# 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[1]

SYLVIA HU, EIICHI OHTSUBO,[2] NORMAN DAVIDSON,* AND HEINZ SAEDLER

*Department of Chemistry, California Institute of Technology, Pasadena, California 91125, and Institüt für Genetik, Universität zu Köln, Cologne, West Germany*

Heteroduplex experiments between the plasmid R6 and one strand of the deoxyribonucleic acid (DNA) of a λ phage carrying the insertion sequence IS1 show that IS1 occurs on R6 at the two previously mapped junctions of resistance transfer factor (RTF) DNA with R-determinant DNA. From previous heteroduplex experiments, it then follows that IS1 occurs at the same junctions in R6-5, R100-1, and R1 plasmids. Heteroduplex experiments with the DNA from a λ phage carrying the insertion sequence IS2 show that one copy of IS2 occurs in R6, R6-5, and R100-1 (but not R1) at a point within the RTF with coordinates 67.5 to 68.9 kilobase units (kb). In an accompanying paper, Ptashne and Cohen (1975) show that the insertion sequence IS3 occurs on R6 and R6-5. R100-25, a *traC* mutant, differs from its parent R100-1 only in that it contains an additional copy of IS1 inserted within the *tra* gene region of 82.1 kb. R100-31, a *traX*, Tc^s mutant of R100-1, is deleted in R100-1 sequences starting at one of the IS3 termini (46.9 kb) and extending with RTF to 61.0 kb. Heteroduplex studies of F plasmids with the DNA of a λ phage bearing insertion sequence IS2 show that the sequence of F with coordinates 16.3–17.6F is IS2. The occurrence of IS1 at the two junctions of R-determinant DNA and RTF DNA in R plasmids provides a structural basis to explain the mechanism of the previously observed formation of molecules containing one RTF unit and several tandem copies of the R-determinant unit, when R plasmids in *Proteus mirabilis* are grown in the presence of antibiotics, and the segregation of an R plasmid into an RTF unit and an R-determinant unit. In general, correlation of our results with previous studies shows that insertion sequences play a role in a variety of F- and R-related intra- and intermolecular recombination phenomena.

Insertion sequences that cause strongly polar mutations have been identified in *Escherichia coli* and coliphages (5, 14, 16, 28, 31; for a review, 36). The present study is particularly concerned with IS1 and IS2, which have lengths of about 800 and 1,400 nucleotide pairs, respectively (8–10, 17, 30). Hybridization studies have indicated that multiple copies of both IS1 and IS2 occur in the chromosome of *E. coli* K-12 (29), that IS2 occurs in the F plasmid (29), and that IS1 occurs in some R plasmids (H. Saedler, personal communication). (Further work has shown that the previous indication [29] that

there is a low level of hybridization of λ *pgal*::IS1 DNA to F plasmids is due to nonspecific hybridization of *gal* sequences; we believe that IS1 does not occur on F [H. Saedler, personal communication]).

In the work reported here, we have applied the electron microscope heteroduplex method to test for the presence of and, where present, to map IS1 and IS2 on F and on several R plasmids. In an accompanying paper, Ptashne and Cohen (24) report their discovery that IS3 occurs in R plasmids and also report their independent observations on the occurrence of IS1 in R plasmids.

The basic idea of the experiment was to attempt to prepare a heteroduplex between a strand of λ deoxyribonucleic acid (DNA) or a λ *gal* DNA bearing an IS insertion with a strand

of a suitable F′ or R plasmid. We are seeking to see heteroduplexes involving a region of homology 800 or 1,400 nucleotide pairs in length. The F′ and R plasmids have lengths of about 90,000 nucleotide pairs and the λ DNA molecules have lengths of about 46,500 nucleotide pairs. Therefore, if one starts with a mixture of equal amounts of the two duplex DNAs, dissociates and reassociates, homoduplex formation will be much, much faster than any heteroduplex formation. Preparations of separated complementary strands of the λ DNAs were available; preparations of separated complementary strands of F′ or R plasmids were not readily available. Heteroduplexes of interest were therefore prepared using a preparation of one of the complementary strands of the λ DNA and a mixture of the two strands of the F′ or R plasmid DNA.

The results of a conceptually independent investigation of the properties of several plasmids derived by P1kc transduction of the plasmid R100-1 also provide information on the occurrence and properties of IS sequences on R plasmids. These results are also reported here.

## MATERIALS AND METHODS

**Strains, plasmids, and bacteriophage.** Bacterial strains and plasmids are listed in Table 1. Bacteriophage were prepared from the lysogens listed in Table 2. Transducing phage were carried as double lysogens with a helper phage.

**Preparation of plasmid DNA.** Bacteria were grown, and closed circular plasmid DNA was extracted and converted to the singly nicked form by

X-rays as described previously (21, 34).

**Preparation of phage lysates.** See Michaelis et al. (19) and Ohtsubo et al. (22) for preparation of phage lysates.

**Preparation of separated strands of phage DNA.** This was accomplished as described by Hradecna and Szybalski (11), except that DNA was denatured by alkali as suggested by Shapiro et al. (32).

**Heteroduplex preparation.** DNA was liberated from bacteriophage by incubation of a small volume of the banded bacteriophage suspension in 20 μl of 0.2 M disodium ethylenediaminetetraacetate, pH 8.5, at room temperature for 5 min. For phage heteroduplexes, the solution contained equal amounts (in micrograms) of the two kinds of DNA molecules. For the preparation of heteroduplexes of plasmids with separated strands of phage DNA, proper amounts of plasmid DNA from the CsCl-EtBr band were mixed with a fivefold excess of the phage DNA strands, and $H_2O$ was added to give a volume of 60 μl. A 20-μl amount of 1.0 N NaOH was added to give an OH⁻ concentration of 0.25 M; after 4 min at room temperature this solution was neutralized with 20 μl of 1.0 M tris(hydroxymethyl)aminomethane-hydrochloride followed by 100 μl of 0.2 M disodium ethylenediaminetetraacetate, pH 8.5. The total concentration of DNA in this renaturation solution was normally 0.5 to 1.0 μg/ml; it was raised to 2 to 3 μg/ml for the reassociation of plasmid DNA with separated strands of phage DNA. This solution was dialyzed into the standard renaturation buffer as described previously; other aspects of our electron microscope technique are all as described previously (21, 34).

## RESULTS

**IS1 in R6.** Membrane filter DNA-DNA hybridization studies have shown that several copies of IS1 are present in the F-like plasmid

TABLE 1. *Bacterial strains and plasmids*

| Strain | Plasmid | Plasmid genotype[a] | Chromosomal genotype | Source and/or reference |
|---|---|---|---|---|
| CSH2(R6) | R6 | Tc Sm Su (Km Nm)Cm | Prototrophic | S. N. Cohen (33) |
| W1485(R6-5) | R6-5 | Sm Su (Km Nm)Cm) | Prototrophic | S. N. Cohen (33) |
| JE513 | F13-4 | *lac⁺* | *lac purB str^r* | Y. Sugino (37) |
| ND3 | FΔ(0-15) | | *gal6 trp arg recA str^r* | (34) |
| | R100-31 | Sm Cm Su *traX⁻* | | (23) |
| | R100-25 | Sm Cm Tc Su *traC⁻* | | (23) |

a Abbreviations: Tc, tetracycline; Sm, streptomycin; Su, sulfonamide; Km, kanamycin; Nm, neomycin; Cm, chloramphenicol.

TABLE 2. *Bacteriophage*

| Strain | Helper phage | Phage or interest | Lysogen genotype | Source and/or references |
|---|---|---|---|---|
| S165 | λCI857t68 | λCI857 t68 dgOP::IS2-308 (λdgal OP308) | Δ *gal his* Sm^r | (9, 10) |
| C600(λCI857 NNr32) | | λCI857 NNr32(IS2) (λr32) | F⁻ *thr leu lac su_{II}⁺* | (5) |
| C600(λCI857 NNr14) | | λCI857 NNr14(IS1) (λr14) | F⁻ *thr leu lac su_{II}⁺* | (5) |
| ND14 | φ80i^λ(I) | φ80darg(G) | MN42λ⁻ [HfrP4X Δ(ppc⁻ argECBH)λ⁻] | Glansdorffand and Maas (22) |

R1-*drd*19, a derepressed mutant of R1 (H. Saedler, personal communication).

Heteroduplex studies show that there is extensive homology between the F-like R factors R1, R6, R6-5, and R100-1 (33). Furthermore, it should be recalled that all these plasmids carry a resistance transfer factor (RTF) sequence of length approximately 70 kilobase units (kb) and an R-determinant segment approximately 20 to 25 kb in length. The RTF segment primarily contains genes that code for conjugal transfer and for autonomous replication, although it is suspected that a tetracycline (Tc) resistance gene is carried in this region (33, 38). The R-determinant region contains all of the other known genes for antibiotic resistance. A part of the RTF segment is homologous to that part of F which codes for *tra* genes (33, 34).

A map of R6, including the positions of the RTF sequences, is shown in Fig. 1a. The system for assigning coordinates to the several R plasmids proposed by Sharp et al. (33) is explained in the legend to Fig. 1. A single strand of R6 contains two duplex regions due to rapid intramolecular reassociation (fold-back) of inverted repeat sequences. In each case, a short single-strand loop separates the complementary sequences. The notation introduced in Fig. 1 for inverted repeat structures is as follows. For each inverted repeat, the duplex stalks are denoted by d. The two different inverted repeat structures on R6 are denoted by subscripts a and b, respectively; they differ in both length and base sequences. For any one inverted repeat, the two complementary sequences that reassociate to form the duplex stalk are denoted by numerical subscripts, for example, $d_{a1}$ and $d'_{a2}$. The prime indicates that $d'_{a2}$ is inverted and complementary to $d_{a1}$. Similarly, $d_{b1}$ and $d'_{b2}$ are mutually complementary sequences on a single strand of R6. The inverted repeat features provide reference points on R6 for heteroduplex analysis; for example, the coordinates on R6 of an IS duplex region in a heteroduplex of R6 with an λ DNA can be measured by measuring the distances to these inverted repeat features.

The DNA of phage λr14 contains an IS1 sequence inserted into the *xycIIOP* operon (5, 10). Its structure is shown in Fig. 1b.

Relaxed circular duplex R6 DNA and an excess of heavy (H) strands of λr14 were treated with NaOH to denature the R6 and neutralized and incubated under renaturing conditions as described in Materials and Methods. The reaction conditions were such that about 50% of the R6 strands reassociated to homoduplexes. Approximately 1% of the R6 strands were seen as R6-λr14 heteroduplexes. The duplex region was

the IS1 sequence, as shown by its length and its position on the λr14 strand. IS1 maps at the two different sites shown in Fig. 1c on R6. (Because of the low probability of heteroduplex formation, only one site was occupied in most of the heteroduplexes seen.) Electron micrographs are shown in Fig. 2 and 3. Quantitative distance measurements show that, within experimental error, the two IS1 sites map at the two junctions of RTF DNA with R-determinant DNA on R6. Previous studies have shown that these junction regions are identical by the heteroduplex criterion in R6, R6-5, R100-1, and R1 (33); thus, within experimental error, IS1 occurs at the two junctions for all of these R plasmids.

**IS2 occurs on R6-5, R6, and R100-1.** The inverted repeat sequence $d_a$ of Fig. 1 has the same length as the IS2 sequence. $d_a$ occurs twice
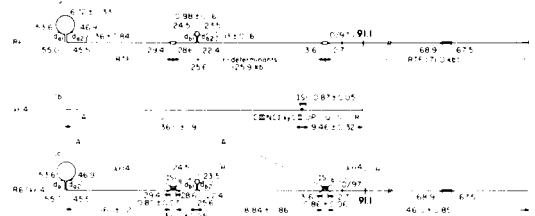


FIG. 1. *Structures of R6, λr14, and of the heteroduplex between them. The positions of the two IS1 sequences on R6 were determined by observing where the IS1 sequences of λr14 hybridized to R6, as indicated in (c). All coordinates and dimensions are given in kilobase units (one kilobase is 1,000 nucleotide or nucleotide pairs for single- or double-stranded DNA, respectively). In order to conserve space, the circular map of R6 is represented in a linear form. The assignment of coordinates on R6 (and on other R plasmids) is based on the system introduced by Sharp et al. (33). A particular point which is present on all of the R factors has been assigned the coordinate 91.1 kb. This is the point at which there is a readily recognizable insertion loop in the heteroduplex between any of the R plasmids and F8-33. The polarity of the coordinate system is also determined from the previously assigned polarity for F and the heteroduplexes between the R and F plasmids. The origin and terminus of the circular map is then determined by the molecular length of the particular plasmid; for example, for R6 this point is 0/97.1 kb. Since different R plasmids have different molecular lengths and different insertions and deletions relative to one another, each R plasmid will have a unique set of coordinates. Thus, the same sequence present on two different R plasmids may have different coordinates. Figure 11, in effect, permits one to correlate coordinates of homologous points on different plasmids. These correlations are based on the heteroduplex studies of Sharp et al. (33). In the heteroduplexes, R6/λr14, the two inverted repeats on R6, $d_{a1}/d'_{a2}$ and $d_{b1}/d'_{b2}$, provided internal reference features for mapping the positions of the two IS1 duplex regions.*
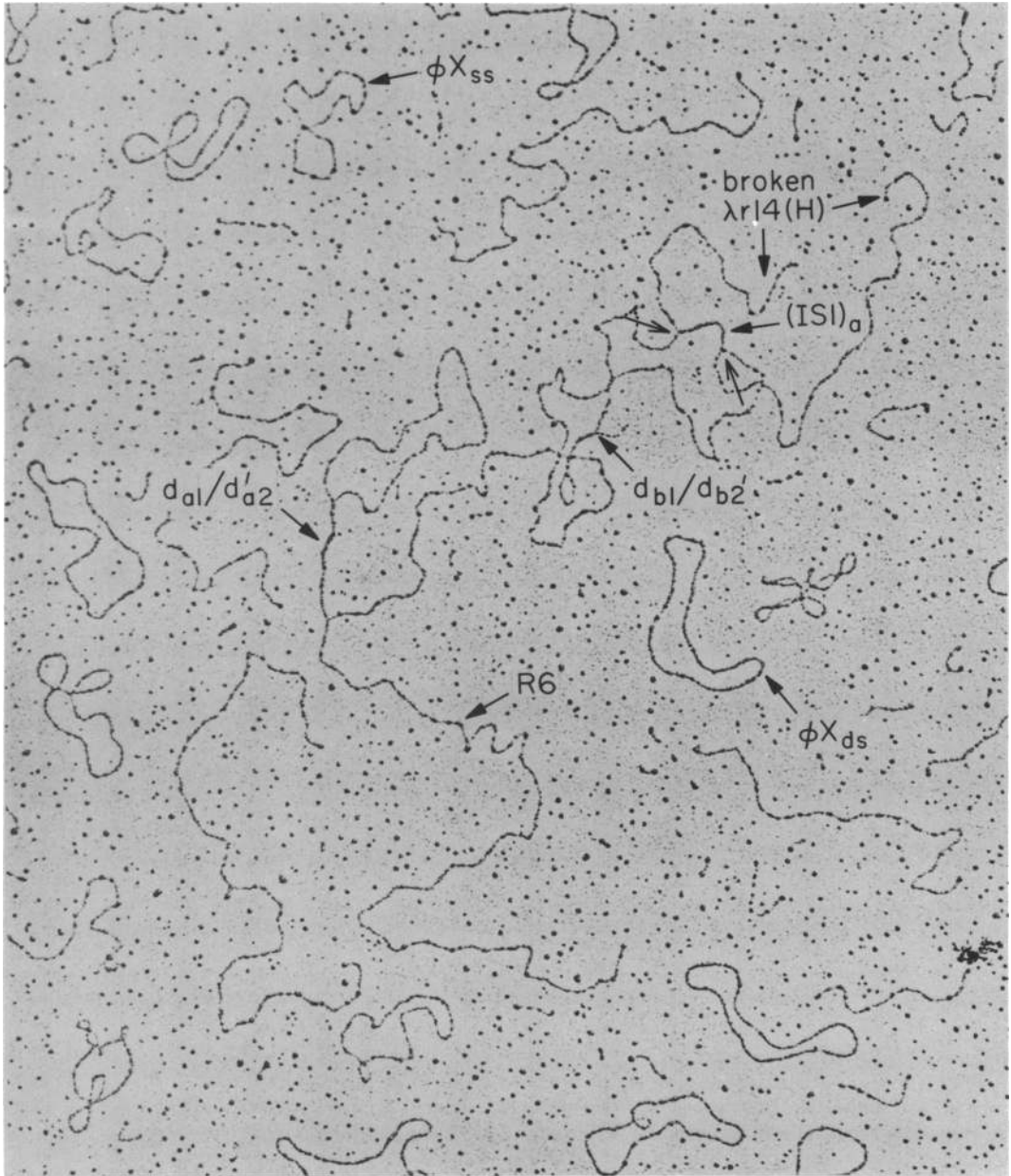
FIG. 2. *Electron micrograph of a heteroduplex of R6 with the separated heavy strand of λr14 DNA. There is a duplex region due to a hybrid of the IS1 sequence on λr14 with the $(IS1)_a$ sequence of R6. The position of this feature relative to the reference features of R6, the $d_{a1}/d'_{a2}$ and $d_{b1}/d'_{b2}$ inverted repeats, can be measured. The arrow labeled $d_{b1}/d'_{b2}$ identifies the points 22.4 and 25.6 kb on R6 (Fig. 1). The single strand of R6 DNA in this micrograph is broken and ends at the coordinate of approximately 22.0 kb. In this and other micrographs, single- and double-stranded φX DNA are present as length standards.*

on R6. R6-5 is a spontaneous $Tc^s$ mutant of R6. It contains a third copy of $d_a$, inserted at the position shown in Fig. 4a and denoted as $d_{a3}$. Sharp et al. (33) suggested that the insertion was actually in the Tc gene, thus causing its inactivation. It was observed, in agreement with this hypothesis, that in a spontaneous revertant to the $Tc^r$ genotype $d_{a3}$ had again disappeared. Thus, in these respects $d_a$ behaves like an insertion sequence. Because of its length, we
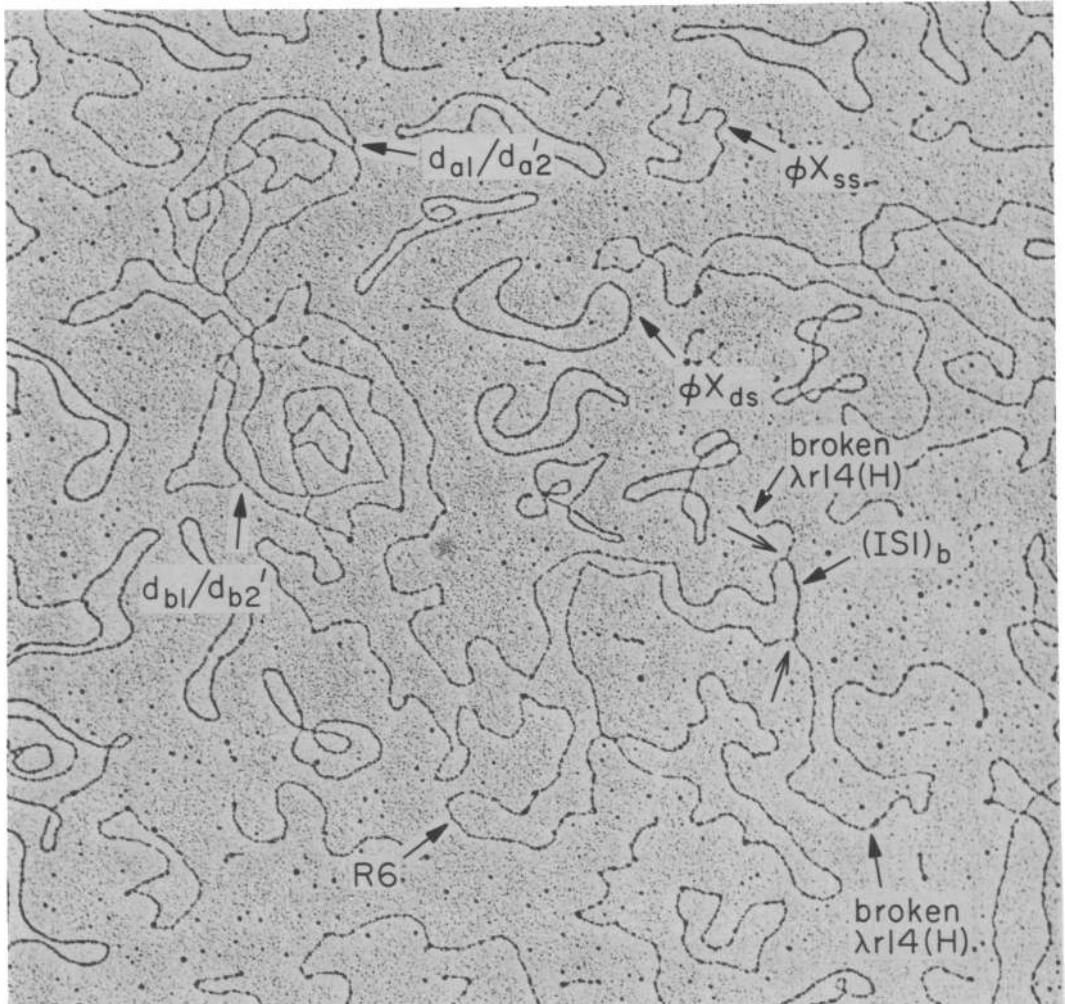
FIG. 3. *Heteroduplex of R6 with λr14 H strand DNA. In this case the IS1 sequence of λr14 DNA forms a duplex with the (IS1)$_b$ sequence of R6. In the present example a complete circular R6 single strand is present.*

suspected that it might be IS2.

The formation of a heteroduplex of one of the two $d_a$ sequences in R6 with a phage strand bearing its complement would be relatively improbable; $d_{a1}$ and $d'_{a2}$ associate to form a duplex by a fast intramolecular reaction. However, there is always one $d_a$ sequence in the single-strand state in a strand of R6-5. Therefore, we attempted to make heteroduplexes of R6-5 with λr32, which carries IS2 in the position shown in Fig. 4b (5, 10). We did not observe any heteroduplexes with λr32 hybridized to the $d_a$ sequence of R6-5; however, we did observe heteroduplexes at the position within the RTF shown in Fig. 4c. The studies by Sharp et al. (33) show that this particular sequence, now

identified as IS2, occurs in R6, R6-5, and R100-1, but is missing in R1.

**Insertion sequences on R100-1 derivatives.** R100-31 and R100-25 are transfer-defective (*tra*) mutants of R100-1 produced by P1kc transduction (23). A study by one of us (E. O.) on the structures of these two plasmids is described here because the results for R100-25 confirm the direct heteroduplex studies on the mapping of IS1 in R plasmids and those for R100-31 provide additional interesting information on the properties of the sequence $d_a$ (cf. Fig. 4).

The final structures of R100-1, as previously determined (33), and of R100-25 and R100-31, as determined in the present investigation, are shown in Fig. 5.

FIG. 4. Structures of R6-5, λr32 DNA, and the heteroduplex between them. The main point is that λr32 DNA contains an insertion of IS2 at the point indicated, so that the heteroduplex R6-5/λr32 identifies the position and orientation of the IS2 sequence on R6-5 and R6. It should be recalled that IS2 in λr32 is a strongly polar mutation preventing rightward transcription. It should also be recalled that R6-5 differs from R6 only in that the former contains the insertion $d_{a3}$ at the position indicated in the large inverted repeat loop. In all other respects R6-5 and R6 are identical; therefore, the positions of the IS1 sequences on R6-5 are known from the positions of these sequences on R6.
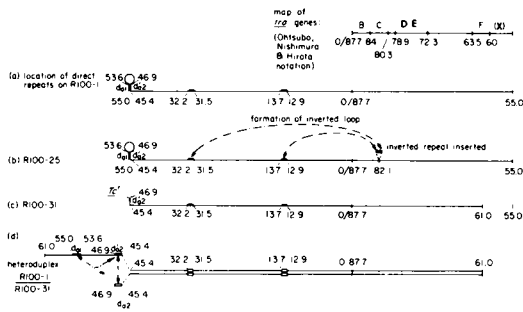


FIG. 5. Structures of (a) R100-1, (b) R100-25, and (c) R100-31. In R100-25 an additional copy of IS1 occurs in an inverted order at coordinate 82.1. It forms inverted repeat duplexes with either the IS1 sequences at 12.9-13.7 or the IS1 at 31.5-32.2. This additional IS1 occurs in the region of the traC gene, using the notation of Ohtsubo et al. (23) and the mapping data of Sharp et al. (34). R100-31 is a deletion variant of R100-1. The deleted sequences are indicated by the dotted lines in (c). Two observed alternate structures for the heteroduplex R100-1/R100-31 are depicted in (d). In one, the inverted repeat of $d_{a1}$ with $d'_{a2}$ of R100-1 is not seen, because $d'_{a2}$ is hybridized to $d_{a2}$ of R100-31. The sequences from 46.9 to 61.0 of R100-1 are seen as a deletion loop. This shows that R100-31 contains $d_{a2}$. In the other observed structure, the inverted repeat is seen, followed by an additional loop (55.0 to 61.0), and a 1.3-kb loop ($d_{a2}$ of R100-31) is seen on the other strand. This shows that the deletion in R100-31 starts at (within experimental error) the left side of $d_{a2}$.

R100-25 is defective in the tra gene denoted as traC by Ohtsubo et al. (23) and as traG by Achtman et al. (1, 3) (see Fig. 5; for further

details, see Fig. 3 of ref. 7). A heteroduplex study of R100-1/R100-25 shows that the latter is completely homologous with R100-1 except that it carries an insertion 760 ± 120 nucleotides in length. From a study of the two inverted repeat-loop structures seen in single strands of R100-25 (Fig. 5b), this insertion maps at R100-1 coordinate 82.1 and is an inverted repeat of a sequence which occurs in duplicate in a direct order on R100-1 with coordinates 12.9-13.7 and 31.5-32.2. This region of the chromosome is homologous between R100-1, R6, and R6-5 (33). Therefore, we now recognize that this repeated sequence on R100-1 is IS1. Therefore the insertion at 82.1 kb on R100-25 responsible for the traC (G) defect is IS1, inserted with a polarity opposite to that of the two IS1 sequences of R6, R6-5, and R100-1. The map position of the insertion at 82.1 was confirmed by a study of the heteroduplex of R100-25 with R100-31, whose structure is described below.

A study of the heteroduplexes of R100-1 with R100-31 shows that the latter is a deletion variant of R100-1. This deletion extends from one end of the $d'_{a2}$ segment (R100-1 and R6 coordinates of 46.9) clockwise to 61.0 kb (Fig. 5). Several heteroduplex studies established this conclusion; the simplest and most direct are the two structures for the R100-31/R100-1 heteroduplex shown in Fig. 5d and discussed in the legend. R100-31 has lost the Tc$^r$ property and is defective in traX; the simplest interpretation of these facts is that these genes map in the interval 46.9 to 61.0 kb. The traX cistron was known to lie clockwise of 60 kb (R100-1 coordinates) but had not otherwise been mapped (23, 24). Sharp et al. (33) inferred from the position of the Tc-inactivating insertion in R6-5 that the Tc gene lies in the R6 loop from 46.9 to 53.6 (Fig. 1), which is consistent with the structure for R100-31.

**Insertion sequence on F.** We chose to use the plasmid F13-4 in our initial search for IS sequences on F. Its structure as determined by Hu et al. (12) is shown in Fig. 6a. F13-4 is deleted in the sequences of F extending from
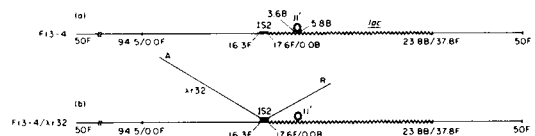


FIG. 6. Structure of F13-4 and of its heteroduplex with λr32 DNA. This heteroduplex shows that the 16.3-17.6 sequence of F is IS2. The small inverted repeat structure jj' of F13-4 is very useful as a reference point for this mapping.

17.6 to 37.8 F. The chromosomal sequences of F13–4, which have a length of 23.8 kb and include the *lac* region, contain an inverted repeat structural feature denoted jj'. This feature is useful as a reference point for mapping duplex regions in heteroduplexes. An electron micrograph illustrative of the jj' feature is shown in Fig. 7.

λr32 DNA, which carries IS2 in the *xycIIOP* operon (5, 10), formed the heteroduplex shown in Fig. 6b with F13–4. A micrograph is shown in Fig. 8. The duplex region is due to the IS2 sequence. Its distance from the jj' feature indicates that IS2 either is a chromosomal sequence on F13–4 with coordinates 9.4–10.7 B or is the F sequence 16.3–17.6 F, which occurs at the

junction between F DNA and chromosomal DNA in F13–4. A heteroduplex was therefore prepared with the plasmid FΔ(0–15) which contains only F sequences but is deleted in the segment 0–15 F. Again, a hybrid of the IS2 region of λr32 was seen. This still leaves open the possibility that there is an IS2 sequence on the F segment 17.6–37.8, which is deleted in F13–4, and a copy of IS2 on the chromosomal sequences of F13–4 at 9.4–10.7 B.

However, as explained below, there were reasons to suspect that 16.3–17.6 F was an insertion sequence of some sort. Previous studies (22) had shown that the phage DNA φ80d*arg*(G) contains the 16.3–17.6 F sequence in the position shown in Fig. 9a. Heteroduplexes
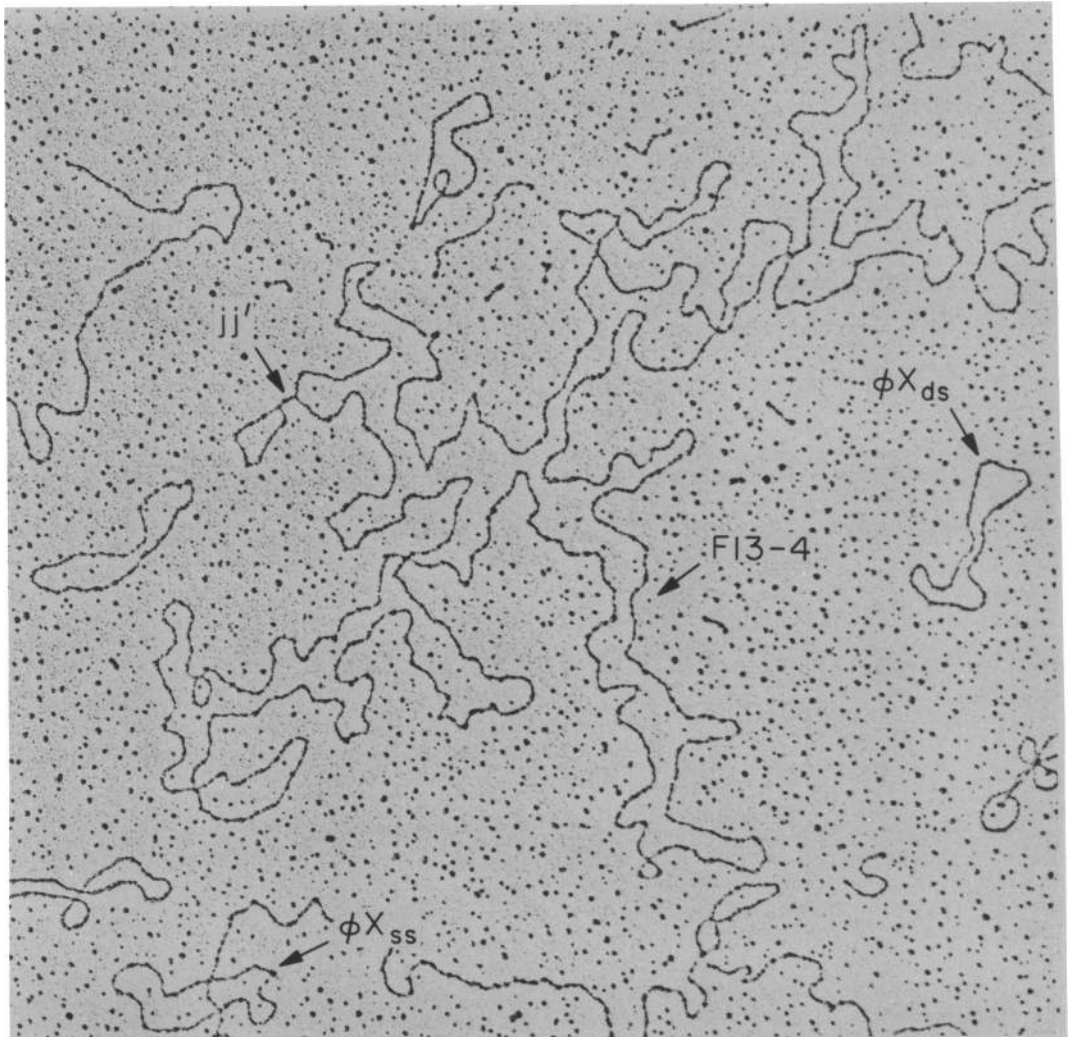


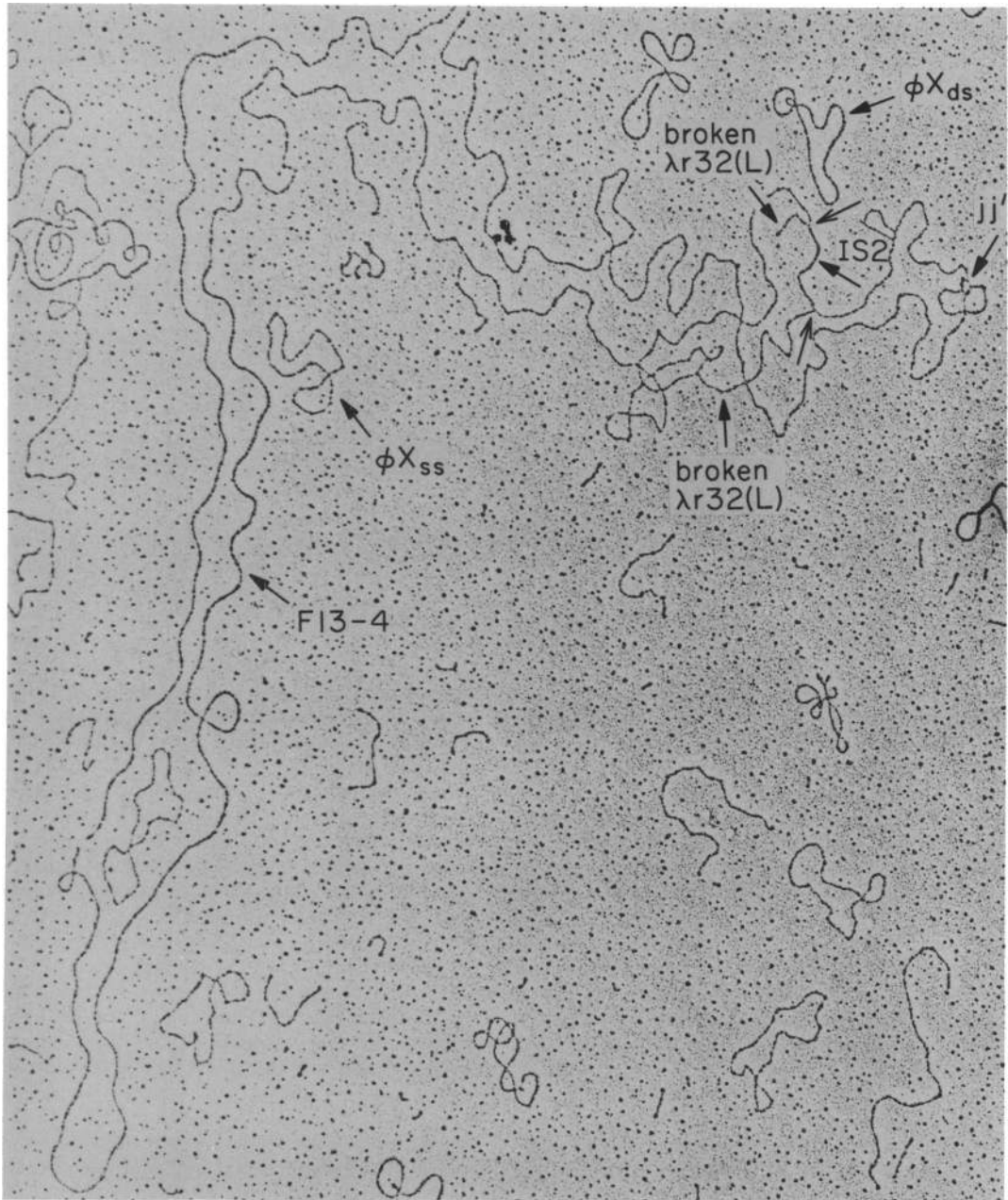FIG. 7. *Electron micrograph of a complete circular single strand of F13-4 showing the jj' feature.*

Fig. 8. *Electron micrograph of a heteroduplex of F13-4 with the separated light strand of λr32 showing the presence of a duplex region due to the homologous IS2 sequences in both DNA molecules. Note the close position of the jj' reference feature. A part of the complete F13-4 circle is double stranded due to reassociation with a broken fragment of the complementary F13-4 strand.*

were therefore prepared between the DNA of φ80d*arg*(*G*) and the DNA of the transducing phage λd*galOP*::*IS2*-308. The latter carries IS2 in a convenient place for this particular study. The heteroduplex shown in Fig. 8c includes the small duplex segment due to the IS2 sequence of λd*galOP*::*IS2*-308 bonded to the 16.3-17.6 F segment of φ80d*arg*(*G*). A micrograph is shown in Fig. 10. This result conclusively shows that the 16.3-17.6 F sequence is IS2 and that there-
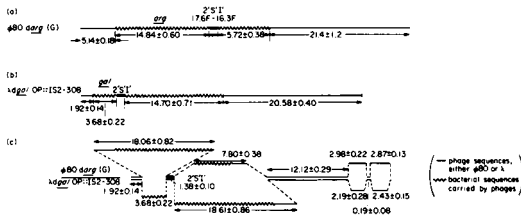
FIG. 9. *Structures of φ80darg(G) as determined by Ohtsubo et al. (22), of λdgalOP::IS2-308 as determined by Hirsch et al. (9), and of the heteroduplex between them (c). Since φ80darg(G) was known to contain the 16.3–17.6F sequence, the heteroduplex structure confirms the identification of this sequence as IS2.*

fore this is the sequence on F13–4 involved in duplex formation with λr32.

## DISCUSSION

The heteroduplex studies reported here show that IS sequences occur at several different sites on R and F plasmids. As discussed below, some of these sites play special roles in F and R plasmid-related recombination phenomena. A map summarizing the structures of the several R plasmids is displayed in Fig. 11.

One important reservation should be noted. What we have actually observed is that authentic, genetically characterized IS sequences on several phage DNAs will form duplex regions with certain segments on R and F plasmids under the conditions of hybridization used in heteroduplex experiments. The formation of hybrids does not guarantee that the sequences are identical. The conditions of hybridization in heteroduplex experiments are not very stringent ($T_m - 25C$); therefore, there may be considerable sequence mismatch between two segments which associate to form a duplex. IS sequences are genetically defined as sequences found as insertions in certain spontaneous, strongly polar mutations (4, 5, 10, 14, 16, 28, 31; see review in 36). This activity may be associated with enzymes coded for by the IS sequence and/or very specific substrate sequences at the two boundaries. It is not known whether any particular one of the homologous sequences found on R and F plasmids has the necessary functional properties of an IS sequence. Having noted this caveat, we shall continue to refer to the sequences on R and F plasmids as IS1 and IS2.

IS1 occurs at the two junctions of RTF DNA with R-determinant DNA in the plasmids R1, R100-1, R6, and R6-5. Several previous studies have shown that when R plasmids are grown in *Proteus mirabilis*, recombination or replication phenomena occur leading to the formation of

separate molecules containing only RTF and only the R-determinant segments of the entire R plasmid (6, 35) and/or of molecules containing one RTF unit and many tandem copies of the R-determinant segment (26, 27).

A plausible and natural mechanism for the segregation of an entire R plasmid into an RTF segment and an R-determinant segment is reciprocal recombination between the two IS1 sequences at the two junctions between the RTF and R-determinant components of the entire R plasmid. This mechanism is illustrated in Fig. 12.

Molecules containing one RTF segment and many R-determinant segments are formed when *P. mirabilis* strains carrying an R plasmid are grown in the presence of high concentrations of antibiotics (26, 27). We surmise that the structure of such a molecule is as shown in Fig. 12, with an IS1 at each junction of RTF with an R-determinant and at each junction of one R-determinant segment with another. It is not known whether the formation of such molecules involves unequal crossing-over in recombination or some unequal form of replication generating tandem repetitions. In either case, the IS1 sequence would play a central role in the formation of the molecule containing multiple copies of the R-determinant segment when a bacterium is grown in the presence of antibiotics. When the high concentration of antibiotics is removed from a *P. mirabilis* culture, the R plasmids revert back to their normal structure with one RTF and one R-determinant. Recombinations of some of the IS1 sequences could be the mechanism for such a segregation. There must be, however, additional specific mechanisms that control the sizes of the plasmids in the presence and absence of antibiotics.

We have observed that IS2 occurs in R6-5 at the position 67.5–68.9 kb (Fig. 4). R6 and R100-1 are homologous to R6-5 in this region; R1 and F are not (33). The R6, R100, and R1 plasmids are independent isolates from Germany, Japan, and England, respectively (2, 15, 18). If they are all derived from a common ancestor, the insertion of IS2 from some other DNA may have occurred after divergence of R1 but before the divergence of R6 and R100 from each other. The *tra* genes of F (other than *traJ*) are transcribed as a single operon under the control of *traJ* (2), which has F coordinates between 62 and 68 and therefore R6 coordinates of approximately 61 to 66 (33). It is not known whether the same situation holds for the *tra* genes of R factors, but it might. IS2 appears to be a stop signal for transcription in one orientation, thus causing strongly polar mutations, and
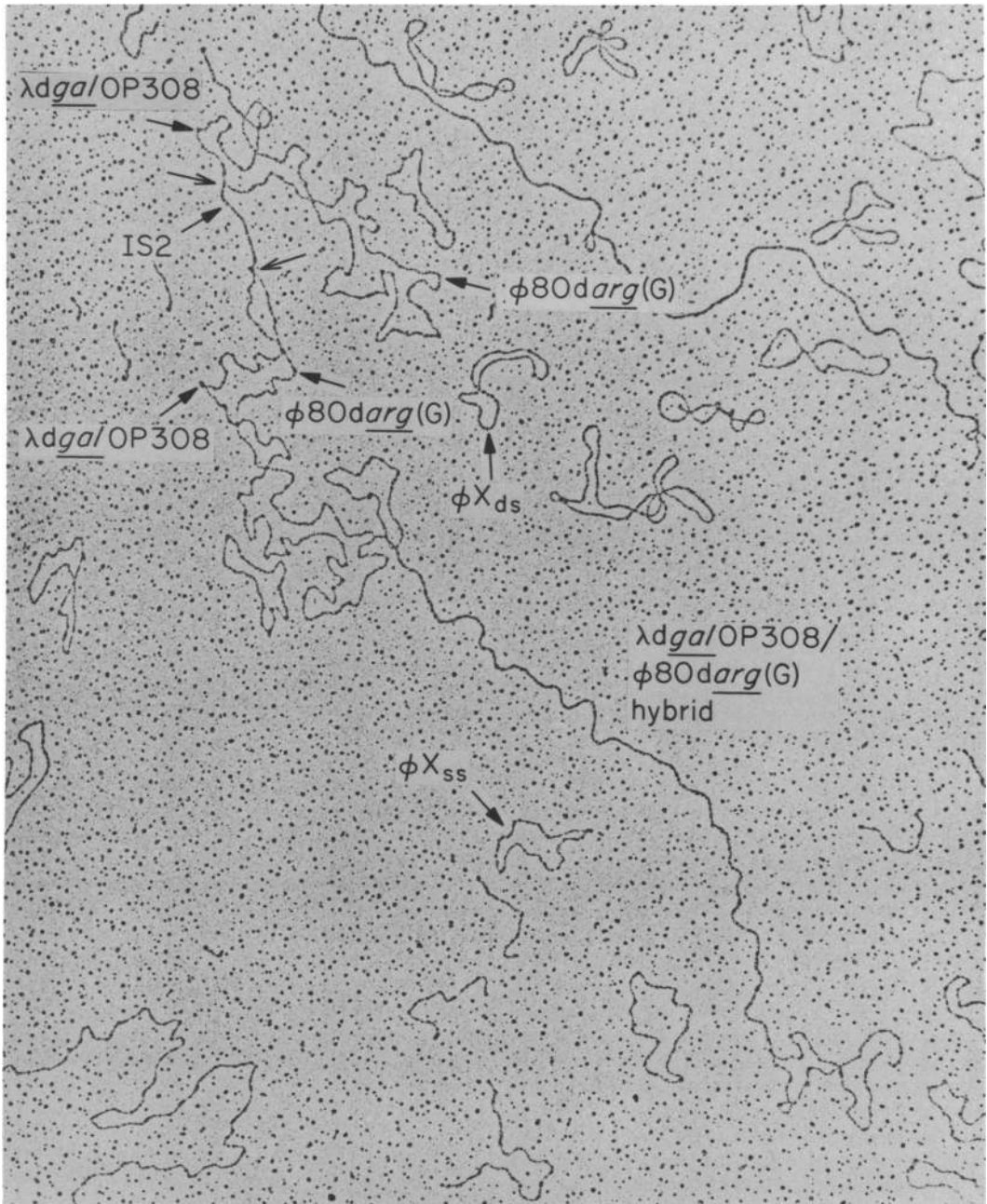
FIG. 10. *Electron micrograph of a ϕ80darg(G)/λdgalOP::IS2-308 heteroduplex. The several features of the molecule can readily be interpreted by comparison with Fig. 9c.*

a promoter that can initiate constitutive transcription in the other (30). In the orientation present in R6-5, etc., IS2 would function as a promoter for transcription of the *tra* genes (i.e., to the left in Fig. 12) and thus might contribute to the *tra*+ character; it would not cause polar mutations.

We strongly suspected that the $d_a$ sequence of the R plasmids (Fig. 1 and 4) is an IS sequence. We have shown that it is not IS2. Ptashne and Cohen have now shown that it is IS3 (24), as reported in the paper accompanying this one.
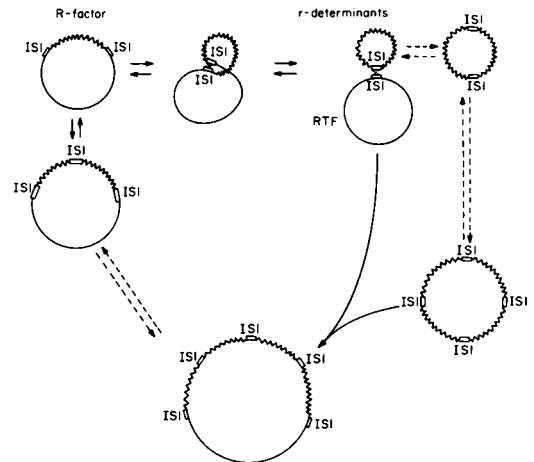
We have observed that the sequence of F with coordinates 16.3–17.6 F is IS2. Previous studies
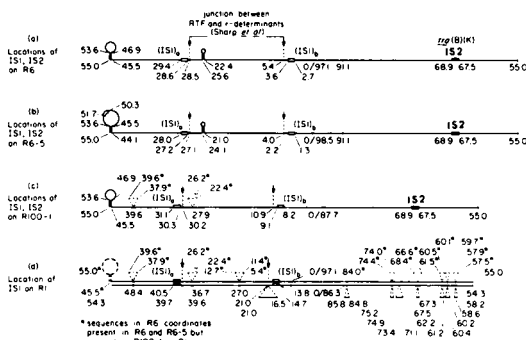
FIG. 11. *Summary of the structures of R6, R6-5, R100-1, and R1 showing the positions of the IS1 and IS2 sequences on these plasmids. Note also that, in the accompanying paper, Ptashne and Cohen (24) report that the sequences 44.1–45.5, 50.3–51.7, and 53.6–55.0 of R6-5 (and therefore 45.5–46.9 and 53.6–55.0 of R6 and R100-1) are IS3. The figure also shows the correlation of coordinates on the several plasmids from the heteroduplex studies of Sharp et al. (33), with some revisions in numerical coordinates based on the present work. In (c), dashed lines show features present on R6 and R6-5 but not on R100-1. Coordinates of these features are given an asterisk and are R6 coordinates. (d) actually shows an R6/R1 duplex. Again R6 coordinates for points not present on R1 are indicated by an asterisk. R6, R6-5, and R100-1 contain an insertion of the IS2 sequence at coordinates 67.5–68.9, which is in the general vicinity of tra (B) (K) cistrons of Achtman et al. (1, 3).*



FIG. 12. *Observed positions of the IS1 sequence at the boundaries between RTF and R-determinant components of R plasmids and proposed positions at each junction in plasmids containing multiple tandem copies of the R-determinant sequences. The top of the figure shows a specific model for segregation of an R plasmid into RTF and R-determinant components by reciprocal recombination between the two IS1 sequences. The lower part of the picture suggests that, in a plasmid containing a single copy of RTF and multiple tandem copies of the R-determinant sequence, IS1 occurs at each of the junctions. However, no specific model is proposed for the formation of the large plasmid.*

from this laboratory have indicated that the 16.3–17.6 F sequence has some of the properties of an insertion sequence. It has appeared as an insertion with the same polarity as 16.3–17.6 F at the point 22.4 F in some preparations of F 42 (Flac) (12). It has appeared as an insertion with the opposite polarity at the point 35.2 F in the plasmid F8-33 (34). The Hfr parent of F13 is Hfr 13. We have previously proposed that the structure of F13 indicates that the chromosome of Hfr 13 was formed by reciprocal recombination of the IS2 sequence of F with an IS2 sequence resident on the chromosome of the F⁺ parent of Hfr 13 (12).

It has been suggested that a plausible mechanism for the translocation of an IS sequence in a DNA molecule consists of two steps: (i) excision with or without duplication of the IS sequence from some other position on a DNA molecule by reciprocal recombination between its two ends to generate a small circular molecule (which has never been detected [11]); and (ii) reintegration of this small circular IS without circular permutation at a new point (8, 10). Thus, in this model, an insertion is due to recombination within an IS sequence. The observations re-

ported here suggest that recombination between two IS sequences plays an important role in R- and F-related integration and excision phenomena.

In addition to participating in reciprocal exchanges involved in segregation or integration events, the special sequences of F and R (including IS2) that are active in recombination events also seem to participate in half-site specific recombination exchanges (7, 25, 37). For example, IS2 at 16.3–17.6 F occurs at the junction of F DNA with chromosomal DNA in many F8 plasmids.

### LITERATURE CITED

1. Achtman, M. 1973. Genetics of the F sex factor in *Enterobacteriaceae*. Curr. Top. Microbiol. Immunol. **60:**79–123.
2. Achtman, M., and P. Helmuth. 1975. The factor carries

an operon of more than $15 \times 10^6$ daltons coding for deoxyribonucleic acid transfer and surface exclusion, p. 95-103. *In* D. Schlessinger (ed.), Microbiology—1974. American Society for Microbiology, Washington, D.C.

3. Achtman, M., N. Willetts, and A. J. Clark. 1971. Beginning a genetic analysis of conjugational transfer determined by the F factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. J. Bacteriol. 106:529-538.

4. Adhya, S. L., and J. A. Shapiro. 1969. The galactose operon of *E. coli* K12. I. Structural and pleiotropic mutations of the operon. Genetics 62:231-247.

5. Brachet, P., H. Eisen, and A. Rambach. 1970. Mutations of Coliphage λ affecting the expression of replicative functions O and P. Mol. Gen. Genet. 108:266-276.

6. 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.

7. Davidson, N., R. C. Deonier, S. Hu, and E. Ohtsubo. 1975. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. X. Deoxyribonucleic acid sequence organization of F and of F-primes, and the sequences involved in Hfr formation, p. 56-65. *In* D. Schlessinger (ed.), Microbiology—1974. American Society for Microbiology, Washington, D.C.

8. Fiandt, M., W. Szybalski, and M. H. Malamy. 1972. Polar mutations in *lac*, *gal* and phage λ consist of a few DNA sequences inserted with either orientation. Mol. Gen. Genet. 119:223-231.

9. Hirsch, H. J., H. Saedler, and P. Starlinger. 1972. Insertion mutations in the control region of the *gal* operon of *E. coli* II. Physical characterization of the mutations. Mol. Gen. Genet. 115:266-276.

10. Hirsch, H. J., P. Starlinger, and P. Brachet. 1972. Two kinds of insertions in bacterial genes. Mol. Gen. Genet. 119:191-206.

11. Hradecna, Z., and W. Szybalski. 1967. Fractionation of the complementary strands of coliphage λ DNA based on the asymmetric distribution of the poly I,G-binding sites. Virology 32:633-643.

12. Hu, S., E. Ohtsubo, and N. Davidson. 1975. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*: structure of F13 and related F-primes. J. Bacteriol. 122:749-763.

13. Jordan, S., H. Saedler, and P. Starlinger. 1967. Strong polar mutations in the transferase gene of the galactose operon in *E. coli*. Mol. Gen. Genet. 100:296-306.

14. Jordan, E., H. Saedler, and P. Starlinger. 1968. 0°—and strong polar mutations in the *gal* operon are insertions. Mol. Gen. Genet. 102:353-363.

15. Lebek, G. 1963. Über die Enstehung mehrfach resistanter Salmonellen. Ein experimenteller Beitrag. Zentralbl. Bakteriol Parasitenkd. Infektionskr. Hyg. Abt. I Orig. 188:494-505.

16. Malamy, M. H. 1970. Some properties of insertion mutations in the *lac* operon. *In* J. R. Beckwith and D. Zipser (ed.), The lactose operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

17. Malamy, M. H., M. Fiandt, and W. Szybalski. 1972. Electron microscopy of polar insertions in the *lac* operon of *Escherichia coli*. Mol. Gen. Genet. 119:207-222.

18. Meynell, E., and N. Datta. 1967. Mutant drug resistant factors of high transmissibility. Nature (London) 214:885-887.

19. 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.

20. Nakaya, R., A. Nakamura, and Y. Murata. 1960. Resistance transfer agents in *Shigella*. Biochem. Biophys. Res. Commun. 3:654-659.

21. Ohtsubo, E., R. C. Deonier, H. J. Lee, and N. Davidson. 1974. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. IV. The F sequences in F14. J. Mol. Biol. 89:565-583.

22. Ohtsubo, E., H. J. Lee, R. C. Deonier, and N. Davidson. 1974. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. VI. Mapping of F14 sequences homologous to φ80d*metBJF* and φ80d*argECBH* phages. J. Mol. Biol. 89:599-618.

23. Ohtsubo, E., Y. Nishimura, and Y. Hirota. 1970. Transfer defective mutants of sex factors in *E. coli*. I. Defective mutants and complementation analysis. Genetics 64:173-188.

24. Ptashne, K., and S. N. Cohen 1975. Occurrence of insertion sequence regions on plasmid deoxyribonucleic acid as direct and inverted nucleotide sequence duplications. J. Bacteriol. 122:776-781.

25. Reif, H. J., and H. Saedler. 1974. IS1 is involved in deletion formation in the *gal* region of *E. coli* K12. Mol. Gen. Genet. In press.

26. Rownd, R., H. Kasamatsu, and S. Mickel. 1971. The molecular nature and replication of drug resistance factors of *Enterobacteriaceac*. Ann. N.Y. Acad. Sci. 182:188-206.

27. Rownd, R., and S. Mickel. 1971. Dissociation and reassociation of RTF and R-determinants of the R factor NR1 in *Proteus mirabilis*. Nature (London) New Biol. 234:40-43.

28. Saedler, H., J. Besemer, B. Kemper, B. Rosenwirth, and P. Starlinger. 1972. Insertion mutations in the control region of the *gal* operon of *E. coli*. I. Biological characterization of the mutations. Mol. Gen. Genet. 115:258-265.

29. Saedler, H., and B. Heiss. 1973. Multiple copies of the insertion-DNA sequences IS1 and IS2 in the chromosome of *E. coli* K12. Mol. Gen. Genet. 122:267-277.

30. 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.

31. Shapiro, J. A. 1969. Mutations caused by the insertion of genetic material into the galactose operon of *E. coli*. J. Mol. Biol. 40:93-105.

32. Shapiro, J., L. MacHattie, L. Eron, G. Ihler, K. Ippen, and J. Beckwith. 1969. Isolation of pure *lac* operon DNA. Nature (London) 224:768-774.

33. 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.

34. Sharp, P. A., M. T. Hsu, E. Ohtsubo, and N. Davidson. 1972. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. I. Structure of F-prime factors. J. Mol. Biol. 71:471-497.

35. Silver, R. P., and S. Falkow. 1970. Specific labeling and physical characterization of R-factor deoxyribonucleic acid in *Escherichia coli*. J. Bacteriol. 104:331-339.

36. Starlinger, P., and H. Saedler. 1972. Insertion mutations in microorganisms. Biochimie 54:177-185.

37. Sugino, Y. 1966. Mutants of *Escherichia coli* sensitive to methylene blue and acridines. Genet. Res. 7:1-11.

38. Tye, D. K., R. K. Chan, and D. Botstein. 1974. Packaging of an oversized transducing genome by *salmonella* phage P22. J. Mol. Biol. 85:485-500.