The Regulatory Region of the L-arabinose Operon: Its Isolation on a 1000 Base-pair Fragment from DNA Heteroduplexes

JOHN T. LIS AND ROBERT SCHLEIF

Department of Biochemistry
Brandeis University
Waltham, Mass. 02154, U.S.A.

(Received 18 November 1974)

A DNA fragment containing the L-arabinose operon regulatory region of Escherichia coli was purified from DNA heteroduplexes formed between opposite strands of two non-defective ara transducing phage. The phage and arabinose gene orientation is such that the heteroduplex contains two single-stranded "bubbles". The ara regulatory region and short portions of the flanking araB and araC genes are in the short duplex between the "bubbles". Extensive regions of homology between the phage genomes allowed nearly half of the DNA renatured from a mixture of the two phage DNAs to be in the form of heteroduplexes. Digestion of the reannealed DNA containing heteroduplexes and homoduplexes with the easily purified, single-strand specific nuclease S1 yielded the 1000 (1017 ± 20, n = 36) base-pair ara DNA duplex plus half and whole phage-length duplexes. The larger DNA duplexes were selectively precipitated by polyethylene glycol before the final purification by preparative electrophoresis on polyacrylamide gels. By these methods 10 to 20 µg of the 1000 base-pair DNA fragment were purified.

1. Introduction

Two of the four elements known to be involved in regulation of the L-arabinose operon of Escherichia coli, RNA polymerase and cyclic-AMP receptor protein, have been purified to homogeneity (Burgess, 1969; Berg et al., 1971; Anderson et al., 1971). The third protein element, the araC gene product, a protein which positively and negatively regulates expression of the arabinose operon, has been partially purified in biologically active form (Lee et al., 1974). Here we report the purification of physical quantities of a fourth component of the control system, the regulatory region of the DNA. The isolation of sizeable quantities of this DNA should facilitate both the nucleotide sequencing of the regulatory region and physical measurements probing the mechanism of positive and negative regulation exhibited by the arabinose operon. Although genetic, physiological, and in vitro studies of the ara operon have identified the regulatory elements, details of their interactions have yet to be determined (Englesberg et al., 1965,1969; Hirsh & Schleif, 1973; Greenblatt & Schleif, 1971; Lee et al., 1974).

The regulatory region DNA was isolated from a heteroduplex between the DNA of two different transducing phage by a variation of the technique originally used for purification of the β-galactosidase gene of E. coli (Shapiro et al., 1969). These methods,
involving heteroduplex formation and digestion with $S_1$ nuclease, augment the use of restriction enzymes for generating specific fragments from a longer DNA molecule (Danna et al., 1973).

2. Materials and Methods

(a) Media and chemicals

Media are as described in Schleif (1969). The crude pancreatic DNAase (DN-25), cytochrome $c$ (C7752), and ethidium bromide are from Sigma. Polyethylene glycol is from Union Carbide and is sold under the trade name Carbowax 6000. PEG† is also known as polyethylene oxide and chemically is $\text{HOCH}_2(\text{CH}_2\text{OCH}_2)_n\text{CH}_3\text{OH}$. Acrylamide and methylene bis acrylamide are from Bio-Rad Laboratories. A solution of cytochrome $c$ at 10 mg/ml was filtered through Millipore filters before use in electron microscopy experiments.

(b) Large scale purification of phage

A ligase overproducing strain, JTL164 (top8 str thi) (from D. Botstein, D5404) is grown at 35°C in a carboy containing 15 liters of 1.5-times concentrated yeast/Tryptone medium (Schleif, 1969) to an $OD_{550}$ value of 0.8, at which time maltose is added to 0.13%. Growth of phage on the ligase overproducing strain appears to give more unfragmented DNA following heteroduplex formation, although we have not quantitated the effect. When the culture reaches an $OD_{550}$ of 1.5, MgSO$_4$ is added to 0.007 M and transducing phage (possessing El837 and 77 markers) is added at a multiplicity of infection of 4. After 20 min at 35°C for absorption, we ensure that the phage are induced by heating the culture to 42°C for 15 min. The culture is cooled to 35°C and grown for 3.5 h with vigorous aeration. The cells are centrifuged by passing the culture through a DeLaval Gyro Tester at a rate of 500 ml/min. The cell pellet, about 90 g, is resuspended in 160 ml of $\lambda$ suspension medium (Lis & Schleif, 1973) and 1 mg of crude pancreatic DNAase is added. After adding 2.5 ml of chloroform the mixture is swirled in a 35°C bath for 30 min. Cell debris is removed by 2 successive 20-min centrifugations at 8000 g. The supernatant is layered on top of a CsCl block gradient and centrifuged as described previously (Lis & Schleif, 1973). The phage collected from the first block gradient is mixed with an equal vol. of CsCl of density 1.9. CsCl of densities 1.6 and 1.4 are layered on top of the phage layer. After a 90-min centrifugation at 21,000 revs/min in the SW27 rotor, the phage form a sharp band between the 1.6 and 1.4 density layers. Typical yields are quantities of phage containing 100 to 200 mg of DNA.

(c) $S_1$ nuclease

The $S_1$ nuclease was partially purified from $\alpha$-amylase powder (Sigma) through the heat step and ammonium sulfate precipitation described by Vogt (1973). For our purposes this preparation was sufficiently free from contaminating nucleases; it produced results indistinguishable in our system from a highly purified fraction supplied by V. Vogt.

(d) Purification of ara regulatory region DNA

Equal amounts of $\lambda$paraB114 and $\lambda$paraC116 phage (Lis & Schleif, 1975), a total of 37.5 mg of DNA, were added to 37.5 ml 0.1 M-NaOH, 20 mM-EDTA. After 20 min at 25°C the solution was brought to pH 8.5 with 3.75 ml of 2 M-Tris-HCl, pH 7.15. Formamide was added to 50%, and the solution was put into a no. 20 dialysis sack. After 3 min, a time sufficient to permit about 70% renaturation, the sack was plunged into 21 of 0°C, 0.01 M-Tris-HCl, pH 8.5, 1 mM-EDTA, to stop renaturation and dialyzed for 60 min. The DNA was then dialyzed to equilibrium in $S_1$ buffer (0.03 M-sodium acetate, pH 4.6, 0.3 M-NaCl, 1 mM-ZnSO$_4$, 5% glycerol), also at 0°C.

The renatured DNA was incubated with 800 units (Vogt, 1973) of the $S_1$ nuclease for 70 min at 35°C. The reaction was stopped by adding 1.5 ml of 2 M-Tris-HCl, pH 8.5, and 5.2 ml of 0.2 M-EDTA and heating to 63°C for 10 min. The larger DNA molecules were

† Abbreviations used: PEG, polyethylene glycol; Hin, Hemophilus influenzae restriction endonuclease.
precipitated with PEG by the addition of NaCl to 0.55 M and PEG to 40 mg/ml (Lis & Schleif, 1975). The solution was maintained at 4°C and centrifuged 3 times at 8000 g for 15 min, at 1.5, 10, and 36 h. The pellets were discarded. The resulting supernatant was 5 to 10-fold enriched in ara fragment DNA. The remaining DNA was precipitated by adding PEG to a final concentration of 90 mg/ml, chilling for 2 h, and centrifuging as above. The precipitated DNA was resuspended in 1/50 the original vol. of 0.01 M-Tris·borate (pH 8.3), 0.1 mM-EDTA, 5% glycerol, and 0.01% each of dyes xylene cyanol and bromophenol blue.

The purification of regulatory region DNA was completed by electrophoresis on a slab gel similar to that described by DeWachter & Fiers (1971), containing 4% acrylamide, 0.33% methylene-bisacrylamide, 0.1 M-Tris·borate (pH 8.3), and 1 mM-EDTA. The 16 cm × 14 cm × 0.3 cm gels with a 10-cm wide slot for sample were run at 45 mA (300 V) until the xylene cyanol marker reached the end. The gel was soaked for 0.5 h in 0.01 M-Tris·HCl, pH 7.5, containing 1 μg ethidium bromide/ml and DNA located by its fluorescence upon excitation from a 4 W, long-wavelength U.V. light. The DNA in the section of the gel containing the ara regulatory region fragment was electrophoretically extracted by a modification of the procedure of Petterson et al. (1973). A 1.7-cm internal diameter glass tubing was constricted at one end to 1.1 cm. The open end of about 2 cm of no. 32 dialysis tubing (Union Carbide) knotted at the other end was fitted over the constricted glass tubing end and held in place with a rubber stopper containing a 1.0 cm hole. The DNA was eluted from the gel within the tube into 8 ml of gel electrophoresis buffer within the dialysis sack by a 30 mA, current, 100 V, for 3 h. All glassware coming into contact with the DNA from this point on was silicone-coated (Dri-Film, sc-87, Pierce Chemical Co.). Using high resolution electron microscopy, it was observed that neither ethanol precipitation nor phenol extraction adequately removed small fragments of polyacrylamide from the DNA (Jay Hirsh, personal communication). Therefore, this DNA was concentrated and purified using a DEAE-column. The 10 to 20 μg of DNA were dialyzed into 0.1 M-Tris·HCl (pH 7.4), 0.1 M-KCl and loaded onto a 1.0 ml DEAE-column. The column was eluted with 3 ml vols. of 0.1 M-Tris·HCl, pH 7.4, with KCl concentrations of 0.4 M and 0.6 M. Most DNA is eluted in the 0.6 M-KCl fraction. In cases where ethidium bromide had been added to locate DNA, it was removed by extraction with isooamyl alcohol (Hudson et al., 1969). The final yield of ara DNA fragment, 15 μg, is approximately 2% of the total ara regulatory region DNA added to the heteroduplex solutions. In our experiments 70% of the DNA renatured and half of this was in useless homoduplex molecules, thus the maximum yield possible is 35%.

(e) Restriction enzyme Hin digestion of λ DNA

Purified Hin enzyme and M₂DNA were the gifts of Tom Maniatis, Harvard University. Digestion of 10 μg DNA in 50 μl 60 mM-NaCl, 6.6 mM-β-mercaptoethanol, 6.6 mM-MgCl₂, 6.6 mM-Tris·HCl (pH 7.4) was for 2 h at 37°C. Following the digestion, bromophenol blue tracking dye to 0.05% and glycerol to 5% were added and the sample was applied to the gel for electrophoresis.

3. Results

(a) The “double-bubble” heteroduplex

The entire ara regulatory region and portions of the flanking araC and araB genes are common to the two non-defective arabinose-transducing phage λparaB114 and λparaC116 (Lis & Schleif, 1975). The location of the ara DNA on the phage and the relative orientations of the ara and phage genes as determined in the accompanying paper (Lis & Schleif, 1975) are illustrated in Figure 1. A heteroduplex between the two phage DNAs is shown in the micrograph of Plate I, and this observed structure is precisely that predicted for the annealing of one strand of λparaC116 to the opposite strand from λparaB114. The small duplex flanked by single-stranded regions contains
the *ara* regulatory region, and the much longer duplexes contain the essential phage genes. The size of the internal duplex region was determined by electron microscopy of the heteroduplexes to be 1017$\pm$20 ($n = 36$) base-pairs.

(b) Polyethylene glycol precipitation and gel electrophoretic purification of the DNA fragment containing *ara* regulatory region

Approximately half of the duplex DNA molecules formed from annealing a mixture of denatured DNA of $\lambda$*paraB114* and $\lambda$*paraC116* are in the form of the double-bubble heteroduplex. The rest are homoduplex molecules. The 1000 base-pair duplex containing the *ara* regulatory region is purified by first digesting the heteroduplexes with single-strand specific nuclease S1. The short fragment containing control region can be separated from the longer duplexes of 15,300 and 28,900 base-pairs and whole length $\lambda$ DNA by electrophoresis on acrylamide gels as shown in Plate II, track 1. The migration velocity of the *ara* fragment corresponds to that of a 1000$\pm$50 base-pair fragment, as determined by comparison of a side-by-side electrophoresis with markers resulting from *Hin* nuclease digestion of $\lambda$ phage (Maniatis *et al.*, 1973). This is in good agreement with the estimate of the size of the undigested internal duplex, 1017 base-pairs, in the heteroduplex of Plate I.

The *ara* fragment constitutes at most about 1% of the total DNA after S1 digestion since 50% of the renatured molecules are homoduplexes and the fragment is only 2% of the length of $\lambda$ DNA. Large-scale purification necessitated development of a convenient high capacity enrichment prior to the final electrophoretic purification.

Polyethylene glycol can precipitate viruses and DNA (Yamamoto *et al.*, 1970; Lerman, 1971). It was noted by Yamamoto *et al.* (1970) that asymmetrical viruses require less PEG for precipitation than symmetrical viruses. Acting on this, we have found for DNA that high molecular weight DNA is precipitated by lower concentrations of PEG than low molecular weight DNA (Lis & Schleif, 1975). Thus large quantities of the *ara* regulatory region DNA fragment were purified by first precipitating the half and whole phage-length DNA molecules. The DNA remaining in solution, usually about tenfold enriched for the desired fragment, was then precipitated for the final purification step of preparative gel electrophoresis by raising the PEG concentration to 9%. The purification of the *ara* regulatory
Plate I. Heteroduplex of \textit{aparaC116} and \textit{aparaB114}. The electron microscopy was as previously described and the distance from the right end of the lambda molecule to the beginning of the substitution in \textit{aparaC116} served as an internal double-stranded DNA length standard of 15,300 base-pairs (Lis & Schleif, 1974). The arrow labeled \textit{ara} points to the location of the 1017 base-pair double-stranded DNA that contains the \textit{ara} regulatory region.
Plate II. Polyacrylamide slab gel electrophoresis of $S_1$ nuclease digested DNA. A mixture of $\lambda$paraB114 and $\lambda$paraC116 was denatured, renatured, digested with $S_1$ and an electrophoresis performed, as described in Materials and Methods, before (track 1) and after (track 2) PEG enrichment. 18 $\mu$g of DNA were loaded onto both tracks 1 and 2. Tracks 3 and 4 are controls in which each phage, $\lambda$paraB114 and $\lambda$paraC116, respectively, was taken separately through the same procedures performed on the mixture, track 2. The arrow labeled "$\lambda + \frac{1}{2}$$\lambda" shows the location of whole and half-length $\lambda$ DNA which are not resolved. The arrow labeled "ara" shows the location of ara regulatory region DNA, the small internal duplex in Fig. 1 and Plate 1.
region fragment achieved by PEG precipitation is demonstrated by the electro-
phoretic DNA separation shown in Plate II, tracks 1 and 2. Equal amounts of DNA
were applied to both tracks, but the sample applied to track 2 had been treated
with PEG. The results in Plate II prove that the combination of both \( \lambda paraB \) and \( \lambda paraC \)
phage are required to produce the 1000 base-pair fragment. Tracks 3 and 4, which
show no 1000 base-pair DNA band, result from either phage alone being denatured,
renatured digested with \( S_1 \) nuclease, and enriched for fragments with PEG precip-
itation. This result in conjunction with the genetic analysis of the two transducing
phage (Lis & Schleif, 1975) prove that the 1000 base-pair band is \( ara \) DNA common
to the two phage and contains a portion of \( araC \), the entire regulatory region and a
portion of \( araB \).

![Fig. 2](image-url)

**Fig. 2.** The length distribution of purified DNA fragments containing the \( ara \) regulatory
region.

(a) Double-stranded lengths. The \( ara \) DNA purified as described in Materials and Methods
but omitting the DEAE-column, was mounted for electron microscopy as described (Lis & Schleif,
1975). All 172 molecules from 4 different fields are included in the histogram.

(b) Single-stranded lengths. The purified \( ara \) DNA was denatured in alkali. 5 \( \mu l \) of DNA at
5 \( \mu g/ml \) were mixed with an equal vol. of alkali solution (0.2 \( M \) NaOH, 0.04 \( M \) EDTA) and after
10 min at 25°C 1 \( \mu l \) of 2 \( M \) Tris-HCl (pH 7.5), 1 \( \mu l \) cytochrome c at 5 \( \mu g/ml \), and 99 \( \mu l \) of form-
amide were added. This hypophase containing 90% formamide was spread on a hypophase con-
taining 55% formamide, 0.01 \( M \) Tris-HCl (pH 8.5) and 0.001 \( M \) EDTA and mounted for electron
microscopy. All 216 molecules from 17 fields are represented. The absolute lengths are approximate
having been computed from magnification settings \( \pm 3\% \), rather than from an internal standard.
The length of molecules in the main peak is 85% that of double-stranded \( ara \) DNA, consistent
with previously measured differences of single and double-stranded DNA lengths (Bujard, 1969).
Quality of the purified ara DNA fragment

The large-scale purification described in detail in Materials and Methods yields 15 μg of purified fragment containing ara regulatory region DNA. The length homogeneity of this DNA was estimated by electron microscopy and electrophoresis on polyacrylamide gels. Measurement of all molecules visible on several fields in the electron microscope shows a size distribution with over 90% of the molecules in a symmetrical peak centered about a mean of 270 nm with a standard deviation of 25 nm (Fig. 2(a)). The standard deviation associated with our measurements of DNA length with the electron microscope generally follows the same relation described by Davis et al. (1971), \( \sigma_L = 2.6 (L)^{1/2} \) where \( \sigma_L \) is the standard deviation and \( L \) is the length in base-pairs of the DNA measured. According to this relationship the standard deviation which should be found in measuring a homogeneous population of DNA molecules 1017 base-pairs long is 83 base-pairs or 8.2% of the length. Since we found a slightly larger standard deviation of lengths in the population, 9.2% or 93 base-pairs, the population could be slightly heterogeneous in length. If the lengths of the molecules were Gaussian distributed, then the population of molecules would have a mean length of 1017 base-pairs with a standard deviation of 43 base-pairs, \( 93 = (83^2 + 43^2)^{1/2} \).

The width of the DNA band following gel electrophoresis can also provide data on the length homogeneity of the population of DNA molecules. We find the width to be no more than 20% greater than the width of the 1125 base-pair band produced by Hin restriction enzyme digestion of λ DNA. From the molecular weight as a function of mobility for these gels, the difference in size of the molecules migrating at the front edge of the band from those migrating at the rear edge of the band is 20 base-pairs. This estimation of the heterogeneity is an upper bound since the Hin enzyme digested DNA fragments are homogeneous, but still possess finite band widths on the gels.

The possibility that single-strand breaks exist within either of the strands of the 1000 base-pair fragment was examined by electron microscopy of denatured fragment. The histogram of lengths of single-stranded DNA (Fig. 2(b)) shows that most of the DNA is unbroken.

4. Discussion

Several proven techniques currently exist for isolating specific nucleotide sequences from bacteria and phage. The first step for bacterial sequences is transferring the desired DNA to a lambdoid phage, providing a 100-fold purification. Following this, three methods have been successful: formation of a heteroduplex structure containing the desired sequences as the only double-stranded region followed by digestion with single-strand specific DNAases (Shapiro et al., 1969), protection of a specific sequence from double-stranded DNAases by DNA-binding proteins (LeTala & Jeanteur, 1971; Heyden et al., 1972; Pirrotta, 1973; Maniatis & Ptashne, 1973a; Gilbert & Maxam, 1973), and digestion by restriction enzymes combined with purification of the desired fragment by electrophoresis (Maniatis & Ptashne, 1973).

We have purified a 1000 base-pair DNA fragment containing the arabinose operon regulatory region by a modification of the heteroduplex method. The heteroduplexes were formed between two λara transducing phage containing this region in common. The orientation of the ara genes on these two phage with respect to the λ genes is such that in the heteroduplex, consisting of opposite strands from each of the phage, most of the phage genes form a double-stranded structure. Near the center of the
molecules the ara regulatory region is located between two single-strand "bubbles". Digestion of the single-strand regions with partially purified S1 nuclease followed by separation of the 1000 base-pair fragments from the whole and half length \( \lambda \) molecules provides purified ara regulatory region DNA.

The orientation of ara and phage genes in the two phage used eliminates the need for the time consuming and low-capacity step of strand purification prior to heteroduplex formation (Shapiro et al., 1969). When denatured DNA from each of our phage are reannealed, half the DNA forms homoduplexes from which no regulatory region can be purified, but half forms heteroduplexes containing the regulatory region flanked by single-strand "bubbles".

Although gel electrophoresis provides a workable final step in our purification, we have found it convenient first to remove 90% of the unwanted large molecules by their selective precipitation with polyethylene glycol. Using these methods, approximately 15 \( \mu g \) of this fragment containing ara regulatory region DNA are purified in preparations of the scale described in Materials and Methods.

Subfragments of the regulatory region may be isolated in quantity using the same techniques by starting with phage containing deletions entering the region. The isolation, genetic mapping and the physical mapping of the respective DNA duplex fragments from such deleted phage is a subject in the accompanying paper (Schleif & Lis, 1975).

Bartok et al. (1974) have recently reported the generation of specific DNA fragments by digestion of adenovirus heteroduplexes with a mixture of two single-strand specific nucleases from Neurospora crassa. Presently it is easier to use S1 nuclease in the digestion of large quantities of DNA as reported in this paper; however, the use of \( N. \) crassa enzymes may be preferable in situations requiring digestion of analytical quantities of DNA at neutral pH.

We thank Tom Maniatis for Hin enzyme, Volker Vogt for the initial preparations of purified S1 nuclease, and Allan Maxam in Walter Gilbert's laboratory, Harvard University for instruction with the gels. This work was supported by National Institutes of Health (research grant GM18277, Career Development award KO4-GM-38797, and training grant GM00212). This is publication no. 1023 from the Graduate Department of Biochemistry, Brandeis University.

REFERENCES


