitation, the fragment was self-ligated and transformed into strain EC146 ( $\lambda AZ-7$ ). The deletions were verified by *Sau*3A digestion, since the 2-bp deletion converts the *Eco*RV site to a *Sau*3A site.

**DNase Footprinting.** Plasmid pER3 was linearized with *Spe* I (Fig. 1) and labeled with [ $\alpha$ - $^{32}$ P]dCTP by treatment with Klenow fragment. After digestion with *Xho* I, a 190-bp fragment, labeled at one end, was isolated from a polyacrylamide gel. About 0.2 pmol of this fragment was incubated for 20 min at 37°C with various concentrations of ArgR with or without 5 mM L-arginine. The incubation mixture contained, in a total volume of 20  $\mu$ l, 0.1 M KCl, 10 mM MgCl<sub>2</sub>, 1 mM

‡To whom reprint requests should be addressed.

Table 1. Description of plasmids

Plasmid	Relevant phenotype	Characteristics, source, and/or ref.
pKC30	Ap <sup>R</sup>	Expression vector carrying <i>P<sub>L</sub></i> promoter (M. Rosenberg; ref. 15)
pDR540	Ap <sup>R</sup>	Expression vector carrying <i>tac</i> promoter (26)
pORF2	Ap <sup>R</sup>	Open reading frame expression vector for LacZ fusion proteins (11)
pMC20	Ap <sup>R</sup> ArgF <sup>+</sup>	pBR322 derivative carrying <i>argF</i> (12)
pER3	Ap <sup>R</sup> ArgR <sup>+</sup>	pUC9 derivative carrying <i>argR</i> (this work)
pDB100	Ap <sup>R</sup> ArgR <sup>+</sup>	pDR540 derivative carrying <i>argR</i> (this work)
pDB101	Ap <sup>R</sup> LacZ <sup>+</sup>	pORF2 derivative carrying <i>P<sub>L</sub></i> promoter and <i>argR-lacZ</i> fusion (this work)
pDB102	Ap <sup>R</sup> LacZ <sup>+</sup>	pORF2 derivative carrying <i>tac</i> promoter and <i>argR-lacZ</i> fusion (this work)
pDB200	Ap <sup>R</sup> ArgR <sup>-</sup>	Two-base pair deletion of pER3 at <i>EcoRV</i> site (this work)

Ap<sup>R</sup>, resistance to ampicillin; ArgF<sup>+</sup>, production of ornithine carbamoyltransferase.

dithiothreitol, 5% (vol/vol) glycerol, 20 mM Tris-HCl (pH 7.6), and 1 μg of pUC18 DNA. Following incubation, 66 ng of pancreatic DNase was added, and the reaction was stopped after 60 sec at 37°C by phenol extraction. After ethanol precipitation, the reaction products were analyzed in an 8% sequencing gel, using a G+A Maxam-Gilbert sequencing ladder of the same 190-bp fragment as a guide.

**In Vitro Run-off Transcription.** This was done with a 280-bp *HindIII-Pst I* fragment of pMC20, containing the promoter-operator region of *argF*. In a final volume of 20 μl, ≈0.05 pmol of this fragment were incubated at 37°C for 15 min with various amounts of ArgR in the presence or absence of 5 mM L-arginine. The reaction mixture contained 20 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 5% (vol/vol) glycerol, 1 mM dithiothreitol, and 0.1 M KCl. RNA polymerase holoenzyme (5 pmol) was added for 5 min, and then the reaction was started by the addition of 150 μM ATP/200 μM CTP/116 μM UTP/1.2 μM [ $\alpha$ -<sup>32</sup>P]GTP. After a 2.5-min incubation period, the GTP concentration was increased to 40 μM by adding unlabeled GTP, and the incubation was continued for 3 min. After phenol extraction and ethanol precipitation the reaction products were analyzed in an 8% sequencing gel.

**Primer Extension.** Total RNA was extracted from *E. coli* strains EC146 ( $\lambda$ AZ-7), EC146 ( $\lambda$ AZ-7/pER3), and EC146 ( $\lambda$ AZ-7/pDB200) as described by Aiba *et al.* (16). Primer extension was performed with the synthetic nucleotide 5'

CCCTGGGAGCTAAATTT 3', as described by Curtis *et al.* (17). The reaction products were analyzed in an 8% sequencing gel, together with the dideoxy chain-termination products obtained from the same primer.

**Purification of the ArgR- $\beta$ -Galactosidase Fusion Protein.** The ArgR- $\beta$ -galactosidase fusion protein from a heat-induced culture of strain N5151 (pDB101) was purified by affinity chromatography using *p*-aminophenyl  $\beta$ -D-thiogalactoside (TPEG; Sigma) CH-Sepharose-4B beads as described by Ullmann (18). To achieve complete purification of the fusion protein two additional chromatography steps were found necessary: anion-exchange chromatography using a MonoQ column and molecular sieve chromatography using a Superose 6 column, both attached to a Pharmacia FPLC System (Pharmacia).

**Purification of ArgR (Table 2).** An extract of strain JM101 (pDB169) was prepared by sonication of a cell suspension in Arg buffer (20 mM Tris-HCl/10 mM MgCl<sub>2</sub>/10 mM 2-mercaptoethanol, pH 7.5). The supernatant from the extract obtained by centrifugation at 40,000 × *g* for 30 min at 4°C was treated with DNase (1 μg/ml of extract) for 30 min at 37°C. After DNase treatment the extract was centrifuged at 40,000 × *g* for 30 min at 4°C, and the supernatant was saved (hereafter referred to as the DNase-treated extract). Solid ammonium sulfate was added to this supernatant to 55% of saturation. After mixing for 1 hr at 4°C, the precipitate was harvested, resuspended in Arg buffer, and dialyzed against this buffer. The final dialyzed product was Millipore filtered (hereafter referred to as the 55% (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> precipitate). Solid L-arginine was added to a final concentration of 5 mM. After incubation for 30 min at 37°C, the precipitate was harvested by centrifugation at 5000 × *g* for 5 min at 25°C. It was resolubilized by gently mixing in Arg buffer containing 1 M NaCl. The solution was dialyzed first against 1 M NaCl/Arg buffer and then against Arg buffer. After centrifugation to remove any insoluble material, this fraction (hereafter referred to as the Arg precipitated fraction) was then chromatographed consecutively on preparative anion- and cation-exchange columns (MonoQ and MonoS HR 10/10 columns, respectively, attached to a Pharmacia FPLC System) using 0–0.5 M NaCl gradients in the Arg buffer for column development. Purification of ArgR was continuously monitored at each step by NaDodSO<sub>4</sub>/PAGE.

**Assays of Proteins.** ArgR levels were assayed by a quantitative rocket immunoelectrophoresis assay (19) using antibody directed against the purified ArgR- $\beta$ -galactosidase fusion protein. Antibody was elicited in rabbits according to a described protocol (19). The immunoglobulin fraction was prepared from immune sera according to Harboe and Ingild (20). Molecular weights were determined by three methods. (i) The subunit value was obtained by NaDodSO<sub>4</sub>/PAGE analysis (12.5% separating gel) of reduced samples and

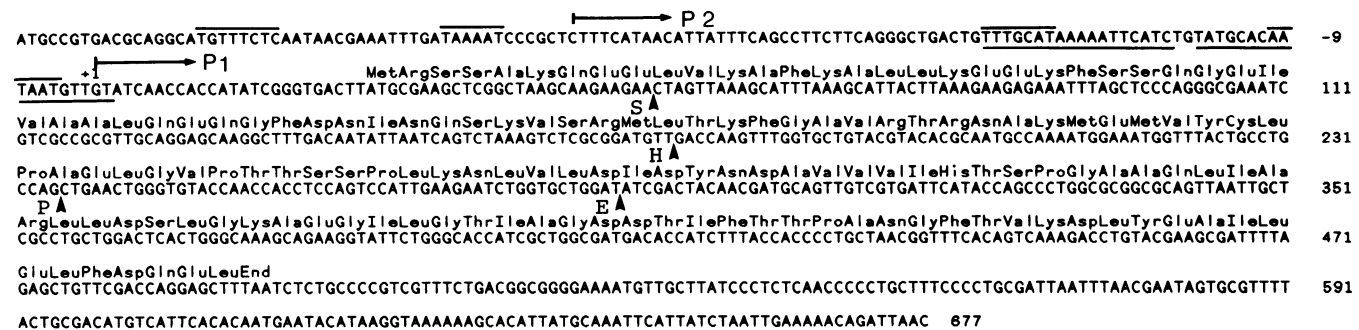


FIG. 1. Nucleotide sequence and deduced amino acid sequence of *argR*. Two transcription start sites are denoted by arrows. The sequence is numbered from the transcription start point of the proximal promoter *P1*. Putative -10 and -35 regions of each promoter are overlined. Two ARG boxes are underlined. The -10 and -35 regions of *P1* overlap with the ARG boxes. The sequence starts from the *Sph I* site of the *Sph I-Acc I* fragment. Arrowheads indicated sites for the following restriction enzymes: S, *Spe I*; H, *HincII*; P, *Pvu II*; E, *EcoRV*.

Table 2. Purification of ArgR

Sample	Volume, ml	Protein		ArgR		
		mg/ml	Total mg	mg/ml	Total mg	S.A.
Sonic extract	140	50	7000	6.95	975	0.14
DNase-treated extract	138	47.5	6555	7.15	980	0.15
55% (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> precipitate	50	108	5400	18	900	0.17
Arg-precipitated fraction	45	18.5	980	16.4	740	0.75
Mono Q peak	42	15	630	14.2	596	0.95
Mono S peak	50	11.25	560	11.25	560	1.0

Protein concentration was determined by the Bradford procedure (Bio-Rad). ArgR was measured by quantitative rocket immunoelectrophoresis. S.A., specific activity.

appropriate standards, using the procedure of Laemmli (21). (ii) The native molecular weights were determined by nondenaturing gradient gel electrophoresis using the Pharmacia PhastGel system and by molecular-sieve chromatography in a Superose 12 HR 10/30 column operated with a Pharmacia FPLC system. In the latter procedure it was necessary to elute ArgR in the presence of at least 0.15 M NaCl to avoid interaction of the purified protein with the Superose beads. Total protein was determined by measuring light absorption at 280 nm, combined with dry-weight determinations, or by the Bradford procedure (Bio-Rad).  $\beta$ -Galactosidase was measured as described by Miller (22). The amino acid sequencing was carried out by D. Schlesinger of the Kaplan Cancer Center at New York University Medical Center.

## RESULTS

**Nucleotide Sequence and Autoregulation of *argR*.** Restriction enzyme mapping of plasmid pER3 indicated the entire *argR* gene is encoded in an 880-bp *Sph*I-*Acc*I fragment. This fragment was sequenced by the method of Maxam and Gilbert (23). Details of the sequencing strategy are available from the authors. The DNA sequence for 806 bp starting at the *Sph*I site is shown in Fig. 1. It contains one major open reading frame, spanning *Hinc*II, *Pvu*II, and *Eco*RV sites and coding for a protein of 156 amino acids.

Inspection of the region upstream of the putative *argR* coding region revealed two ARG boxes, which suggested that the *argR* gene is autoregulated. To verify this supposition, *argR*-*lacZ* fusions were constructed *in vivo* (24). It was found in these strains that in the presence of arginine and the *argR*<sup>+</sup> gene, but not with either one alone,  $\beta$ -galactosidase formation is inhibited (data not shown).

To study the autoregulation of *argR* further, we analyzed the transcriptional products of *argR* by the primer extension method. Two plasmids were used: pER3, which contains intact *argR*, and the derivative pDB200, which has a 2-bp deletion in the *Eco*RV site of *argR* (Fig. 1) and does not produce active repressor. About 300 ng of a synthetic 17-nucleotide oligomer, complementary to nucleotides at positions 88-104 in Fig. 1, was annealed with total RNA extracted from *E. coli* EC146 ( $\lambda$ AZ-7/pER3), *E. coli* EC146 ( $\lambda$ AZ-7/pDB200), and *E. coli* EC146 ( $\lambda$ AZ-7) and extended with reverse transcriptase. As shown in Fig. 2, there are two transcriptional start sites and presumably two promoters, *P*<sub>1</sub> and *P*<sub>2</sub>, in *argR*. In the proximal promoter, *P*<sub>1</sub>, transcription starts from a guanosine in position +1. The -10 and -35 regions of the *P*<sub>1</sub> promoter are entirely within the ARG boxes (Fig. 2). Since ArgR binds to these ARG boxes (see below), it seems likely that *P*<sub>1</sub> is controlled by ArgR. This is confirmed by a comparison of the *in vivo* transcriptions of

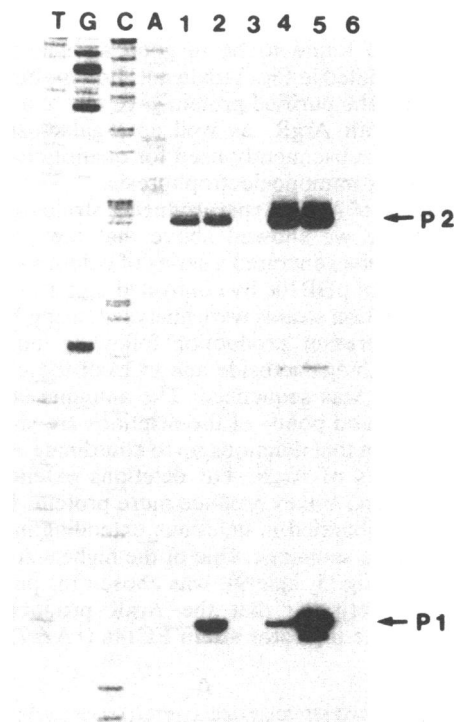


FIG. 2. Primer extension of *argR*. Total RNA (10  $\mu$ g) was annealed with 300 ng of synthetic oligonucleotide 5' CCCTGG-GAGCTAAATTT 3' and extended by reverse transcriptase in the presence of [<sup>32</sup>P]dCTP. The reaction products were loaded onto an 8% sequencing gel. The same primer was used for the dideoxy chain-termination, to serve as a marker for the size of the extended fragments. A, C, G, and T denote the dideoxy reaction products. Lanes: 1 and 4, primer extension of RNA from *E. coli* EC146 ( $\lambda$ AZ-7/pER3); 2 and 5, from *E. coli* EC146 ( $\lambda$ AZ-7/pDB200); 3 and 6, from *E. coli* EC146 ( $\lambda$ AZ-7). In lanes 4, 5, and 6, five times as much sample was applied. In lane 6, faint transcripts from the chromosomal *argR* gene are seen in the corresponding positions. *P*<sub>1</sub> and *P*<sub>2</sub> are proximal and distal promoters, respectively.

*argR* in the intact pER3 plasmid and the deletion mutant pDB200 plasmid (Fig. 2, compare lanes 1 and 4 with lanes 2 and 5). The transcription from *P*<sub>1</sub> in the presence of arginine is about 15 times higher in pDB200 than in pER3.

The distal promoter *P*<sub>2</sub> is located upstream from the ARG boxes and transcription starts from cytidine in position -76. The transcriptional activity of this promoter in the presence of arginine is the same in pER3 and in pDB200. Therefore, this promoter is not regulated by ArgR. In pER3, under conditions of repression, *P*<sub>2</sub> is the major promoter and the ratio of activities of repressed *P*<sub>1</sub>/*P*<sub>2</sub> is about 1:7. Under derepressed conditions, *P*<sub>1</sub> is the major promoter and the ratio of activities of *P*<sub>1</sub>/*P*<sub>2</sub> is  $\approx$ 2.5:1 (see Fig. 2).

**Hyperproduction and Isolation of an ArgR- $\beta$ -Galactosidase Fusion Protein.** Since initially it was difficult to do quantitative assays for ArgR, we constructed a chimeric gene in which part of *argR* is joined to *lacZ*, coding for a fusion protein containing the N-terminal portion of ArgR fused to  $\beta$ -galactosidase. The chimeric gene containing the *argR* leader sequence was put under the control of either the *P*<sub>L</sub> promoter or the *tac* promoter.

The fusion protein was extracted and purified from strain N5151 (pDB101), containing the *P*<sub>L</sub> promoter. Cells, induced by growing at 42°C for 4 hr, were lysed by sonication, and the protein was subsequently purified. The final product was considered to be pure because after NaDodSO<sub>4</sub>/PAGE a single band was visible, and because the preparation had a specific enzyme activity similar to that of pure  $\beta$ -galacto-

sidase. The sequence of 14 amino acids at the N terminus was determined and found to be in good agreement with the sequence postulated in Fig. 1 (data not shown). Immunization of rabbits with the purified protein gave rise to an antiserum that reacted with ArgR, as well as  $\beta$ -galactosidase. This antiserum was subsequently used for quantitative assays of ArgR by rocket immunoelectrophoresis.

**Construction of ArgR Hyperproducing Strains and Isolation of ArgR.** Since we showed above that the *argR* gene is autoregulated, we generated a series of deletions in the *argR* control region of pDB100 by controlled digestion with BAL-31. Ninety deletion strains were analyzed, using NaDodSO<sub>4</sub>/PAGE, for repressor production following induction with isopropyl  $\beta$ -D-thiogalactoside and in 14 of these the control region of *argR* was sequenced. The amounts of ArgR produced and the end points of the deletions are shown in Fig. 3. It can be seen that deletions up to coordinate -81 produce similar amounts of ArgR, but deletions extending further through the ARG boxes produce more protein. Even higher expression is obtained in deletions extending into the *argR* Shine-Dalgarno sequence. One of the highest ArgR producers, pDB169 (Fig. 3, lane 9), was chosen for purification of ArgR, after verifying that the ArgR produced by it is functional in the indicator strain EC146 ( $\lambda$ AZ-7) (6).



FIG. 3. Construction of hyperproducing strains following BAL-31 deletion of pDB100. BAL-31 deletion was carried out to remove the *argR* control region. The *tac*-promoter cassette (*Eco*RI-*Bam*HI fragment of pDR540) was placed on the upstream side of the deleted *argR* region. Sixteen of the deleted strains are shown in B and in 14 of these the deletion end points were determined by sequencing appropriate restriction fragments cloned in M13mp9 by the method of Sanger *et al.* (25). (A) BAL-31 deletion end points. Arrows indicate the deletion end points. The numbers correspond to the numbers in B. The ARG boxes and putative Shine-Dalgarno sequence of *argR* are underlined. (B) NaDodSO<sub>4</sub>/PAGE of deletion clones. Each deletion clone was treated as described in the legend of Fig. 3 except a 15% gel was used. The lane numbers correspond to the numbers in A. pDB169 in lane 9 was used for ArgR purification. Arrow indicates ArgR.

Initially the purification of ArgR was carried out in a similar way as for the ArgR- $\beta$ -galactosidase fusion protein. However, during purification we noted that a precipitate was formed when L-arginine was added, which could be redissolved in Arg buffer containing at least 0.1 M NaCl. NaDodSO<sub>4</sub>/PAGE of the redissolved precipitate showed that it had the mobility expected of the ArgR. Furthermore, the precipitation was specific for L-arginine. We assumed that the precipitate was ArgR and developed a procedure for its purification. Results of a typical experiment are shown in Table 2. It can be seen that ArgR constitutes  $\approx$ 14% of the initially extracted protein. The yield of pure ArgR is 58%, with the greatest increase in the purification being due to the precipitation by L-arginine. The amount of ArgR at each step of purification was determined by quantitative rocket immunoelectrophoresis using as standard the material recovered from the final cation-exchange column. This material gives a single band in polyacrylamide gels, and we assume that it is pure.

Some of the properties of the purified ArgR have been studied. The molecular weight of the monomer determined by NaDodSO<sub>4</sub>/PAGE analysis was found to be 16,500, in good agreement with that predicted from the DNA sequence. The molecular weight of the native protein was found to be 98,000, indicating that ArgR exists normally as a hexamer. The native protein is soluble in dilute buffer (Arg buffer), but insoluble in distilled water. The protein is precipitated completely in 0.8–1.0 mM L-arginine. This effect is very specific. The following substances, in similar or higher concentrations, did not bring about precipitation: D-arginine, L-canavanine, agmatine, L-arginine hydroxamate, L-homoarginine, spermidine, and streptomycin.

**Interaction of ArgR with the Control Regions of *argR* and *argF*.** To test for the presumed biological activity of the purified ArgR, *in vitro* run-off transcription was carried out with a DNA fragment containing the control region of *argF*. About 0.02 pmol of the 280-bp *Pst*I-*Hind*III *argF* fragment was incubated with 1 pmol of ArgR with or without L-arginine. As shown in Fig. 4A, it is clear that ArgR inhibits transcription of *argF*, but only in the presence of L-arginine. Fig. 4B shows that the inhibition is dependent on the concentration of ArgR. This result confirms the previous conclusion that ArgR acts at the level of transcription and that it requires L-arginine.

The site of interaction of ArgR with DNA in *argR* was localized by DNase footprinting experiments. Both strands of a fragment containing the *argR* control region were labeled

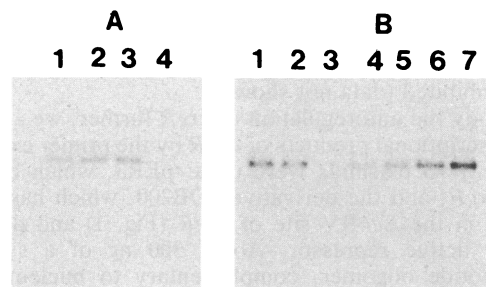


FIG. 4. *In vitro* run-off transcription of *argF* with ArgR, with or without L-arginine. The reaction products were analyzed in an 8% sequencing gel. (A) Dependence of inhibition of transcription on L-arginine. Lanes: 1, control run-off transcription without ArgR and L-arginine; 2, with 5 mM L-arginine, no ArgR; 3, with 100 ng of ArgR, no L-arginine; 4, with 100 ng of ArgR in 5 mM L-arginine. (B) Dependence on ArgR concentration. Lanes: 1, control without ArgR and L-arginine; 2, with 50 ng of ArgR, no L-arginine; 3, with 50 ng of ArgR in 5 mM L-arginine; 4, with 25 ng of ArgR in 5 mM L-arginine; 5, with 12.5 ng of ArgR in 5 mM L-arginine; 6, with 6.2 ng of ArgR in 5 mM L-arginine; 7, with 3.2 ng of ArgR in 5 mM L-arginine.

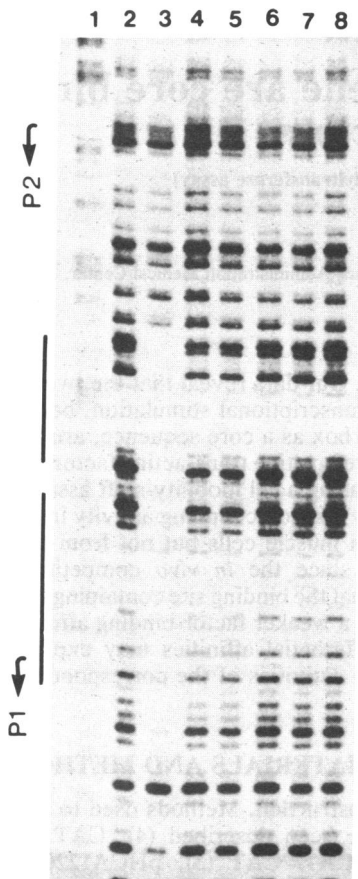


FIG. 5. DNase footprinting of *argR*. The sense strand was labeled in the *Spe* I site (Fig. 1) with Klenow fragment. The G+A ladder of the same fragment obtained by Maxam-Gilbert sequencing was used for determining the position of the protected region. The arrows indicate the transcription start points. ARG boxes are marked. Lanes: 1, G+A ladder; 2, without ArgR and L-arginine; 3, with 136 ng of ArgR in 5 mM L-arginine; 4, with 68 ng of ArgR in 5 mM L-arginine; 5, with 34 ng of ArgR in 5 mM L-arginine; 6, with 17 ng of ArgR in 5 mM L-arginine; 7, with 272 ng of ArgR, no L-arginine; 8, with 136 ng of ArgR, no L-arginine. P1 and P2 are proximal and distal promoters, respectively.

with <sup>32</sup>P and incubated with the purified ArgR in the absence or presence of L-arginine. After treatment with DNase, the reaction products were analyzed in sequencing gels. As shown in Fig. 5 (only the sense strand is shown, the antisense strand gave similar results), ArgR binds to the ARG boxes, but only with L-arginine. The protected sites are between coordinates -37 and +5. These results confirm the suggestion from sequence comparisons of *arg* genes and mutational studies that the ARG box sequences are the binding sites for ArgR. The concentration of ArgR required for 50% protection against DNase is  $\approx 2 \times 10^{-8}$  M.

Similar results were obtained in DNase footprinting experiments with the control region of *argF* (data not shown). The ARG boxes of *argF* conform more closely to the ARG box consensus sequence (1) than do those of *argR*.

### DISCUSSION

In this paper we have determined the nucleotide sequence of the *argR* gene and have described the isolation and purifi-

cation of its product, ArgR. We have also described the isolation of an ArgR- $\beta$ -galactosidase fusion protein. The properties of these two proteins support the structure of the ArgR subunit postulated from the nucleotide sequence. In its native state, ArgR appears to be a hexamer.

We have shown that the purified ArgR is active at low concentrations in inhibiting *in vitro* run-off transcription of *argF*, indicating that it has retained its biological activity. DNase footprinting experiments have shown that ArgR interacts with two adjacent palindromic ARG boxes in the control region of *argR* and *argF*. D. Charlier, R. Cunin, and N. Glansdorff (personal communication) in similar experiments have observed binding of our purified repressor to the ARG boxes of *argECBH* and *carAB*. The purified ArgR thus behaves as postulated on the basis of studies with mutants.

This work was supported by Public Health Service Grant GM06048 from the National Institute of General Medical Sciences. W.K.M. is the holder of Public Health Service Career Award GM15129 from the National Institute of General Medical Sciences.

- Cunin, R., Glansdorff, N., Pierard, A. & Stalon, V. (1986) *Microbiol. Rev.* **50**, 314-352.
- Gorini, L., Gundersen, W. & Burger, M. (1961) *Cold Spring Harbor Symp. Quant. Biol.* **26**, 173-182.
- Maas, W. K. (1961) *Cold Spring Harbor Symp. Quant. Biol.* **26**, 183-191.
- Maas, W. K., Maas, R., Wiame, J. M. & Glansdorff, N. (1964) *J. Mol. Biol.* **8**, 359-364.
- Maas, W. K. & Clark, A. J. (1964) *J. Mol. Biol.* **8**, 365-370.
- Eckhardt, T. (1980) *Mol. Gen. Genet.* **178**, 447-452.
- Saint-Girons, I., Belfaiza, J., Guillou, Y., Perrin, D., Guiso, N., Barzu, O. & Cohen, G. N. (1986) *J. Biol. Chem.* **261**, 10936-10940.
- Smith, A. A., Greene, R. C., Kirby, T. W. & Hindenack, B. R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6104-6108.
- Arvidson, D. N., Bruce, C. & Gunsalus, R. P. (1986) *J. Biol. Chem.* **261**, 238-243.
- Gunsalus, R. P. & Yanofsky, C. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7117-7121.
- Weinstock, G. M., Ap Rhys, C., Berman, M. L., Hampar, B., Jackson, D., Silhavy, T. J., Weisemann, J. & Zweig, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4432-4436.
- Crabeel, M., Charlier, D., Cunin, R. & Glansdorff, N. (1979) *Gene* **6**, 207-231.
- Gottesman, M. E., Adhya, S. & Das, A. (1980) *J. Mol. Biol.* **140**, 57-75.
- Vieira, J. & Messing, J. (1982) *Gene* **19**, 259-268.
- Rosenberg, M., Ho, Y.-S. & Shatzman, A. (1983) *Methods Enzymol.* **101**, 123-138.
- Aiba, H., Adhya, S. & de Crombrughe, B. (1981) *J. Biol. Chem.* **256**, 11905-11910.
- Curtis, S. E. (1987) *J. Bacteriol.* **169**, 80-86.
- Ullmann, A. (1984) *Gene* **29**, 27-31.
- Oppenheim, J. D., Nachbar, M. S. & Blank, M. (1983) *Electrophoresis* **4**, 53-62.
- Harboe, N. & Ingliid, A. (1973) in *A Manual of Quantitative Immuno-electrophoresis*, eds. Axelsen, N. H., Kroll, J. & Weeke, B. (Universitetsforlaget, Oslo), pp. 161-164.
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680-685.
- Miller, J. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560.
- Casadaban, M. J. & Cohen, S. N. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4530-4533.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
- DeBoer, H. A., Comstock, L. J. & Vasser, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 21-25.