Proc. Natl. Acad. Sci. USA
Vol. 73, No. 7, pp. 2316–2320, July 1976
Biochemistry

Isolation of inverted repeat sequences, including IS1, IS2, and IS3, in *Escherichia coli* plasmids

(S1 nuclease/gel electrophoresis/electron microscope)

HISAKO OHTSUBO AND EIICHI OHTSUBO

Department of Microbiology, Health Sciences Center, State University of New York at Stony Brook, Stony Brook, N.Y. 11794

Communicated by Norman Davidson, April 30, 1976

ABSTRACT A method is described for isolation of inverted repeat DNA sequences that occur in *E. coli* plasmids. The procedures of the isolation involved: (*a*) denaturation of intact plasmid DNA, (*b*) a rapid, 30 sec, renaturation of inverted repeat sequences in the genome, (*c*) digestion of the single-stranded portion by S1 nuclease to recover duplex DNA, and (*d*) detection and purification of the duplexes using 1.4% agarose gel electrophoresis. If a plasmid DNA carried inverted repeats of either one type or two different types of special DNA sequences, these procedures enabled us to observe either one or two characteristic DNA bands, respectively, in the agarose gels. If a plasmid DNA did not carry any inverted repeats, or if the plasmid DNA only carried direct repeat sequences, no characteristic DNA bands were recovered. Cleavage of the spacer DNA between inverted repeat sequences generated no gel bands. This indicated that the inverted repeat sequences must be in the same strand. Using this method, we isolated and purified several repeated sequences, including IS1, IS2, and IS3, from derivatives of F and R plasmids.

Electron microscope heteroduplex mapping methods have shown that bacterial F and R plasmids, certain bacteriophages, and the *Escherichia coli* chromosome contain several kinds of repeated DNA sequences (1–8). These DNA sequences appear either in the same orientation (direct repeats) or in an inverted orientation (inverted repeats) in the genome. Some of these sequences have been identified as the insertion (IS) sequences, IS1, IS2, and IS3, which cause strongly polar mutations when inserted into the *gal* and *lac* operons (4, 9–11). Repeated sequences, including these IS sequences, have been shown to also play important roles in fusion of two different genomes, deletion of a part of a genome, and the transposition of a DNA segment (3–5, 12–18).

In this paper, we report an isolation method for some of these repeated sequences, including IS1, IS2, and IS3, which occur as inverted repeats in the *E. coli* plasmid DNA. The method used is based on the fact that the inverted repeat DNA sequences form duplexes very rapidly after denaturation and renaturation of a genome containing these sequences (1, 2, 19). The duplex portion can be recovered by digestion of the remaining single-stranded DNA with S1-nuclease and then isolated by gel electrophoresis.

By the method described in the *text*, quite pure DNA of inverted repeat sequences can be obtained. This method will be applicable for the isolation of any inverted repeat sequence that occurs in a plasmid DNA molecule, a bacteriophage DNA, or in the *E. coli* chromosome.

MATERIALS AND METHODS

Plasmids. The plasmid strains used are listed in Table 1. Fig. 1 shows schematic representations of the physical structures of the R factors used. Repeated sequences are indicated both in Table 1 and Fig. 1.

Abbreviation: kb, kilobase(s).

The method of Sharp *et al.* (1) was used to isolate and convert covalently closed circular plasmid DNAs to open circular molecules; approximately 50% of the closed circles so obtained showed breaks in one strand. Ethidium bromide was removed from plasmid DNA by the use of isopropanol saturated with CsCl solution (1.50 g/ml) in TES buffer (pH 8.5) containing 0.05 M NaCl, 0.05 MTris·HCl, 0.005 M EDTA.

Procedure for Isolation of Inverted Repeats. The standard procedure used was as follows: (*i*) Purified plasmid DNA was denatured with 0.15 M NaOH for 15 min at room temperature; after neutralization with HCl, the solution was chilled in ice, and the Na^+ concentration was adjusted to 0.3 M with NaCl. (*ii*) Renaturation of inverted repeats was performed by incubating the DNA solution at 68° for 30 sec. (*iii*) Duplex DNA was recovered by digestion of single-stranded DNA with S1 nuclease. The reaction mixture for single-stranded DNA digestion contained 30 mM sodium acetate buffer at pH 4.6, 4.5 mM $ZnCl_2$, 0.3 M NaCl, and 200 units/μg of DNA from S1 nuclease (product of Miles Chemical Co.). Under these conditions, duplex DNA molecules are stabilized according to the paper by Shenk *et al.* (20). After incubation for 30 min at 37°, the reaction mixture was chilled in ice and the reaction was terminated by adding 30 mM EDTA (pH 8.5). (*iv*) The solution was then subjected to 1.4% agarose gel electrophoresis as described below. With use of these procedures, an initial concentration of 2–4 μg of plasmid DNA was enough to see characteristic DNA bands of inverted repeats with ethidium bromide.

Gel Electrophoresis. One and four-tenth percent agarose gels (0.3 × 15 × 13 cm slab gels or 0.6 × 15 cm disc gels) were prepared in E-buffer (pH 8.0) composed of 40 mM Tris·acetate, 2 mM EDTA, and 18 mM NaCl as described (21). Electrophoresis was carried out at 60 V for 5 hr for slab gels, or at 80 V for 4 hr for disc gels. After the run, the DNA bands were stained with ethidium bromide (0.4 g/ml) and visualized by a short wavelength ultraviolet light. The fluorescent bands were photographed on Polaroid type 55 P/N film with use of a Kodak no. 23A red filter. A Joyce Loebl photodensitometer was used to trace the negative films of the gels.

Molecular weights and amount of DNA in the gel bands were estimated from the electrophoretic mobility and density of the gel bands, respectively. *Eco*RI digested DNA fragments were used as controls for the determinations.

DNA fragments in the gels were eluted as described by Allet *et al.* (22).

Cleavage of DNA with *Eco*RI Endonuclease. *Eco*RI was purchased from Miles Laboratory. The reaction mixture contained 2–4 μg of DNA, 50 mM NaCl, 20 mM Tris·HCl (pH 7.4), 10 mM $MgCl_2$, and 2.5 μl of *Eco*RI endonuclease (100 units/ml). After incubation for 2 hr at 37°, the reaction was stopped by the addition of 30 mM EDTA, 0.02% bromophenol blue, and 8% glycerol.

The sites on R100 that are sensitive to *Eco*RI endonuclease

Table 1. Presence of the inverted repeats in plasmids used

Plasmid[a]	Molecular length in kb	Inverted repeat[b] sequence	Ref.
R100	89.3	IS3	2
RTC	69.2	IS3	—[c]
R100-27	75.8	None	—[c]
R100-25	90.1	IS1 and IS3	4
R6	94.8	IS3 and $\eta\theta$	2
F8(N33)	113.1	IS2	1

[a] The physical structures of R factors are shown in Fig. 1.

[b] IS3 (9, 10) is the same sequence as $\alpha\beta$ in F (3, 11–13). IS2 (9, 10) is the same sequence as $\epsilon\zeta$ in F (3, 4, 13). IS1 (9, 10) is the sequence which is directly duplicated in R100 and R6 (4). The sequence $\eta\theta$ has been named by us in this paper. This sequence was described by Sharp *et al.* (2). Lengths of these sequences are given in Table 2 of the *text*.

[c] E. Ohtsubo, manuscripts in preparation.

have been mapped as shown in Fig. 1 (unpublished results).

Electron Microscopy. Samples for electron microscopy were prepared by the aqueous spreading technique described by Davis *et al.* (23) and the formamide spreading technique described by Sharp *et al.* (1, 2). ColE1 DNA (4.2×10^6 daltons or 6.34 kb), kindly supplied by Dr. J. Tomizawa, was usually used as a standard for determination of the molecular lengths of duplex DNA.

RESULTS

Conditions for the isolation of inverted repeat sequences

Fig. 2A-d shows that with use of the procedures described in *Materials and Methods*, a plasmid, RTC, carrying the inverted repeat sequences of IS3 (see Table 1 and Fig. 1) generated a characteristic band, approximately 1.3 kilobases (kb) or 1300 bases long, in an agarose gel. Both denaturation and renaturation of intact plasmid DNA molecules and the digestion of single-stranded DNA with S1 nuclease are essential for detection of the new bands in agarose gels. When the DNA was not denatured with NaOH, such a characteristic DNA band did not appear, either with or without S1 nuclease treatment (Fig. 2A-a and 2A-b). It is shown in Fig. 2A-c that even when DNA was denatured and renatured, no gel bands appeared at the position of DNA of 1.3 kb if there was no S1 nuclease treatment.

Renaturation of inverted repeat sequences to form a fold-back structure is a very rapid process. Fig. 2B shows the effects of incubation time on renaturation. A 30 sec incubation of a denatured DNA solution at 68° gives a maximum density of the characteristic band that is 1.3 kb in length after S1 nuclease treatment (Fig. 2A-b). There is a decrease in the density of gel bands in longer incubations that is possibly due to homoduplex formation of plasmid genomes. We sometimes observed that storage of the DNA solution in ice for several hours produced the characteristic DNA band (see Fig. 2B-a) after S1 nuclease treatment, although the density of the gel band is less than that for a 30 sec incubation. This may also indicate that the formation of fold-back structures in the inverted repeat sequences is a very rapid process.

Fig. 2C shows the effects of different amounts of S1 nuclease in the reaction mixture. S1 nuclease in the amount of 200 units/μg of DNA is approximately enough to separate duplexes (Fig. 2C-a). The addition of less enzyme caused the loss of the characteristic band (Fig. 2C-c and -d). A large excess of enzyme (more than 500 units/mg of DNA) did not increase the density of the characteristic band (data not shown).

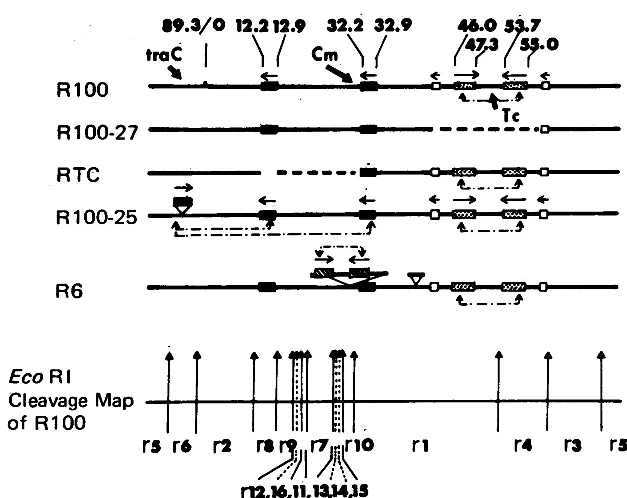

FIG. 1. Physical structures of R factors listed in Table 1. The molecules are actually circular duplexes, but the structures are displayed in a linear representation, by cutting the circle of a single strand at a point which lacks interesting features. All the numbers are assigned in kb from a selected origin. Deleted regions in R100-27 and RTC are shown by broken lines, whereas insertions in R100-25 and R6 are also shown as in the figure. Special repeated sequences are indicated as follows: ■ for IS1; ▨ for IS3; □ for the DNA sequence of about 0.5 kb in length which we call $\iota\kappa$ (unpublished results); and ▩ for the $\eta\theta$ sequence. The arrows above the repeated sequences indicate the sequence orientations. The two copies of the sequences in an inverted orientation in a molecule form duplexes after renaturation as indicated by ←-----→. Approximate map positions of the genes responsible for genetic transfer (*traC*), resistance to chloramphenicol (*Cm*), and resistance to tetracycline (*Tc*) are also shown on the R100 map. *Eco*RI cleavage sites on R100 are shown at the bottom of the figure; the same restriction sites are on R6 homologous to R100 (unpublished results). R100 produces 16 fragments, r1 ~ r16, by *Eco*RI digestion (see Fig. 2A-e).

In the experiments shown here and later, we usually observed some DNA remaining at, and close to, the top of the gels. This DNA was not eliminated by any steps in the procedure. We assume, from the experiments shown in Fig. 2, that circular DNA molecules remained at the top of the gels, whereas any DNA with a linear structure, including double-stranded and/or single-stranded DNA, formed a broad band at the position of duplex DNA with molecular lengths of 10–90 kb or else appeared as a diffuse background.

Under the conditions described in *Materials and Methods*, recovery of the inverted repeat sequences from plasmid DNA was between 40 and 80%. This was estimated from the density of the gel bands in which *Eco*RI-digested fragments of R100 were used as a control.

Isolation of inverted repeats from derivatives of F and R plasmids

The plasmids strains listed in Table 1 have been analyzed by the electron microscope heteroduplex method. Except for R100-27, these plasmids contain either one or two different sets of inverted repeat sequences in their genomes (see Table 1 and Fig. 1). These plasmids were used to isolate the DNA sequences of inverted repeats, including IS1, IS2, and IS3. Fig. 3A and B show the results of gel electrophoresis after S1 nuclease treatment of various R plasmids and of F8(N33), an F derivative. When R100, RTC, and F8(N33) which carried one kind of inverted repeat were examined, a single DNA band was observed in agarose gels (Fig. 3A-a and -c, and 3B). When the plasmids, R6 and R100-25, which carried two different sets of inverted repeats were examined, two bands appeared in the gels (Fig.

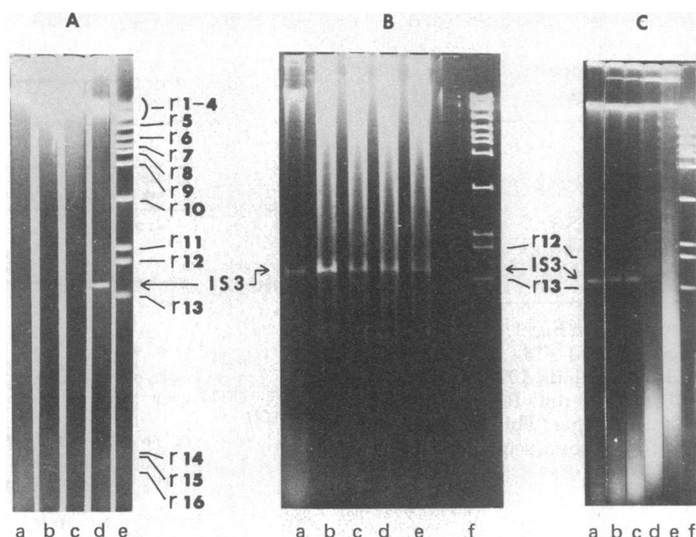

FIG. 2. Electrophoresis of DNA in 1.4% agarose gels, showing generation of a characteristic DNA band from RTC plasmid DNA which carries inverted repeats of IS3. A. Effects of denaturation, renaturation, and S1 nuclease treatment. (a) Native RTC duplex DNA, (b) Native RTC duplex DNA with S1-treatment, (c) Denatured and renatured RTC DNA, and (d) Denatured and renatured RTC DNA with S1-treatment. The procedures for denaturation, renaturation, and S1-treatment are described in *Materials and Methods*. B. Effects of incubation time on renaturation. RTC DNA was denatured, renatured, and treated with S1 nuclease as described in *Materials and Methods*, except the times for renaturation were (a) 0 sec, (b) 30 sec, (c) 3 min, (d) 5 min, and (e) 10 min. C. Effects of amount of S1 nuclease. Denatured and renatured RTC DNA was digested with S1 used in the following amounts: (a) 200, (b) 100, (c) 50, and (d) 10 units/μg of DNA, respectively. Panel A column (e), Panel B column (f), and Panel C column (f) are the *Eco*RI digests of R100 which are used as standards for determining length. The lengths of fragments, r1, r11, r12, and r13 are approximately 21.1 kb, 1.7 kb, 1.5 kb, and 1.2 kb, respectively (unpublished results).

3A-d and -e). R100-27 which carried no inverted repeats did not produce any characteristic bands in a gel (Fig. 3A-b). These results indicate that inverted repeats can be isolated by the procedures employed here. As seen in Fig. 3A, the common DNA bands of 1.3 kb in length, which are derived from R100, RTC, R100-25, and R6, must be IS3. The other DNA band derived from R100-25 was estimated to have about 750 bases and must be IS1. R6 produced a second band of 1.0 kb in length as expected. The single band derived from F8(N33) must be IS2.

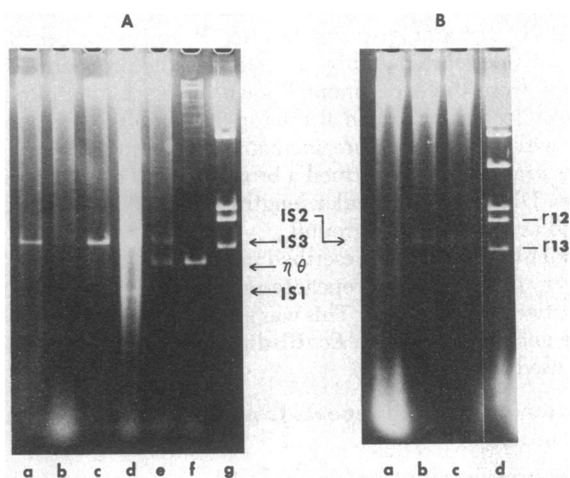

FIG. 3. Electrophoresis of DNA in 1.4% agarose gels. Panel A shows generation of characteristic DNA bands of inverted repeat sequences present in various R plasmids: (a) R100, (b) R100-27, (c) RTC, (d) R100-25 (see also Fig. 4A), (e) R6, and (f) DNA fragments of R6 cleaved by *Eco*RI. The DNA was denatured, renatured, and treated with S1 nuclease before gel electrophoresis as described in *Materials and Methods*. Panel B shows generation of inverted repeat sequences of IS2 in F8(N33) under the condition described in *Materials and Methods*, except that denatured F8(N33) DNA was incubated for renaturation for (a) 0 sec, (b) 30 sec, (c) 3 min. Panel A column (g) and Panel B column (d) are the *Eco*RI digests of R6 and R100, respectively, which were used as standards for determining length. The lengths of the fragments r12 and r13 are approximately 1.5 kb and 1.2 kb, respectively.

As schematically shown in Fig. 1, the R plasmids used here were known to have direct repeats of IS1 (750 bases) and of the sequences denoted ικ of length 500 bases in their genomes. It is obvious from the results above that these direct repeats could not be isolated by the method used.

Our isolation procedure requires that the two copies of an inverted repeat be linked together on a single molecule. This can be demonstrated by the following experiment. An R plasmid, R6, was digested by *Eco*RI endonuclease which cleaves the linkage between two repeats of IS3, but holds the linkage of the two copies of the sequence (ηθ) that is 1.0 kb long in the same strand (see *Eco*RI cleavage map shown in Fig. 1). The *Eco*RI-digested R6 DNA was then denatured and renatured and treated with S1 nuclease. As expected, after gel electrophoresis, only one single band that is 1.0 kb long was seen (Fig. 3A-f).

Electron microscopic observation of inverted repeat sequences

All the characteristic DNA bands seen in Fig. 3 were eluted from gels by gel electrophoresis as described by Allet *et al.* (21), and examined in the electron microscope. The preparation contained duplex DNA that was quite homogeneous in size, but usually also contained a small portion of a single-stranded DNA contaminant that could be removed by nitrocellulose treatment or by extensive gel electrophoresis. In Fig. 4, electron micrographs of purified IS1 and IS3 are shown. Table 2 summarizes the molecular lengths of isolated DNA determined by electron microscopy.

We also examined the lengths of inverted repeat sequences which were seen as duplex stems in the inversion loops formed after denaturation and renaturation of whole plasmid genome. In Fig. 5, an electron micrograph of an R100-25 molecule which

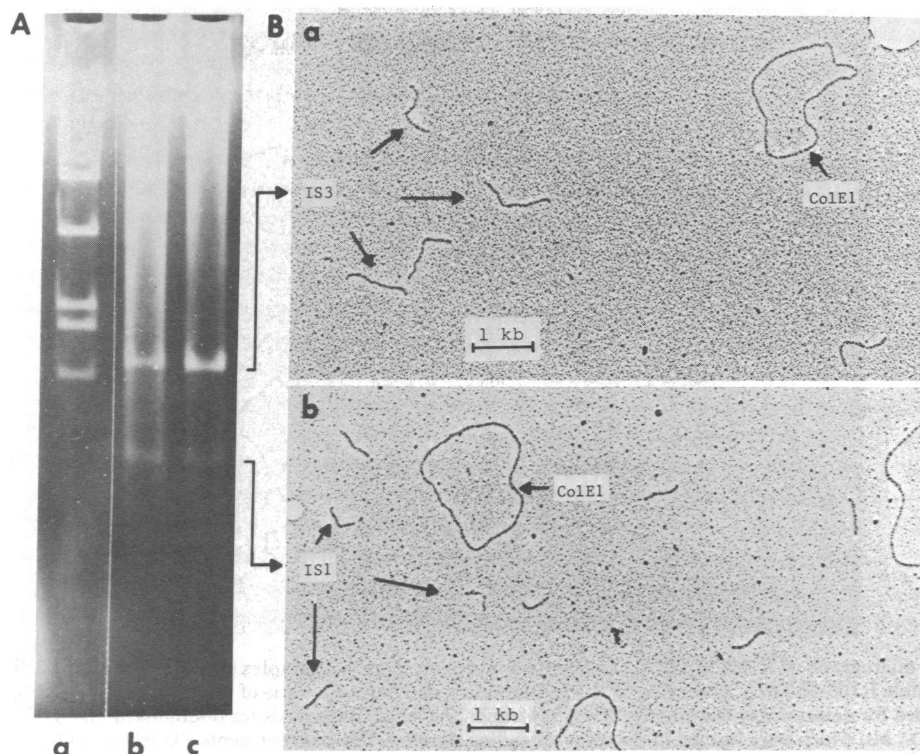

FIG. 4. A. Agarose (1.4%) gel electrophoresis of (b and c) R100-25 DNA which was denatured, renatured, and treated with S1 as described in *Materials and Methods* and of (a) *Eco*RI-digested DNA fragments of R100 which were used as standards for determining length. B. Electron micrographs of isolated inverted repeat sequences of (a) IS3 and (b) IS1 derived from R100-25. The DNA samples were eluted and purified further to remove a small portion of contaminated single strand DNA as follows: DNA fragments were precipitated in ethanol, then resuspended in a buffer containing 10 mM Tris·HCl at pH 8.5, 1 mM EDTA, and 0.3 M NaCl. Nitrocellulose powder was added and the mixture was shaken gently for 45 min at 4°. After removing the nitrocellulose powder by centrifugation (4,400 × g, 15 min), the DNA solution was treated with phenol, which was saturated with TE buffer (10 mM Tris·HCl at pH 8.5, 1 mM EDTA), and was then dialized against TE buffer. The aqueous spreading technique (23) was used to prepare the DNA samples for electron microscopy. In the micrographs, ColE1 double-stranded molecules (approximately 6.34 kb long) are shown and were used as standards for determining length.

forms two inversion loops with duplex stems of IS1 and IS3 is shown. All the measurements of stem lengths are shown in Table 2. The values obtained by the two methods are in good agreement.

DISCUSSION

A method is described in this paper to isolate repeated sequences in the *E. coli* plasmids. We have shown that the method can be applied to the isolation of inverted repeat sequences, but not to the isolation of direct repeat sequences that occur in the plasmid genome. Obviously, the method can be used for the isolation of inverted repeated sequences from other plasmids, from bacteriophage DNAs such as λ and P1, and also from the *E. coli* chromosome. It is known that inverted repeat sequences occur in these genomes (2, 5, 8, 17, 18).

Table 2. Determination of molecular lengths of repeated sequences by electron microscope

Repeated sequence	Plasmid	I	II
IS1	R100-25	0.71 ± 0.05 (92)	0.73 ± 0.13 (17)
IS2	F8(N33)	1.22 ± 0.07 (51)	1.27 ± 0.04 (12)
IS3	R100-25	1.29 ± 0.07 (80)	1.33 ± 0.14 (21)
IS3	R6	1.29 ± 0.10 (50)	1.34 ± 0.10 (45)
$\eta\theta$	R6	0.96 ± 0.05 (57)	1.02 ± 0.06 (26)

I. Molecular lengths, in kb, of the fragments eluted from gels. ColE1 DNA (6.34 kb) was used as the standard for determining length. II. Lengths, in kb, of the duplex stems in the inversion loops formed after denaturation and renaturation of whole plasmid DNA. The lengths of IS1 and IS3 in R100-25 and IS2 in F8(N33) were measured from homoduplexes of R100-25 (90.1 kb) and F8(N33) (113.1 kb), respectively, that were used as the standard for determining length. The lengths of IS3 and $\eta\theta$ in R6 were measured with use of ColE1 (6.34 kb) as the length standard. Numbers in parentheses are the number of molecules measured.

Mutant plasmids with known physical structures can be used not only to isolate several kinds of repeated sequences, but also to characterize some of these sequences as IS1, IS2, and IS3. However, we have been developing a different approach in our laboratory for characterization of sequences by using restriction endonucleases. Preliminary studies indicate that a restriction endonuclease, *Hin*dIII, cleaves the IS2 sequence, but not IS1 and IS3, whereas other restriction endonucleases, such as *Eco*RI and *Bam*HI, do not cleave any of these sequences. This approach can be applied directly to characterize the duplex DNA of inverted repeats isolated by the method described here.

Electron microscope heteroduplex studies have revealed that insertion sequences cause not only insertion duplications into any one of the sites of DNA, but also deletion formation and transposition of a special DNA sequence (2, 5, 14, 15, 17, 18). These phenomena indicate that the end points of insertion sequences are effectively sensitive to recombination. It will therefore be of interest to look for any special characteristics of the sequences at the ends of the duplex IS segments. Such studies will require the use of digestion conditions in which a nuclease specific for a single strand that does not result in "nibbling" effects at the ends (Shenk *et al.* ref. 20) is used.

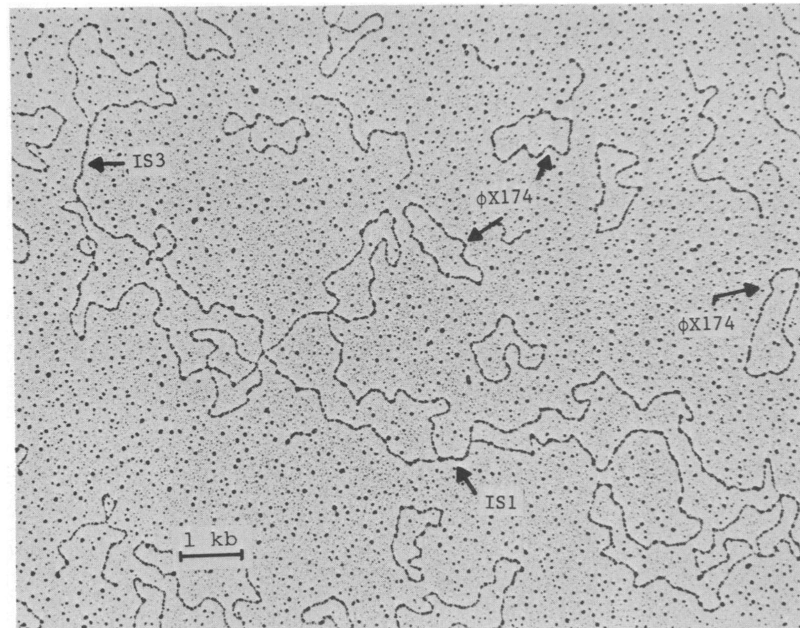

FIG. 5. An electron micrograph of R100-25 DNA, showing two inversion loops with duplex stems of IS1 and IS3. This molecule was seen after denaturation of whole R100-25 DNA, according to the formamide spreading technique of Sharp *et al.* (1). ϕX174 single-stranded DNA was added as the standard for determining the lengths of single strands. As length standards for double-stranded DNA, we used homoduplex molecules of R100-25 (90.1 kb) which were formed in the same spreading samples. The bar represents a length of 1 kb for double-stranded DNA.

We would like to thank Dr. N. Davidson for his generous criticism. We would also like to thank Dr. J. Tomizawa for sending ColE1 DNA. This investigation was supported by Public Health Service Grant no. GM22007-02.

1. Sharp, P. A., Hsu, M.-T., Ohtsubo, E. & Davidson, N. (1972) *J. Mol. Biol.* **71**, 471–497.
2. Sharp, P. A., Cohen, S. N. & Davidson, N. (1973) *J. Mol. Biol.* **75**, 235–255.
3. Ohtsubo, E., Deonier, R. C., Lee, H. J. & Davidson, N. (1974) *J. Mol. Biol.* **89**, 565–584.
4. Hu, S., Ohtsubo, E., Davidson, N. & Saedler, H. (1975) *J. Bacteriol.* **122**, 764–775.
5. Lee, H. J., Ohtsubo, E., Deonier, R. C. & Davidson, N. (1974) *J. Mol. Biol.* **89**, 585–597.
6. Tye, B. K., Chan, R. K. & Botstein, D. (1974) *J. Mol. Biol.* **85**, 485–500.
7. Hsu, M.-T. & Davidson, N. (1975) *Virology* **58**, 229–239.
8. Hu, S., Ohtsubo, E. & Davidson, N. (1975) *J. Bacteriol.* **122**, 749–763.
9. Hirsch, H. J., Starlinger, P. & Brachet, P. (1972) *Mol. Gen. Genet.* **119**, 191–206.
10. Malamy, M. H., Fiandt, M. & Szybalski, W. (1972) *Mol. Gen. Genet.* **119**, 207–222.
11. Hu, S., Ptashne, K., Cohen, S. N. & Davidson, N. (1975) *J. Bacteriol.* **123**, 687–692.
12. Ptashne, K. & Cohen, S. N. (1975) *J. Bacteriol.* **122**, 776–781.
13. Davidson, N., Deonier, R. C., Hu, S. & Ohtsubo, E. (1975) in *Microbiology 1974* (American Society for Microbiology, Washington, D.C.), pp. 56–65.
14. Saedler, H., Reif, H. J., Hu, S. & Davidson, N. (1974) *Mol. Gen. Genet.* **132**, 265–289.
15. Kleckner, N., Chan, R. K., Tye, B.-K. & Botstein, D. (1975) *J. Mol. Biol.* **97**, 561–575.
16. Heffron, F., Rubens, C. & Falkow, S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3623–3627.
17. Berg, D. E., Davis, J., Allet, B. & Rochaix, J. D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3628–3632.
18. Reif, H. J. & Saedler, H. (1975) *Mol. Gen. Genet.* **137**, 17–28.
19. Schmidt, C. W., Manning, J. E. & Davidson, N. (1975) *Cell* **5**, 159–172.
20. Shenk, T. E., Rhodes, C., Rigby, P. W. J. & Berg, P. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 989–993.
21. Greene, P. J., Betlach, M., Goodman, H. M. & Boyer, H. (1974) in *Methods in Molecular Biology Series: DNA Replication and Biosynthesis*, ed. Wickner, R. B. (Marcel Dekker, New York), Vol. 7, pp. 87–111.
22. Allet, B., Jeppesen, P. G., Katagiri, K. J. & Delius, H. (1973) *Nature* **241**, 120–122.
23. Davis, R. W., Simon, M. & Davidson, N. (1971) in *Methods in Enzymology*, eds. Grossman, L. & Moldave, K. (Academic Press, New York), Vol. 21, pp. 413–428.