Proc. Nat. Acad. Sci. USA
Vol. 68, No. 9, pp. 2145–2148, September 1971

Purification of the *araC* Protein

(affinity chromatography/positive regulator/anti-inducer/*ara* operon)

GARY WILCOX*, KENNETH J. CLEMETSON†, DANIEL V. SANTI†, AND ELLIS ENGLESBERG*

* Department of Biological Sciences and † Department of Chemistry, University of California, Santa Barbara, Calif. 93106

Communicated by Charles Yanofsky, June 28, 1971

ABSTRACT **The *araC* gene product, a regulatory protein required for expression of the L-arabinose operon, has been purified by affinity chromatography on Sepharose 4B to which 4-aminophenyl-β-D-6-deoxygalactopyranoside (an anti-inducer of the L-arabinose operon) had been covalently attached by means of a 4-aminophenylbutanamido side chain. Evidence is presented showing that the protein binds specifically to *ara* DNA.**

Positive control (1) of gene expression, as distinct from negative or repressor control (2), was first elucidated in the L-arabinose system of *Escherichia coli* B/r and has also been found to be the method of gene regulation of the maltose regulon (3) and probably the L-rhamnose operon (4). Such a system of positive control is also an integral part of the regulation of gene expression in bacteriophage lambda (5). The product of the specific regulatory gene is absolutely required for the normal expression of structural genes involved in operons under positive control. Superimposed upon the specific positive and negative control of inducible operons is a general system of catabolite positive control or catabolite deactivation‡ mediated through 3′,5′-cyclic AMP and the catabolite gene activator protein (CGA protein) (6, 7). In inducible operons in *E. coli*, the inducer determines the extent of the positive or negative control, as the case may be, as well as the extent of catabolite positive control (8, 9).

In the L-arabinose system in *E. coli* (see Fig. 1), positive control by the product of regulatory gene *araC* has been demonstrated by the finding that deletion or nonsense mutations (also some missense mutations) result in the production of a pleiotropic, L-arabinose-negative phenotype (C^-) uninducible for the first three enzymes involved in L-arabinose metabolism, L-arabinose isomerase (gene *araA*), L-ribulokinase (gene *araB*), L-ribulose 5-phosphate 4-epimerase (gene *araD*), and also for the L-arabinose-permease system, which includes the L-arabinose-binding protein (1, 10–14). The C^- alleles have also been shown to be recessive to the wild-type inducible allele (C^+) (1, 10). The functional product of the *araC* gene is a protein as has been demonstrated by the characterization of heat-labile mutants in gene *araC* [uninducible at 42°C and inducible at 28°C (1)]. In addition, it has been shown that nonsense suppressor mutations convert an uninducible phenotype produced by nonsense mutations in gene *araC* into an inducible one (12). C^c mutations, producing a constitutive phenotype, have been mapped within the *araC* gene. The dominance of C^+ to C^c, and the epistatic effect of the C^+ allele on the expression of I^c mutations (initiator constitutive), have led to the proposal that the normal product of the *araC* gene formed in the absence of L-arabinose exists mainly in the form of a repressor (15–17). L-Arabinose interacts with the repressor and/or the repressor–operator complex and converts the repressor into an activator. The activator, either acting directly on the DNA at the *araI* site or as a subunit of the RNA polymerase, stimulates the expression of the *araOIBAD* operon (Fig. 1, see ref. 18 for a more detailed description of the *ara* system). Some recent investigations suggest that both repressor and activator function at the level of transcription (18, 19, and unpublished data).

For several years, attempts have been made to isolate the *araC* gene product by routine protein-fractionation procedures coupled with assays for the *araC* gene product by specific binding to *ara*DNA, or to D-fucose or L-arabinose by equilibrium dialysis. These methods have failed to yield any consistent results. We have recently been encouraged to try affinity chromatography as a result of the pioneering work of Tomino and Paigen (21) and Cuatrecasas *et al.* (22). Since D-fucose is an anti-inducer of the L-arabinose regulon (1, 9), we made a series of fucoside derivatives and tested them for antiinducer effects. We found that 4-aminophenyl-β-D-6-deoxygalactopyranoside was a very effective antiinducer of the *ara*OIBAD operon, and made a column of Sepharose 4B to which 4-[4-(4-aminophenyl)butanamido]phenyl-β-D-6-deoxygalactopyranoside (PhBPhGal) was covalently attached. Using extracts of an F′ homogenote containing at least two copies of the *araC* gene, we have isolated a protein in about 20% purity that has the characteristics required for the *araC* gene protein. Besides binding to the anti-inducer, this protein binds specifically to φ80λ d*ara* DNA and is missing in extracts of a homogenote lacking the *araC* gene.

Abbreviations: CGA protein, catabolite gene activator protein; PhBPhGal, 4-[4-(4-aminophenyl)butanamido-]phenyl-β-D-6-deoxygalactopyranoside; SDS, sodium dodecyl sulfate.

‡ Since negative control, as defined in the *lac* operon (2), refers to the regulation of one operon by a specific repressor produced by a regulatory gene, we suggest that the term positive control be reserved for the direct opposite effect of negative control, that is, the regulation of one operon by a specific activator produced by a regulatory gene. For systems of general regulation, encompassing many operons, as in the case of the catabolite effect, we propose the terms catabolite deactivation and catabolite positive control. Since catabolite repression is actually not repression of operon expression (no repressor is involved), but rather reduction of activation, these terms will also serve the purpose of more closely describing what is taking place.

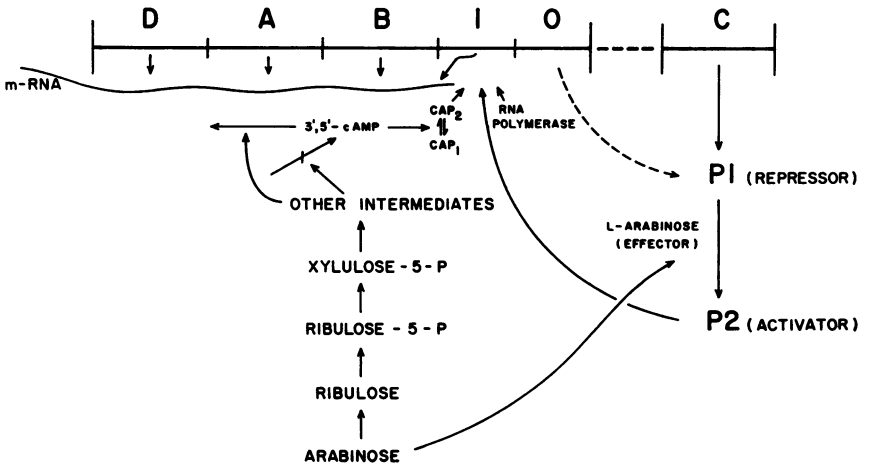

FIG. 1. The L-arabinose gene–enzyme complex (18), CAP = CGA protein.

MATERIALS AND METHODS

Preparation of crude extract

The strains used in this study were SB3102 (F'A39/A39), which was constructed by the procedures of Sheppard and Englesberg (11), and SB3142 (F'Δ719/Δ719), which is described by Sheppard and Englesberg (11). The strains are isogenic except for the *ara* markers. They were grown in medium that contained 7 g K_2HPO_4, 3 g KH_2PO_4, 1 g $(NH_4)_2SO_4$, 0.1 g $MgSO_4 \cdot 7H_2O$, 10 g of Casamino Acids (Difco), and 6 g of yeast extract/liter and were harvested when the culture had reached a density of 1×10^9 cells/ml. All of the following operations were performed at 4°C. A crude extract was prepared by suspension of the bacteria in 10 mM magnesium acetate–10 mM KCl–0.1 mM EDTA–0.1 mM dithiothreitol–10 mM Tris·HCl [pH 7.4 (Buffer A)], at a ratio of one part cells to two parts buffer. The cells were broken in a French press and the volume of the solution was tripled with Buffer A before removal of cell debris by centrifugation at 30,000 × *g* for 20 min. DNase was added to the supernatant to give a final concentration of 10 μg/ml and the solution was stirred for 12 hr.

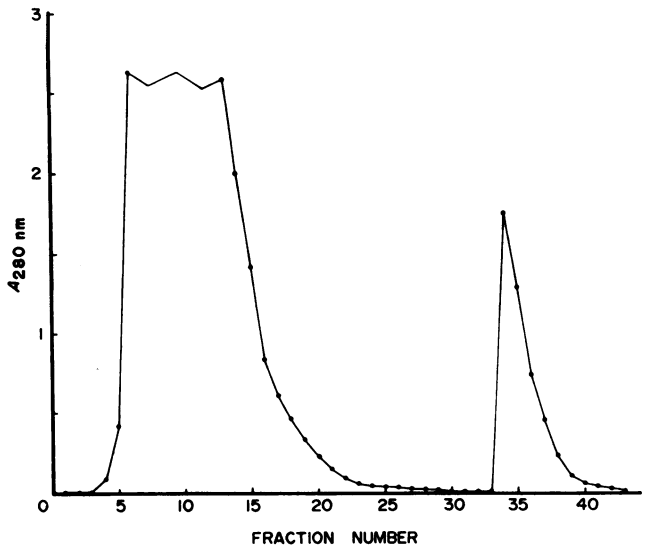

FIG. 2. Typical elution profile from the Sepharose column. The first peak represents material that washes through the column. Borate was applied at the beginning of fraction 31. Fractions containing 30 drops (0.62 ml) were collected. The flow rate was 18 ml/hr.

Affinity chromatography

The column material was prepared from Sepharose 4B (Sigma) to which PhBPhGal (manuscript in preparation) was covalently attached by activation with cyanogen bromide (23). The column material and standards of PhBPhGal were hydrolyzed in 6 N HCl at 100°C for 12 hr and assayed for aromatic-amine content by the methods of Daniel (24). The column material was shown to contain 4 μmol of aromatic amine/ml of packed Sepharose. The derivatized Sepharose was packed in a column (0.7 × 5 cm) and was equilibrated with 50 mM Tris·HCl [pH 7.0 (Buffer B)]. Crude extract (30–300 mg), prepared as described above, was applied and the column was washed with Buffer B until no more material that absorbed at 280 nm was eluted. The column was then eluted with 10 mM sodium borate (pH 10). The eluant was dialyzed against Buffer B for 12 hr and reapplied to the column. The above procedure was repeated twice. Chromatography and dialysis were performed at 4°C. The final eluant after dialysis was stored at 0°C.

DNA-binding assay

φ80λ d*ara* [^{32}P] DNA and φ80λ [^{32}P] DNA were prepared from a defective double-lysogen by the methods of Wilcox *et al.* (19) except that when the phage was heat-induced, 25 μCi/ml of [^{32}P]phosphate was added to the culture. The binding of the *araC* gene product to DNA was detected by the nitrocellulose-membrane-filter technique developed by Riggs and coworkers (20), except that dimethylsulfoxide was omitted from the buffers (BB buffer contains 10 mM magnesium acetate–10 mM KCl–0.1 mM EDTA–0.1 mM dithiothreitol–10 mM Tris·HCl (pH 7.4)–50 μg/ml of bovine-serum albumin).

Gel electrophoresis

Sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis was performed according to the procedures of Weber and Osborn (25) with slight modifications. After dialysis, the final borate eluant was concentrated 4-fold by ultrafiltration. After incubation in 8 M urea–1% SDS–1%

2-mercaptoethanol (37°C, 2 hr), the protein preparation was layered on the gels in 100 μl of sample buffer containing the incubation mixture in 10% (v/v) glycerol and 0.004% Bromophenol Blue. The gels were run for 4 hr at 8 mA/gel, removed from the tubes, and stained for 4 hr in 0.25% Coomassie Brilliant Blue in methanol–acetic acid–water 5:1:5. The gels were then kept overnight in 7.5% acetic acid–5% methanol and destained electrophoretically in the same solution.

RESULTS AND DISCUSSION

4-Aminophenyl-β-D-6-deoxygalactopyranoside is an anti-inducer of the L-arabinose operon in *E. coli* B/r (manuscript in preparation). The antiinducer activity of this compound suggested that it might interact with the product of the regulatory gene, *araC*. If this antiinducer has a sufficient affinity for the *araC* protein and could be covalently attached to Sepharose so that it would not be removed when a crude extract is passed through the column, we reasoned that it should be possible to purify the *araC* gene product, as well as other 6-deoxygalactopyranoside-binding proteins, from crude extracts. (We had previously found that *araC* protein, along with most other proteins in a crude extract, did not bind to underivatized Sepharose.)

Upon the first passage of a crude extract through the affinity column (see *Methods*), a considerable amount of material was bound that was released in a sharp peak when the column was eluted with 0.1 M sodium borate (pH 10). A typical elution profile is presented in Fig. 2. A portion of the first sodium borate eluant was no longer retained by the column upon recycling. By two successive recyclings, we were able to eliminate material that was (presumably) non-specifically bound to the column. Some of this initial, non-specific binding could be due to weak binding to the aromatic residues of the PhBPhGal side chain. Extensive washing of the column is not as efficient a way to remove the nonspecific material as the recycling process. Attempts to elute the column with L-arabinose or D-fucose did not remove any material that absorbed at 280 nm. The inability of a substrate to elute an enzyme from a substituted Sepharose column is not unprecedented. Steers, *et al.* (26) used a galactoside-Sepharose column to purify β-galactosidase, but were unable to elute the desired material with galactosides.

TABLE 1. *DNA-binding experiment*

Fraction	φ80λ *dara* DNA cpm bound	φ80λ DNA cpm bound	Ratio
1	635	615	1.03
2	775	626	1.24
3	1180	700	1.69

1.0 μg of protein from each fraction of the final dialyzed borate eluate from SB3102 was mixed with 0.1 μg of [^{32}P]DNA in BB buffer, to give a final volume of 1. 0 ml. After 10-min incubation at room temperature, two separate 0.5-ml aliquots were filtered through Schleicher and Shuell B6 membrane filters. Each point is the average of duplicate filtrations. There was 3000 cpm of [^{32}P]DNA in the reaction mix. The specific activities of the two DNAs were the same. The background obtained in the absence of protein has been substracted from each value.

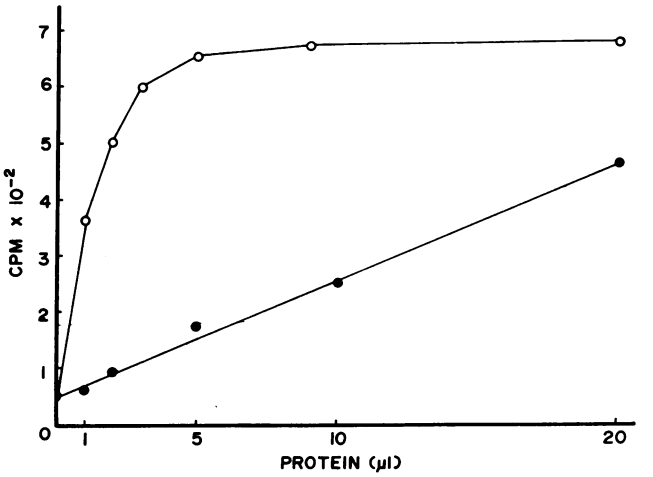

FIG. 3*a*. DNA-binding curves. After dialysis of the final borate eluate, the concentration was adjusted to 200 μg of protein/ml and the appropriate volume was mixed with 0.1 μg of [^{32}P]DNA in BB buffer to give a final volume of 1.0 ml. Duplicate filtrations were performed, as in Table 1. The reaction mix contained 1600 cpm of [^{32}P]DNA. ○ = φ80λ *dara* [^{32}P]DNA; ● = φ80λ [^{32}P]DNA.

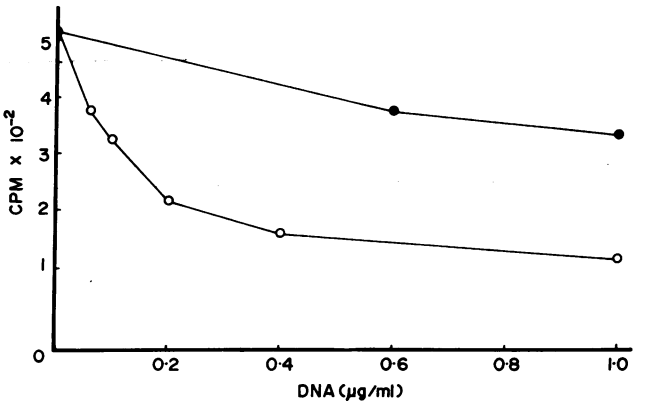

FIG. 3*b*. Competition experiment. The φ80λ *dara* [^{32}P]DNA and protein concentrations were held constant at two-thirds saturation. Conditions were the same as in the DNA-binding experiment (Fig. 3*a*). Various amounts of unlabeled DNA were mixed with the φ80λ *dara* [^{32}P]DNA before the protein was added. ○ = unlabeled φ80λ *dara* DNA; ● = unlabeled φ80λ DNA.

A crude extract prepared from strain SB3102 was purified on the column and three fractions were collected across the final eluted peak, dialyzed against Buffer B, and assayed for DNA-binding activity. As can be seen in Table 1, fractions 2 and 3 contained material that binds preferentially to φ80λ *dara* DNA. The material from fractions 1 and 3 was mixed in equal quantities; the *ara*DNA specific binding was reduced by half, indicating that no inhibitors of the specific binding of *ara*DNA were present in fraction 1.

Each fraction was examined by SDS–polyacrylamide gel electrophoresis. The gel of fraction 3 contained four major bands; it appears that the material in one band is associated with the binding to φ80λ *dara* DNA, since this is the only band that disappears in fraction 1 (which does not bind φ80λ *dara* DNA). The material in this band represents about 20% of the protein put on the gel.

With a crude extract prepared from SB3142, an elution profile similar to that shown in Fig. 2 was obtained on the Sepharose column. After recycling, the amount of material in the final eluant was about 70% of that which had been obtained with strain SB3102. Assay of the final eluant from the preparation from the SB3142 strain did not reveal any specific binding activity for ϕ80λ d*ara* DNA; examination by SDS–polyacrylamide gel electrophoresis showed that only two major bands were present. Furthermore, the band that contained the specific binding activity for ϕ80λ d*ara* DNA in SB3102 was absent from the SB3142 gel.

Increasing amounts of the final sodium borate eluate from strain SB3102 were mixed with a constant amount of ϕ80λ d*ara* [^{32}P] DNA, and a saturation curve was observed (Fig. 3*a*). A similar experiment, also presented in Fig. 3*a*, with ϕ80λ[^{32}P] DNA shows that much less DNA binding occurred with a given amount of eluant, indicating that this binding was nonspecific. Further evidence supporting specific binding to the ϕ80λ d*ara* DNA was obtained from competition experiments with unlabeled ϕ80λ d*ara* DNA (see Fig. 3*b*). The small amount of competitive binding observed with the ϕ80λ DNA and the amount of protein required to saturate *ara* DNA suggest that the C protein–DNA interaction may not involve as tight a binding as in the case of *lac* repressor–*lac* operator DNA (20). However, determinations of a binding constant based upon our experiments performed with crude *araC* protein would probably be open to considerable error. These binding studies indicate that the *araC* gene product binds to *ara* DNA. Addition of L-arabinose or D-fucose to the reaction mixture has no effect on the specific binding to *ara* DNA. Hybridization (19 and unpublished results) and genetic evidence (18) suggest that P1 and P2 (see Fig. 1) have functions at the transcriptional level; thus C protein, in either or both conformations, could conceivably bind to *ara* DNA.

The results described above strongly suggest that the protein that has been purified is the product of the L-arabinose C gene. The protein binds to 4-aminophenyl-β-D-6-deoxygalactopyranoside, an antiinducer of the L-arabinose operon and to *ara* DNA. The further purification, properties, and functions of this protein are currently under investigation.

We thank Dr. Nancy Lee for many helpful discussions and assistance with the polyacrylamide gel electrophoresis, and Dr. J. M. Clemetson for help in making the Sepharose column. We also acknowledge the work of Dr. Richard E. Laughlin and Dr. Joan Marie Frey, who tested other methods for purifying the *araC* gene product and, thereby, saved us much time and travail. This work was supported by grants from the National Science Foundation (GB-24093) to E. Englesberg and from the U.S. Public Health Service (CA-10266), from the National Cancer Institute to D. V. Santi.

1. Englesberg, E., J. Irr, J. Power, and N. Lee, *J. Bacteriol.*, **90**, 946 (1965).
2. Wilson, C., D. Perrin, M. Cohn, F. Jacob, and J. Monod, *J. Mol. Biol.*, **8**, 582 (1964).
3. Schwartz, M., *Ann. Inst. Pasteur, Paris*, **112**, 673 (1967).
4. Power, J., *Genetics*, **55**, 557 (1967).
5. Thomas, R., *J. Mol. Biol.*, **22**, 79 (1966).
6. Zubay, G., D. Schwartz, and J. Beckwith, *Proc. Nat. Acad. Sci., USA*, **66**, 104 (1970).
7. Emmer, M., B. de Crombugghe, I. Pastan, and R. Perlman, *Proc. Nat. Acad. Sci., USA*, **66**, 480 (1970).
8. Katz, L., and E. Englesberg, *J. Bacteriol.*, **107**, 34 (1971).
9. Beverin, S., D. E. Sheppard, and S. S. Park, *J. Bacteriol.*, **107**, 79 (1971).
10. Sheppard, D., and E. Englesberg, *Cold Spring Harbor Symp. Quant. Biol.*, **31**, 345 (1966).
11. Sheppard, D., and E. Englesberg, *J. Mol. Biol.*, **25**, 443 (1967).
12. Irr, J., and E. Englesberg, *Genetics*, **65**, 27 (1970).
13. Kessler, D., and E. Englesberg, *J. Bacteriol.*, **98**, 1159 (1969).
14. Hogg, R. W., and E. Englesberg, *J. Bacteriol.*, **100**, 423 (1969).
15. Englesberg, E., D. Sheppard, C. Squires, and F. Meronk, Jr., *J. Mol. Biol.*, **43**, 281 (1969).
16. Englesberg, E., C. Squires, and F. Meronk, Jr., *Proc. Nat. Acad. Sci., USA*, **62**, 1100 (1969).
17. Kessler, D., and E. Englesberg, *J. Bacteriol.*, **98**, 1159 (1969).
18. Englesberg, E., in *Metabolic Pathways*, ed. D. Greenberg and H. Vogel, in press.
19. Wilcox, G., J. Singer, and L. Heffernan, *J. Bacteriol.*, **108**, in press.
20. Riggs, A. D., H. Suzuki, and S. Bourgeois, *J. Mol. Biol.*, **48**, 67 (1970).
21. Tomino, S., and K. Paigen, in *The Lac Operon*, ed. D. Zipser and J. Beckwith (Cold Spring Harbor Laboratory for Quantitative Biology, 1970), p. 233.
22. Cuatrecasas, P., M. Wilchek, and C. B. Afinsen, *Proc. Nat. Acad. Sci. USA*, **61**, 636 (1968).
23. Porath, J., R. Axen, and S. Ernback, *Nature*, **215**, 1491 (1967).
24. Daniel, J. W., *Analyst (London)*, **86**, 640 (1961).
25. Weber, K., and M. Osborn, *J. Biol. Chem.*, **244**, 1406 (1969).
26. Steers, E., P. Cuatrecasas, and H. B. Pollard, *J. Biol. Chem.*, **246**, 196 (1971).