JOURNAL OF BACTERIOLOGY, May 1975, p. 776-781
Vol. 122, No. 2

Printed in U.S.A.

Occurrence of Insertion Sequence (IS) Regions on Plasmid Deoxyribonucleic Acid as Direct and Inverted Nucleotide Sequence Duplications

KAY PTASHNE AND STANLEY N. COHEN*

Department of Medicine, Stanford University School of Medicine, Stanford, California 94305

Received for publication 16 December 1974

Insertion sequence (IS) regions have been identified previously as a cause of strongly polar mutations in *Escherichia coli* and several bacteriophages. The present experiments indicate that genetically characterized IS regions occur on bacterial plasmid deoxyribonucleic acid (DNA) as both direct and inverted DNA sequence duplications. The DNA insertion which has been shown previously (Sharp et al., 1973) to control expression of tetracycline resistance in the R6-5 plasmid, and which occurs as directly and inversely repeated DNA sequences adjacent to the region believed to contain the tetracycline resistance gene, has been identified as IS3. A second genetically characterized insertion sequence (IS1) has been identified as a direct DNA duplication occurring at both junctions of the resistance transfer factor and R-determinant components of R6-5 and related plasmids. A model is presented for the reversible dissociation of resistance transfer factor and R-determinant components of co-integrate R plasmids at the sites of DNA sequence homology provided by the repeated IS regions.

Although direct and inverted tandem duplications of deoxyribonucleic acid (DNA) nucleotide sequences have been observed as a feature of both prokaryotic (7, 10, 12, 15, 19) and eukaryotic (26) DNA, the biological function of such duplications is unknown. Adelberg and Berquist (1) have suggested that direct duplications of nucleotide sequences may provide sites of homology on DNA for genetic recombination; other recent evidence suggests that DNA duplications in the form of inverted repeats (palindromes) may have a role in *recA*-independent recombination and translocation of plasmid segments (12a).

Recently, certain DNA segments repeated at multiple sites on the *Escherichia coli* chromosome have been shown by electron microscopy to be identical to the insertion sequences (IS regions) identified earlier as a cause of strongly polar mutations in several operons of *E. coli* and its bacteriophages (9, 14, 19). Such preformed genetic duplications, which range in length from 800 to 1,400 nucleotides, have the capacity for insertion at multiple loci of bacterial and phage genomes. At least one such sequence, IS2, has been shown to act as a controlling element capable of affecting the expression of bacterial gene activity (20).

R6-5 is a tetracycline-sensitive spontaneous mutant of the tetracycline resistance plasmid R6. Earlier studies (21) have shown that the R6-5 plasmid contains a 1.4-kilobase (kb) DNA segment repeated in both direct and reverse orientation near the tetracycline resistance gene of the plasmid. Moreover, electron microscope heteroduplex investigations indicate that insertion of a directly repeated copy of this genetic duplication prevents expression of the adjacent tetracycline resistance gene (21). The inserted DNA sequence duplication is excised spontaneously at low frequency (21) and is excised at higher frequency after transformation of *E. coli* by R6-5 plasmid DNA (3), leading to regeneration of a molecule that is phenotypically and structurally identical to the tetracycline-resistant R6 parent plasmid.

The present studies identify the DNA duplication controlling expression of the tetracycline resistance gene of the R6-5 plasmid as one of the IS units characterized previously (i.e., IS3). In addition, a second genetically characterized insertion sequence (IS1) occurs as a direct tandem duplication located at both junctions of the resistance transfer factor (RTF) and R-determinant components of the R6-5 plasmid.

MATERIALS AND METHODS

Bacterial strains, phages, and plasmids. Bacterial strains, plasmids, and methods used for isolation

of covalently closed, circular plasmid DNA have been described previously (25). The lysogenic *E. coli* strain containing the bacteriophage λ mutants λ*cI857plac*5S7-MS505 (carrying the IS3 insertion sequence in the ω gene of the *lac*5 operon) (14) was a gift from M. H. Malamy. Lysogenic *E. coli* strain C600 (λ*cI857NNr14*) carrying the IS1 insertion sequence was obtained from S. Hu and N. Davidson. Lambda phage were isolated after heat induction as described previously (2); in the case of λ*r14*, the defective phage particles carrying the IS1 segment were separated from λ*cI857t68* helper phage by cesium chloride gradient centrifugation (16).

Separation of λ DNA strands and heteroduplex analysis. Since the IS regions carried by λ represent only a small part of the λ phage genome, any heteroduplex formation between λ-IS DNA and plasmid DNA would be expected to occur much less commonly than the formation of homoduplexes involving the two DNA strands. Separated λ DNA strands were therefore used to favor heteroduplex reactions.

Separated strands of λ DNA were isolated directly from suspensions of purified phage by a modification of procedures described earlier (4, 23, 24). In the present experiments, λ strands were preparatively and analytically separated in CsCl density gradients by centrifugation in the presence of the ribonucleotide polymer poly(U,G). Reaction mixtures for strand separation contained the following: 0.1 ml of phage suspension in 1 mM ethylenediaminetetraacetic acid, pH 8.5 (150 μg of DNA); 100 μg of poly (U,G) (Miles Laboratories, Inc.); 0.15% sarkosyl NL97 (Geigy Industrial Chemicals Corp.); 0.003 M tris (hydroxymethyl)aminomethane-hydrochloride, pH 8.5; 1 mM ethylenediaminetetraacetic acid; and enough 0.1 SSC (SSC, 15 mM NaCl, 1.5 mM sodium citrate) to obtain a final volume of 7.0 ml. The mixture was immersed in boiling water for 5 min and then cooled rapidly. Approximately 9.5 g of CsCl (optical grade) was added, and the solution was adjusted to ρ_{25} = 1.4020 to 1.4025. A sample (0.35 ml) was removed for analytical ultracentrifugation (Beckman model E ultracentrifuge) for 20 h at 44,000 rpm and 25 C. The remaining volume of reaction mixture (6.65 ml) was centrifuged in a Beckman 50Ti fixed-angle rotor for 60 h at 40,000 rpm and 10 C. After centrifugation, 6-drop fractions were collected and mixed with 1 ml of 2× SSC. The absorbancy at 260 nm was measured for each fraction. Pooled fractions containing each of the two principal peaks were digested in 0.1 M KOH for 8 h at 37 C to hydrolyze the poly(U,G) and then were dialyzed extensively against 2× SSC containing 2 mM HCl to remove ribonucleotide residues. A final dialysis in 2× SSC was carried out.

Heteroduplex analysis using equal microgram quantities of R6-5 plasmid DNA and one of the two separated strands of λ DNA was carried out as described by Sharp et al. (21). Molecules were examined using a Philips EM201 electron microscope. Although DNA contour length measurements were found to be about 2% greater than those obtained by Hu et al. (11), our values were normalized to their values for comparison. The plasmid DNA map notations of Hu et al. (11) were also adopted.

RESULTS

Separation of λ DNA strands containing IS regions. Bacteriophage λ particles purified by cesium chloride gradient centrifugation were mixed with poly(U,G) in a ratio of 2 to 1 and subjected to heating at 95 C for 3 min in the presence of 0.15% sarkosyl NL97. An analytical cesium chloride gradient centrifugation (ρ=1.715 g/cm^3) pattern of the separated strands is shown in Fig. 1A. The preparative centrifugation pattern is shown in Fig. 1B. The shaded areas of the preparative gradient represent the fractions pooled as the heavy and light strands of λ DNA.

Heteroduplex analysis of IS3 regions on R6-5 plasmid. An electron photomicrograph of a representative heteroduplex between an intact R6-5 plasmid DNA molecule and the heavy strand of λ*plac*5S7-MS505, which carries the

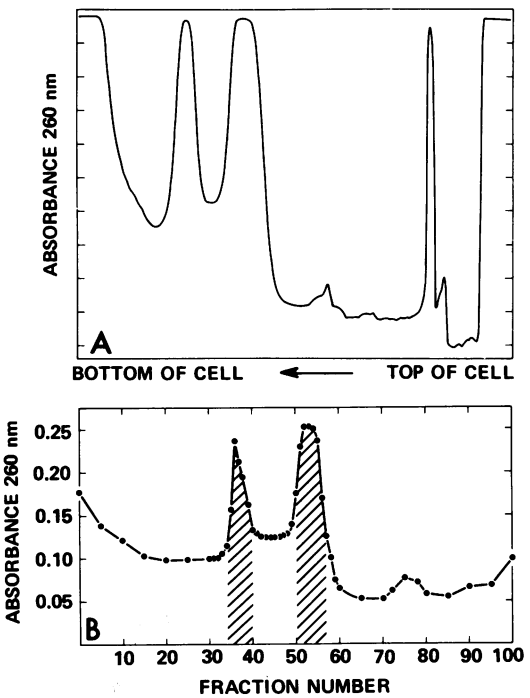

FIG. 1. *Analytic and preparative ultracentrifugation of λplac5S7-MS505 DNA in the presence of poly (U,G). Conditions are described in Materials and Methods. (A) Densitometer (Gilford) tracing of photograph taken during analytical ultracentrifugation. (B) Preparative separation of λplac5S7-MS505 DNA strands; denatured DNA and the poly(U,G) nucleotide phosphate ratio, 2:1 (ρ=1.715 g/cm^3) in CsCl solution were centrifuged in a 50Ti fixed-angle rotor for 60 h at 40,000 rpm and 10 C. Fractions 35 to 40 were pooled and designated as the "heavy" strand; fractions 50 to 57 were pooled and designated as the "light" strand.*

IS3 sequence, is shown in Fig. 2. The double-stranded region of homology between the λ-IS3 DNA strand and the plasmid molecule is identified by arrows (a) pointing to the ends of the double-stranded region. The precise site of homology between the strands of λ and plasmid DNA can be mapped in reference to a second inverted repeat sequence identified previously on the R6-5 molecule at map coordinates 21.8–24.9 (arrow b). The homology between λ and plasmid DNA occupies a position on the plasmid strand precisely at the site occupied by one of the limbs of the inverted repeat sequences which form the Tc resistance "stalk" of R6-5. The inset shows the region of homology diagramatically.

Identical procedures were used for separation and isolation of the strands of λr14 containing IS1 and for heteroduplex analysis of the λ-IS1 strands with the R6-5 plasmid. In confirmation of the results of Hu et al. (11), homology between IS1 and R6-5 was observed at two discrete sites on the R6-5 plasmid where IS1 exists as a directly duplicated DNA sequence.

A map of the R6-5 plasmid showing the location of the repeated DNA sequences identified as IS is shown in Fig. 3A. As seen in the map, IS3 is located at the site of the inversely repeated DNA sequences identified previously (21) to be the DNA segment believed to contain the genes for Tc resistance. Since the IS3 sequence of λ*plac*5S7-MS505 forms a duplex segment with one of the two repeated copies of IS3 included in the Tc resistance "stalk" of R6-5, inverted repeats are not seen at this site of the R6-5 molecule (Fig. 2). The length of the IS3 sequence on five heteroduplex molecules measured in our laboratory is 1.33 kb ± 0.08. This value agrees closely with the previous measurements of Sharp et al. (21) for this inverted

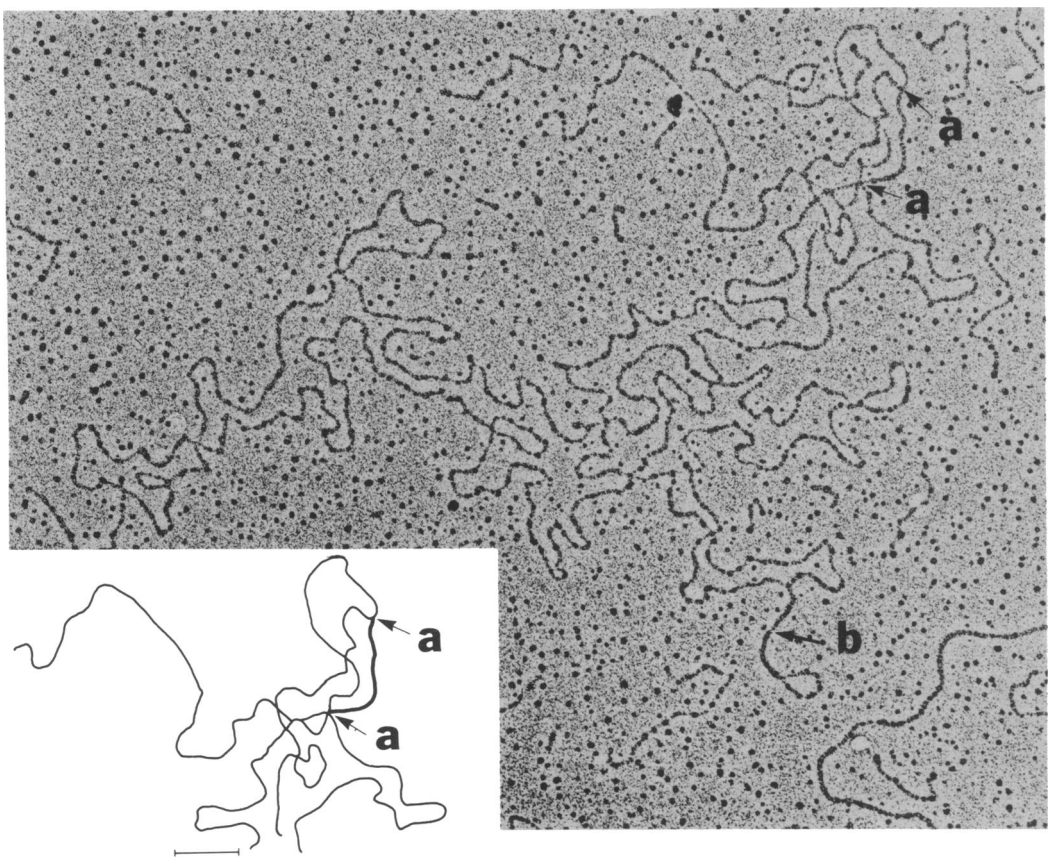

FIG. 2. *Electron micrograph of a heteroduplex of R6-5 plasmid DNA with λplac5S7-MS505 DNA carrying the IS3 sequence. The double-stranded region of homology between the λ-IS3 DNA strand and the R6-5 plasmid is identified by arrows (a). The inset shows the region of homology diagramatically. The precise site of the double-stranded region of homology between λ and plasmid DNA was mapped in reference to a second inverted repeat sequence identified previously (21) on the R6-5 molecule (arrow b). Procedures for heteroduplex formation are described in Materials and Methods. The scale represents 0.84kb.*

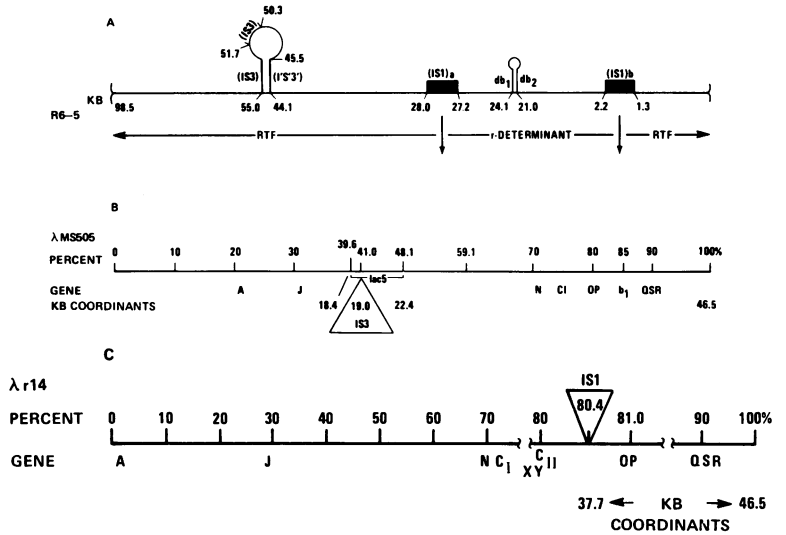

FIG. 3. *(A) Map of R6-5 plasmid showing the kilobase coordinates for the IS sequences identified by heteroduplex analysis. The plasmid notations described by Hu et al. (11) were adopted. (B) Map of λplac5S7-MS505 DNA locating the IS3 segment as 41% of 46.5 kb (14). (C) Map of λr14DNA locating the IS1 segment at 80.4% of 46.5 kb (16).*

repeat of R6-5, but is somewhat longer than the value reported by Malamy et al. (14) for the IS3 segment on λ DNA. Since the heteroduplex shown in Fig. 2B contains an intact strand of λMS505 DNA, it permits localization of the IS3 segment on the phage DNA molecule. This sequence is located at 41% of the 46.5-kb length of the λ DNA in agreement with the data of Malamy et al. (14) (Fig. 3B).

The R6-5 map shown in Fig. 3A also indicates that copies of the IS1 sequence are located between 1.9 to 2.2 kb and between 27.2 to 28.0 kb of the R6-5 molecule. Within experimental error, the IS1 sites map at the two junctions of the RTF and R-determinant components of the R6-5 plasmid. Since R6-5 is homologous with the R6, R1, and R100 plasmids at these map positions (21), as well as along substantial other portions of the RTF and R-determinant segments, IS1 must also occur at the junction of the RTF and R-determinant components of these other plasmids. The occurrence of IS1 at these map positions of R6, R1, and R100 has been shown directly by heteroduplex experiments carried out by Hu et al. (11). The position of IS1 on λr14 was determined as indicated above for IS3 and is shown at 80.4% of the 46.5-kb length of this phage DNA (Fig. 3C).

DISCUSSION

The present investigations provide evidence that genetically characterized authentic IS sequences are present on bacterial antibiotic resistance plasmids as both direct and inverted DNA nucleotide sequence duplications. IS3 occurs as both a direct duplication and an inverted repeat in the region of the R6-5 plasmid believed to contain the gene for tetracycline resistance (21). IS1 occurs as a direct repeat at both junctions of the RTF and R-determinant components of R6-5 and at the same sites on the related plasmids R6, R1, and R100. One of the two directly repeated copies of the IS3 sequence (map position, 50.3 to 51.7 kb) appears to inhibit expression of the Tc resistance gene carried by the R6-5 plasmid in much the same way that the various IS sequences have been shown previously to affect expression of the *E. coli gal* and *lac* operons. Loss of this sequence results in a Tc-resistant revertant of the R6-5 plasmid that is phenotypically and structurally identical to the R6 parent.

Earlier investigations carried out in our laboratory and elsewhere have indicated that, in *Proteus mirabilis*, the RTF and R-determinant segments of the R factors R1, R6, and R100 can undergo reversible dissociation into separate replicons (5, 8, 13, 17). The observed structural and genetic similarity of RTF plasmids which have been derived from separate clones of the same co-integrate R factor suggests that dissociation of such R factors occurs at discrete plasmid sites. On the basis of data obtained in similar experiments, Rownd et al. (18) have disputed the interpretation that substantial dissociation of R factors into separate RTF and R-determinant units occurs in *P. mirabilis*. According to the view put forth by Rownd (18),

the RTF unit ordinarily does not exist as an autonomous replicon in *Proteus*, but instead is always associated with one or more copies of the R-determinant segment; growth in the presence of antibiotics results in a "transition" from a plasmid molecule containing single copies of both the RTF and R-determinant units to a form containing poly R-determinant copies in covalent linkage with a single RTF.

The identification of directly repeated IS sequences at the two junctions of the RTF and R-determinant components of co-integrate R plasmids provides a possible mechanism for reversible dissociation of such plasmids into separate component segments by intramolecular recombination. A model for such reciprocal recombination at the ends of the two IS1 sequences that form boundaries between the R-plasmid components is shown in Fig. 4. Moreover, recombination between IS1 sequences could also yield molecules containing a series of tandemly repeated R-determinant units (Fig. 4B); the number of R-determinant copies included on each of the dissociated component replicons would depend on which two of the IS1 regions participate in the recombinational event. Recombination between the most distant pair of IS units in Fig. 4B would lead to formation of a discrete and autonomous RTF replicon (6, 22) plus a poly R-determinant plasmid, whereas recombination between intermediately situated IS1 units would yield a DNA molecule containing both an RTF segment and a varying number of R-determinant copies (18). Thus, separate RTF and R-determinant replicons (Fig. 4A), as well as plasmids containing multiple R-determinant copies (Fig. 4B), could be produced by the same recombinational process.

Taken together, the experiments described in the present report and in the accompanying paper by Hu et al. (11) indicate that at least three different types of genetically characterized IS regions occur on plasmids as directly or inversely repeated DNA sequence duplications. The regulation of expression of Tc resistance by IS3 on the R6-5 plasmid is consistent with the

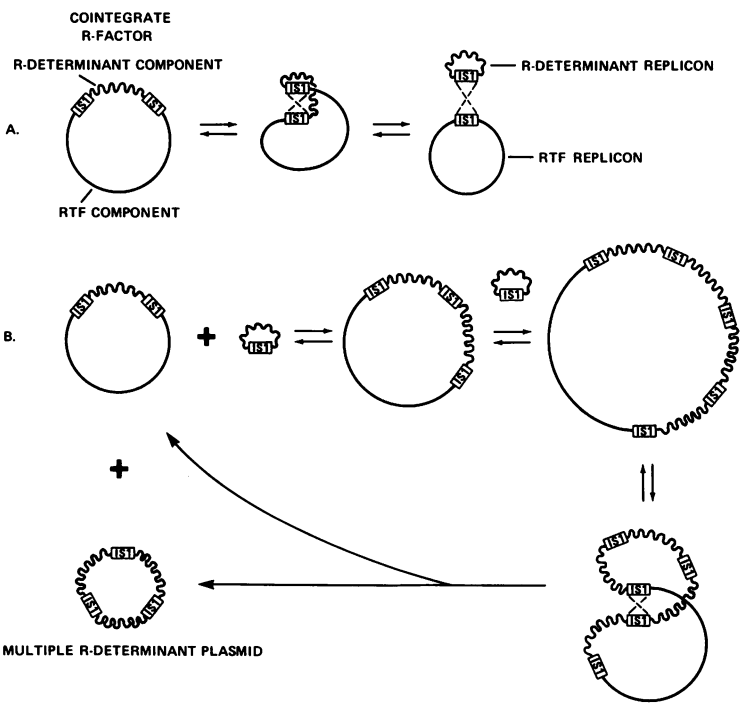

FIG. 4. *(A) Possible mechanism for reversible dissociation of RTF and R-determinant units of co-integrate R factors by intramolecular recombination between IS1 segments. Dissociation of R-factors into RTF and R-determinant components is based on data obtained in P. mirabilis and presented in ref. 8, 12, 13, 18. (B) Possible mechanism for accumulation of multiple R-determinant copies in covalent linkage with a single RTF unit, as proposed by Rownd (18). The model illustrated also provides a possible mechanism for the formation of multiple R-determinant plasmids by recombination between IS1 segments.*

previously demonstrated genetic role of insertion sequences in other bacterial and phage systems. In addition, it now appears that IS regions may also have an important role in recombination between, and the evolution of, complex plasmid genomes.

ACKNOWLEDGMENTS

We gratefully acknowledge gifts of λ phage strains containing genetically characterized IS regions as noted in Materials and Methods. We also acknowledge a free exchange of unpublished information with S. Hu and N. Davidson in the course of these experiments, and the helpful advice of these investigators and J. Zisslar.

Our studies were supported by Public Health Service grant Al 08619 from the National Institute of Allergy and Infectious Diseases, by grant GB-30581 from the National Science Foundation, and by grant VC-139 from the American Cancer Society. Kay Ptashne is the recipient of a Public Health Service Postdoctoral Research Fellowship 5F02ES54436 from the Division of Environmental Health Sciences.

LITERATURE CITED

1. Adelberg, E. A., and P. Berquist. 1972. The stabilization of episomal integration by genetic inversion: a general hypothesis. Proc. Natl. Acad. Sci. U.S.A. **69**:2061-2065.
2. Cohen, S. N., and A. C. Y. Chang. 1970. Genetic expression in bacteriophage λ. III. Inhibition of *Escherichia coli* nucleic acid and protein synthesis during λ development. J. Mol. Biol. **49**:557-575.
3. Cohen, S. N., and A. C. Y. Chang. 1973. Recircularization and autonomous replication of a sheared R-factor. DNA segment in *Escherichia coli* transformants. Proc. Natl. Acad. Sci. U.S.A. **70**:1293-1297.
4. Cohen, S. N., and J. Hurwitz. 1967. Transcription of complementary strands of phage λDNA *in vivo* and *in vitro*. Proc. Natl. Acad. Sci. U.S.A. **57**:1759-1766.
5. Cohen, S. N., and C. A. Miller. 1970. Non-chromosomal antibiotic resistance in bacteria: II. Molecular nature of R factors isolated from *Proteus mirabilis* and *Escherichia coli*. J. Mol. Biol. **50**:671-687.
6. Cohen, S. N., and C. A. Miller. 1970. Non-chromosomal antibiotic resistance in bacteria: III. Isolation of the discrete transfer unit of the R-factor R1. Proc. Natl. Acad. Sci. **67**:510-516.
7. Gilbert, W., and A. Maxam. 1973. The nucleotide sequence of the *lac* operon. Proc. Natl. Acad. Sci. U.S.A. **70**:3581-3584.
8. Haapala, D. K., and S. Falkow. 1971. Physical studies of the drug-resistance factor in *Proteus*. J. Bacteriol. **106**:294-295.
9. Hirsch, H. J., P. Sarlinger, and P. Bracket. 1972. Two kinds of insertions in bacterial genes. Mol. Gen. Genet. **119**:191-206.
10. Hsu, M., and N. Davidson. 1974. Electron microscope heteroduplex study of the heterogeneity of mu phage and prophage DNA. Virology **58**:229-239.
11. Hu, S., E. Ohtsubo, N. Davidson, and H. Saedler. 1975. Electron microscope heteroduplex studies of sequence relations among bacterial plasmids: identification and mapping of the insertion sequences IS1 and IS2 in F and R plasmids. J. Bacteriol. **122**:764-775.
12. Kelly, T. J., and H. O. Smith. 1970. A restriction enzyme from *Hemophilus influenza*. II. Base sequence of the recognition site. J. Mol. Biol. **51**:393-409.

12a. Kopecko, D. J., and S. N. Cohen. 1975. Site-specific *recA*-independent recombination between bacterial plasmids. Proc. Natl. Acad. Sci. U.S.A. **72**:1373-1377.

13. Kopecko, D. J., and J. D. Punch. 1971. Regulation of R factor replication in *Proteus mirabilis*. Ann. N.Y. Acad. Sci. **182**:207-216.
14. Malamy, M. H., M. Fiandt, and W. Szybalski. 1972. Electron microscopy of polar insertions in the *lac* operon of *Escherichi coli*. Mol. Gen. Genet. **119**:207-222.
15. Maniatis, T., M. Ptashne, B. G. Barrell, and J. Donelson. 1974. Sequence of a repressor binding site in the DNA of bacteriophage λ. Nature (London) **250**:394-397.
16. Michaelis, G., H. Saedler, P. Venkov, and P. Starlinger. 1969. Two insertions in the galactose operon having different sizes but homologous DNA sequences. Mol. Gen. Genet. **104**:371-397.
17. Nisioka, T., M. Mitani, and R. Clowes. 1969. Composite circular forms of R factor deoxyribonucleic acid molecules. J. Bacteriol. **97**:376-385.
18. Rownd, R. H., D. Perlman, and N. Goto. 1975. Structure and replication of R-factor DNA in *Proteus mirabilis*, p. 76-94. *In* D. Schlessinger (ed.), Microbiology—1974. American Society for Microbiology, Washington, D.C.
19. Saedler, H., and B. Heiss. 1973. Multiple copies of the insertion-DNA sequences IS1 and IS2 in the chromosome of *E. coli* K-12. Mol. Gen. Genet. **122**:267-277.
20. Saedler, H., H. J. Reif, S. Hu, and N. Davidson. 1974. IS2. A genetic element for turn-off and turn-on of gene activity in *E. coli*. Mol. Gen. Genet. **132**:265-289.
21. Sharp, P. A., S. N. Cohen, and N. Davidson. 1973. Electron microscope heteroduplex studies of sequence relations among plasmids of Escherichia coli. II. Structure of drug resistance (R) factors and F factors. J. Mol. Biol. **75**:235-255.
22. Silver, R. P., and S. Falkow. 1970. Studies on resistance transfer factor deoxyribonucleic acid in *Escherichia coli*. J. Bacteriol. **104**:340-344.
23. Szybalski, W., H. Kubinski, Z. Hradecna, and W. C. Summers. 1971. Analytical and preparative separation of the complementary DNA strands, p. 383-413. *In* L. Grossman and K. Moldave (ed.), Methods in enzymology, vol. 21, part D. Academic Press Inc., New York.
24. Taylor, K., Z. Hradecna, and W. Szybalski. 1967. Asymmetric distribution of the transcribing regions on the complementary strands of coliphage λDNA. Proc. Natl. Acad. Sci. U.S.A. **57**:1618-1625.
25. van Embden, J., and S. N. Cohen. 1973. Molecular and genetic studies of an R factor system consisting of independent transfer and drug resistance plasmids. J. Bacteriol. **116**:699-709.
26. Wilson, D. A., and C. A. Thomas, Jr. 1974. Palindromes in chromosomes. J. Mol. Biol. **84**:115-144.