JOURNAL OF BACTERIOLOGY, May 1975, p. 749-763

Vol. 122, No. 2
Printed in U.S.A.

Electron Microscope Heteroduplex Studies of Sequence Relations Among Plasmids of *Escherichia coli*: Structure of F13 and Related F-Primes[1]

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

Department of Chemistry, California Institute of Technology, Pasadena, California 91125

Received for publication 6 February 1975

The structure of F13, a plasmid containing *lac*, *purE*, and *proC*, has been determined by heteroduplex analysis. As expected for an F-prime formed by a type II excision event, it contains all the sequences of F plus a large segment of *Escherichia coli* chromosomal deoxyribonucleic acid. There is a sequence of F with coordinates 16.3-17.6F which has been shown in other studies to be the insertion sequence IS2. This IS2 occurs twice on F13, once at each of the two junctions of F deoxyribonucleic acid with chromosomal deoxyribonucleic acid. The sequence $\alpha\beta$ which occurs twice on F with coordinates 93.2-94.5/0F and 13.7-15.0F occurs an additional three times, twice in an inverted order relative to the $\alpha\beta$ sequences of F, on the chromosomal sequences of F13. The structures of the plasmids F13-4 and F210 have been determined. The common sequences of F13 with F152-1 (a derivative of F152, the classical F_2gal) and with F13-4 and F210 have been mapped. These results partially map *lac*, *proC*, *tsx*, and *purE* on F13. On the basis of all of these results, it is proposed that Hfr 13 (the parent of F13) was formed by reciprocal recombination between IS2 on F and an IS2 resident at a point between *lac* and *proC* on the chromosome of the F^+ parent of Hfr 13. It is proposed that this IS2 and the several $\alpha\beta$ sequences on the chromosomal part of F13 are hot spots for recombination with F, i.e., for Hfr formation. The point of origin and direction of transfer of many Hfr's can be explained by this hypothesis. In particular, the sequence relations of F42-1 (F*lac*) and of F152-1 (F_2gal) with F13 are completely consistent with this model.

In several previous investigations from this laboratory, the electron microscope heteroduplex method has been applied to study the sequence organization of bacterial F' plasmids (18, 21). Figure 1 depicts the recombination events involved in formation of an F' plasmid, starting with an F^+ cell. The sequences on the circular deoxyribonucleic acid (DNA) molecule, F, are denoted by $f_1, f_2, \ldots f_8$. The sequences on the main bacterial chromosome are denoted by $b_1, b_2, \ldots b_n$. We shall refer to the DNA of the main bacterial chromosome as chromosomal DNA and to that of F as F DNA. The point on F at which transfer of DNA into a recipient female originates and the direction of transfer are indicated by the arrow between f_3 and f_4.

An Hfr cell is formed when the circular DNA molecule, F, integrates by reciprocal recombination into the main bacterial chromosome. In Fig. 1 this recombination is depicted as occurring at a site between the sequences f_1 and f_8 of F and between b_1 and b_n of the bacterial chromosome. We assume that the structural origin for conjugal transfer is not affected by integration of F. Then the sequences of DNA transferred by the donor upon mating are in the order, $f_3, f_2, f_1, b_1, b_2, \ldots b_{n-1}, b_n, f_8, \ldots f_4$. Sequences b_1 and b_2 thus carry early or proximal chromosomal markers; b_{n-1} and b_n are late or distal markers. Since the property of fertility is itself transferred late in an Hfr mating, some of the genes of F that are essential for that function must be included within $f_4, f_5, \ldots f_8$. (For genetic evidence as to the position of the region for conjugal transfer on F, see ref. 24.)

By the reverse excision process, the Hfr can revert to the original F factor and the original bacterial chromosome. By an aberrant excision, a circular F' carrying F sequences and some bacterial chromosome sequences is formed. If, as in type II excision (Fig. 1), the recombination occurs between bacterial chromosome se-

[1] Contribution no. 4956.

[2] Present address: Department of Microbiology, State University of New York at Stony Brook, Stony Brook, N.Y. 11790.

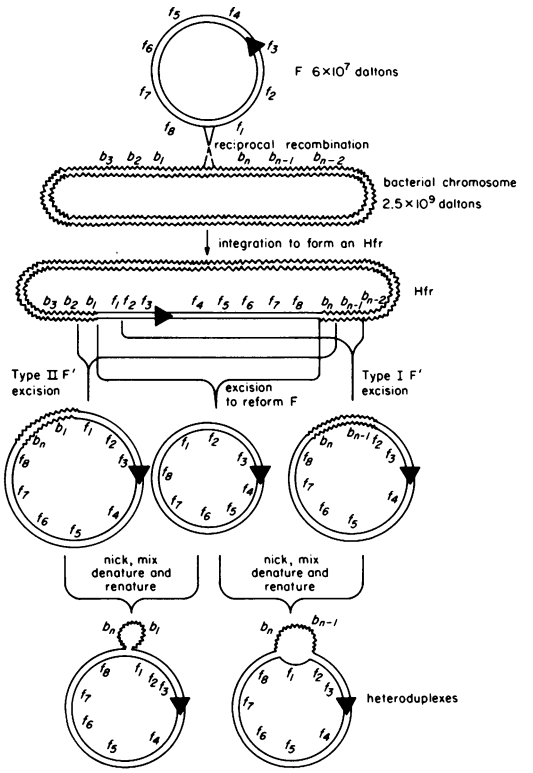

FIG. 1. *Models for: (i) the integration of F into the bacterial chromosome to give the chromosome of an Hfr cell; (ii) excision of a type II F, of one kind of type I F', and of F from the Hfr chromosome; and (iii) heteroduplex structures of a type I and a type II F' with F.*

quences on both sides of F, the resulting F' will contain all of F and both proximal and distal chromosomal sequences depicted as b_1 and b_n. If the recombination occurs between an F sequence and a chromosomal sequence, as illustrated for the type I excision in Fig. 1, the F' will be missing a part of F, shown as f_1, and carry only the distal chromosomal sequences. To be fertile, this F' must carry the genes essential for fertility, some of which lie in the interval $f_4, \ldots f_8$. The notations I and II in Fig. 1 are those introduced by Scaife (19).

Two important general points in the heteroduplex analysis of F' plasmids are as follows. (i) Heteroduplexes between F and an F' give two different general kinds of structure, depending on whether the F' was formed by a type I or a type II excision event. These two kinds of heteroduplexes are depicted in Fig. 1. In a heteroduplex of F with a type II F', which carries all of the sequences of F, there is a complete circle of duplex DNA consisting of the F sequences. The chromosomal sequences b_1 and b_n are seen as a single-strand insertion loop, which leaves and returns to the duplex at a point.

For the type I F', the heteroduplex with F consists of an incomplete circular duplex containing only those sequences of F present on the F' ($f_2, \ldots f_8$ in Fig. 1). The circle is completed by a substitution loop, one arm of which is the F sequence missing in the F' (in this case, f_1) and the other arm of which is the chromosomal sequences ($b_1 b_n$) that are present on the F' but absent from F.

(ii) The site on F and the site on the bacterial chromosome at which recombination occurred when a particular Hfr was formed can be mapped by heteroduplex analysis of the F'. In Fig. 1, the recombination sites are shown to lie between f_1 and f_8 of F and between b_1 and b_n of the bacterial chromosome. In the Hfr chromosome, the junctions lie within the recombinant sequences $b_1 f_1$ and $f_8 b_n$. Both of these junctions are present in a type II F'. The $b_1 f_1$ junction is missing, but the $f_8 b_n$ junction is present in the type I F' depicted in Fig. 1. In the heteroduplexes depicted in Fig. 1, the junctions cannot be mapped relative to other features of F because there are no reference points on the circular parts of the molecule. However, as exemplified in our several previous papers as well as in the present one, heteroduplexes of the F' of interest with a deleted F or with another suitable reference F' can be used to map the recombination junctions.

Sharp et al. (21) and Ohtsubo and Hsu (manuscript in preparation) have studied the parental F-primes F100 (F_1*gal*), F152 (F_2*gal*), and F8 (F_8*gal*) and/or some progeny F-primes derived by repeated culture in different laboratories of strains originally bearing these parental plasmids. The structures of all these molecules suggest that the parental plasmids were formed by a type I excision. This result is consistent with the genetic properties of the plasmids.

F14 has a structure suggesting that it was formed by a type II excision, but there is no positive genetic evidence that both distal and proximal markers from the Hfr are present on F14 (8, 18).

By genetic tests, F13 was formed from Hfr 13 by a type II excision; it carries both proximal markers (*proC* and *purE*) and a distal marker (*lac*) of the Hfr (11, 20). We have therefore studied the physical structure of F13 as an example of a genetically well-characterized F' plasmid formed by a type II excision. We have also studied several other F-primes that were

derived from different Hfr's, but contain some of the same chromosomal sequences, and F13-4, which was derived from F13 by P1 transduction (22).

MATERIALS AND METHODS

The bacterial strains used are listed in Table 1. A cross-streaking technique was used to test for transfer of various markers on the plasmids. Tests were carried out on ethyl methylene blue indicator plates or Davis minimal medium plates supplemented with suitable nutrients. As recipients, we used the F⁻ strains of PB314, ED1111, and χ478 in testing for transfer of *lac* and *purE* and χ478 in testing for *proC*. All other aspects of our experimental techniques are described in previous papers (18, 21) and in a forthcoming paper by Ohtsubo and Hsu (manuscript in preparation).

RESULTS

It will be conducive to clarity, we believe, to first present the final results. The evidence supporting these conclusions will be briefly explained in later sections.

The structures determined for all of the F-primes studied are shown in Fig. 2. We use a linear representation for the circular molecules. It is understood that the two ends of the linear molecule are joined to complete the circle. Sawtooth lines denote bacterial chromosomal sequences; straight lines are F sequences. Certain special sequences are indicated as thicker regions.

Our coordinate system is explained in detail in Ohtsubo et al. (18). All coordinates are given in kilobase (kb) units (1,000 nucleotide pairs or nucleotides for double- and single-stranded DNA, respectively). F itself is a circular molecule 94.5 kb in length. We have previously assigned a coordinate system on F with a convention for the clockwise direction and with

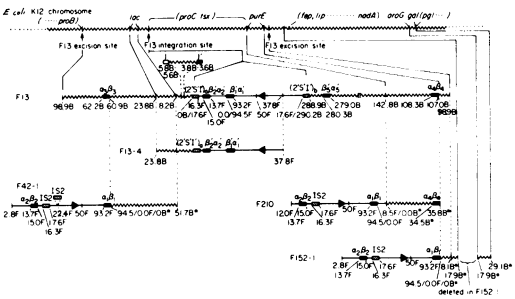

FIG. 2. *Structures of the several plasmids studied here. For an explanation of the notation and for further discussion, see text. F, F coordinate; B, F13 chromosomal DNA coordinate; B*, other F-primes chromosomal DNA coordinate. Symbols:* ∿∿ *, chromosomal DNA;* ——*, F DNA;* ▬ *,* ▨ *, special sequences;* ▶ *, transfer orientation of F-primes;* ——*,* -----*, construction lines.*

a defined terminus and origin of the circular map at 94.5/0F. Coordinates of an F sequence on an F or an F′ are given the suffix F.

Coordinates of the bacterial chromosome sequences of an F′ are denoted by the suffix B. We usually assign the value 0B to one junction of F DNA with chromosomal DNA. The B coordinates then increase along the chromosomal sequences until one reaches the other junction. According to this convention, the B coordinates on a given F′ are unambiguously defined; however, there is a disadvantage in that the same chromosomal sequence will usually have different B coordinates when it is present on two different independently derived F-primes.

With these conventions, we now describe the structure of F13 and the other F-primes studied here, all as depicted in Fig. 2. The genetic map of the relevant part of the *Escherichia coli* K-12

TABLE 1. *Bacterial strains*

Strain	Plasmid	Plasmid genotype	Chromosomal genotype	Source and reference
W1485	F		*thi*	D. Freifelder (21)
W3747	F13	*lac*⁺ *proC*⁺ *tsx*ʳ *purE*⁺	*met strʳ tsx*ʳ	Y. Hirota (11)
JE513	F13-4	*lac*⁺	*lac purB strʳ*	(22)
AT3155	F42-1	*lac I78(i*ˢ*)*	*lac Δ175(i*⁻*) strʳ*	A. L. Taylor (17)
χ790	F210	*purE*⁺	*his strʳ*	R. Curtiss III
ND6	F152-1	*gal*⁺ *aroG*⁺	*gal recA pyrD try his tyr thi mal xyl mtl*	B. Low; Ohtsubo and Hsu (personal communication)
W1655	FΔ(33-43)	*T3*ˢ *φII*ˢ *τ*ˢ	*metB strʳ tsx*ˢ (λ)-λʳ	P. Broda (1)
JE3513	F8-33	*gal*⁺ *aroG*⁺ *nadA*⁺	*thr his lac galK pil met strʳ fla*	(21)
PB314	F⁻		*lac purE thi tsx*ʳ *strʳ*	P. Broda (4)
ED1111	F⁻		*lac purE thi tsx*ʳ *strʳ recA*	P. Broda (5)
χ478	F⁻		*ara leu T1*ʳ *lacZ proC tsx*ʳ *purE* λ-*gal trp lys strʳ mtl xyl metE thi*	R. Curtiss III

chromosome is shown at the top; a left-to-right direction corresponds to clockwise on the circular map (23). The chromosomal sequences of the several F-primes are similarly ordered. In the present paper, we use suffix B to denote F13 chromosomal coordinates and B* for chromosomal coordinates on any other plasmid. Identical chromosomal sequences on different F-primes (but not on the *E. coli* map at the top) have the same vertical position in Fig. 2. Because of the common order for B sequences, the order of the F sequences is different for those F-primes derived from Hfr's which transfer DNA in a counterclockwise order (F42-1, F152-1, F210) and those F-primes derived from an Hfr which transfers in a clockwise direction (F13, F13-4). It should be noted, however, that in later figures in which the heteroduplex structures of individual F-primes are presented F sequences are always oriented so that their clockwise order is left to right. In such figures, the chromosomal sequences of F13 and F13-4 have an opposite order to that shown in Fig. 2.

F13 has a molecular length of 384.6 ± 6.3 kb, corresponding to a molecular weight of 2.54×10^8. It contains 290 kb of chromosomal DNA. As expected from the model of Fig. 1 for a type II F-prime, it contains the entire F genome. The points of the type II joining event of the distal (to the left of *lac*) sequence of Hfr 13 with the proximal (to the right of *purE*) sequences are indicated in Fig. 2.

Several special sequences of F, sequences that previous studies have shown to be especially active in F-related recombination phenomena (7), occur in interesting ways in F13.

(i) There is a sequence of F with coordinates 16.3–17.6F which we have previously denoted as the $\epsilon\zeta$ sequence (7). This sequence has been shown (15) to be identical (at the criterion of base pairing used in heteroduplex analysis) to the insertion sequence IS2 that causes strongly polar mutations in *E. coli* and in coliphages (12, 13). IS2 occurs twice on F13, once at each junction of F DNA with chromosome DNA. (In Fig. 2 the sequences are denoted as 2'S'I' because of our convention about the clockwise order on F; the two 2'S'I' sequences are identified by subscripts a and b. In general, special sequences that have a clockwise order on F according to our convention and that are depicted in a particular linear representation in a left-to-right order are unprimed, and inversely.)

It should be noted that in the present paper we consistently refer to the 16.3–17.6F sequence as IS2. The two sequences can hybridize together to form a heteroduplex; this result, however, does not guarantee that they are identical. IS2 is genetically defined as a sequence found as an insertion in some spontaneous strongly polar mutations. This activity may be associated with enzymes coded for by IS2 and with the very specific sequences at the two boundaries. It is not known whether or not 16.3–17.6F does in fact have the necessary functional properties of IS2.

(ii) There is another special sequence, $\alpha\beta$, which previous work had shown occurs twice on F with coordinates 93.2–94.5/0F($\alpha_1\beta_1$) and 13.7–15.0F($\alpha_2\beta_2$) (Ohtsubo et al. [18]). We now observe that $\alpha\beta$ occurs an additional three times on the chromosomal sequences of F13. These sequences are designated $\alpha_3\beta_3$, $\alpha_4\beta_4$, and $\beta'_5\alpha'_5$ in Fig. 2; in F13 the first two have an inverted order and the last the same order relative to the $\alpha\beta$ sequences on F.

(iii) There is a small inverted repeat structure with a loop in the middle on the chromosomal part of F13, designated jj' in Fig. 2. This small feature, which is seen about 80% of the time when the appropriate region of F13 is otherwise single stranded, is useful as a reference identification feature in the study of various heteroduplexes. We do not know what, if any, biological significance it has.

The F-primes F42-1 (F*lac*), F210, F152-1, and F13-4 all have chromosomal sequences in common with F13. It was useful, in determining the structure of F13, to determine their structures and to study their heteroduplexes with F13. These results may be summarized as follows. The structure of F42-1 is shown below that of F13 in Fig. 2. All of its chromosomal sequences are included in the chromosomal sequences of F13. The polarity of the F sequence relative to the chromosomal sequences is opposite in the two F-primes. One junction of F DNA with chromosomal DNA in F42-1 occurs at the $\alpha_1\beta_1$ sequence of F. The $\alpha_1\beta_1$ sequence and the adjoining chromosomal sequences on F42-1 are congruent with $\alpha_3\beta_3$ and the adjoining chromosomal sequences on F13 (Fig. 2). This identity is a very important result for our later discussion.

An additional point is that F42-1 DNA extracted from AT3155 carries an insertion of IS2 at the point 22.4F (Fig. 2) (14). This insertion has the same polarity as the 16.3–17.6F IS2 sequence. It is not present in most of the F42 molecules extracted from JC1553 and studied previously by Sharp et al. (21). We refer to the plasmids without and with the additional IS2 at 22.4F as F42 and F42-1, respectively; the distinction is not important in considering the structure of F13.

The structure of F210 is also shown in Fig. 2. All of its chromosomal sequences are present on

F13. The two junctions of F DNA with chromosomal DNA in F210 occur at 8.5F and 12.0F. However, the $\alpha_4\beta_4$ chromosomal sequence of F13 (F13 coordinates 108.3–107.0B) is present in F210 with coordinates 34.5*–35.8B*/12.0F, i.e., at one junction of chromosomal DNA with F DNA.

The complete structure of F152-1 will be reported by Ohtsubo and Hsu (personal communication). The chromosomal sequences start within the region spanned by F13 but extend further to the right. These additional sequences are not shown in detail in Fig. 2. The right junction of F DNA with chromosomal DNA in F152-1 is the 93.2–94.5F/0F $\alpha_1\beta_1$ sequence of F. The $\alpha_1\beta_1$ sequence and the adjacent chromosomal sequences of F152-1 are congruent with the $\alpha_4\beta_4$ sequence and rightward chromosomal sequences of F13, which is important in our later discussion.

F13-4 is a deletion variant of F13 and includes the sequences depicted in Fig. 2. The jj′ inverted repeat is retained in F13-4.

Molecular lengths. Plasmids are isolated as closed circular molecules, essentially free of linear or open circular contaminants. They are then nicked to an open circular form by X rays. The homogeneity of each F-prime DNA preparation was tested and the molecular length was measured by spreading the open circular duplex DNA from aqueous ammonium acetate in the presence of added F as a length standard. These results are presented in Table 2.

Three classes of molecules were present in the preparation of F13 DNA: (i) F13, with a molecular length of 4.07 ± 0.07F or 384.6 ± 6.3 kb or 2.54×10^8 daltons, 95% by number; (ii) middle-size DNA, 46.1 ± 1.8 kb or 3.04×10^7 daltons, 5% by number; (iii) small DNA, 21.95 ± 0.86 kb or 1.45×10^7 daltons, less than 1%. Thus, the F13 molecules are homogeneous in length and are about four times the size of F. The nature and significance of the middle size and small DNAs, which occur at a low frequency, are not known. No circular DNA with the size of F was found in the F13 DNA preparation, indicating that no significant amount of segregation to form F molecules occurs. The DNA preparations of the other plasmids isolated were homogeneous in size. These data are also shown in Table 2.

Interpretation of heteroduplex and self-renatured structures of F13. In our several previous papers we have presented in explicit detail the arguments by which the structure of an F′ or other plasmid was deduced from the experimental evidence, i.e., from the several observed heteroduplex and self-renatured structures. In the present case we shall present the evidence and arguments in a very abbreviated form. In most cases, the observed heteroduplex structures are shown in the figures, but the reasoning that leads to the interpretations and coordinates is not given in full detail.

The molecules F, FΔ(33–43), F8-33, and F42-1 (1, 21) were used as reference molecules and heteroduplexed with F13. The structures of these reference F-primes are shown at appropriate places in the figures. Each of them has a different sequence that makes it useful for a study of the structure of F13 and F13-4.

Figures 3 and 4 show observed heteroduplex structures of F13 with FΔ(33–43) and F8-33. It should be noted that several of the figures depict incomplete structures, i.e., heteroduplexes in which one or both of the two strands is fragmented. Incomplete heteroduplexes often provide evidence for repeated sequences which is not obtainable from complete structures.

The structure in Fig. 3a(i) shows that all of the sequences of F occur in F13. Figure 5 is an electron micrograph showing such a heteroduplex molecule. Figure 3a(i) plus Fig. 4c map the junction of chromosomal DNA with F DNA as within the IS2 at 16.3–17.6F. The distance from the 33–43 deletion loop to the junction is measured as 15.0 ± 0.4 kb in Fig. 3a(ii); it is measured as 16.3 ± 0.4 kb in Fig. 3a(iii) and 3b

TABLE 2. *Molecular weights of F-prime DNA molecules*

DNA	Frequency of occurrence (%)	Mol wt	
		kb	$\times 10^{-6}$
F13	95	384.6 ± 6.3	253.8 ± 4.1
	~5	46.1 ± 1.8	30.4 ± 1.2
	~1	21.95 ± 0.86	14.5 ± 0.6
F13-4	100	98.4 ± 2.6	64.9 ± 1.7
F42-1	100	146.7 ± 3.0	96.8 ± 1.9
F210	100	124.3 ± 2.1	82.0 ± 1.3
F152-1	100	120.2 ± 2.0	79.6 ± 1.3
F	100	94.5 ± 2.4	62.6 ± 1.6

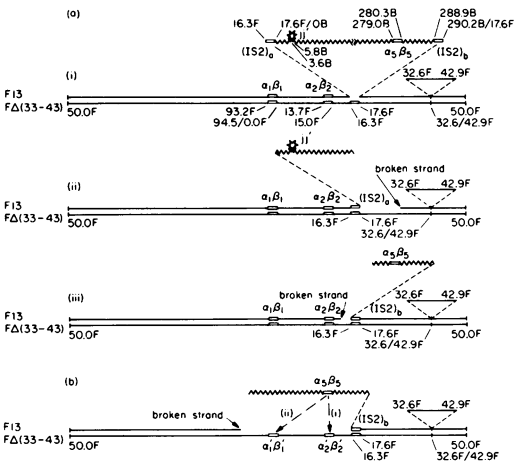

FIG. 3. *Observed heteroduplex structures of F13 with FΔ (33-43). In (a) (i), the single-strand loop due to the chromosomal sequences of F13 leaves and joins the duplex region at a point which varies between 15.9 and 16.3 kb in its distance from the Δ(33-43) reference loop, depending on what share of the duplex is contributed by $(IS2)_a$ and $(IS2)_b$, respectively. In (a) (ii) this distance is measured as 15.0 ± 0.4 kb; in (a) (iii) it is measured as 16.3 ± 0.4 kb. The several structures with a broken strand of F13, depicted in (b), show that $\alpha_5\beta_5$ occurs on the chromosomal DNA of F13 and confirm the presence of two copies of αβ ($\alpha_1\beta_1$ and $\alpha_2\beta_2$) on FΔ(33-43) and thus on F.*

and varies between these two extremes for the structures of Fig. 3a(i). These data suggest that IS2 occurs at each junction of F DNA with chromosomal DNA in F13. Several structures in Fig. 3a(iii) and b provide convincing evidence for the presence of $(IS2)_b$ on F13; the structure in Fig. 4b shows that $(IS2)_a$ is present.

Several of the structures in Fig. 3 and 4 map the αβ sequences on the chromosomal part of F13.

Inverted repeat structures in single strands of F13 were seen on the grids and on samples of F13 which were denatured and renatured in the absence of other DNA molecules. Many structures are possible, involving duplex formation between one or two of the three αβ sequences with one or two of the two α′β′ sequences, respectively. These are all shown in Fig. 6. An example is shown in Fig. 7. As indicated in Fig. 6, most of these possible structures were actually seen, although they are frequently broken. The ones that were not are topologically complex and therefore likely to be tangled and difficult to recognize. The jj′ inverted repeat feature is very useful for identifying the different αβ/α′β′ pairs.

Structure of F13-4. This F′ which was isolated by P1 transduction of F13 is fertile; it transfers *lac* but not *proC* or *purE* (Nishimura, cited by Sugino [22]). The molecular length of F13-4 was measured as 98.4 ± 2.6 kb (Table 2).

Figure 8 shows several heteroduplexes that were used to determine the structure of F13-4. The structure of Fig. 8a shows that all of the sequences present in F13-4 are present in F13, and that F13-4 is a simple deletion mutant of F13. F13-4 contains the jj′ inverted repeat which is useful for many identifications. The F segment 17.6–37.8F and the chromosomal segment 23.8–290.0B of F13 are deleted in F13-4 (Fig. 2). The incomplete structure depicted in Fig. 6c in which the $IS2_b$ of F13 is duplexed to $IS2_a$ of F13-4 is a very convincing confirmation of the evidence previously cited that F13 does indeed have two IS2 sequences at the two junctions of F and chromosomal sequences. The presence of $\alpha_5\beta_5$ in F13 is again confirmed by its being hybridized with $\alpha_2\beta_2$ in F13-4 (Fig. 8c).

Relations of F42-1 to F13 and F13-4. F42-1, F13-4, and F13 carry *lac* genes. The chromosomal sequences of F42-1 and F13-4 are much shorter than those of F13. We can place limits on the positions of the *lac* genes in the three plasmids by identifying where their chromosomal sequences overlap. The relevant heteroduplex structures are shown in Fig. 9 and 10. Because of the opposite polarity of the F sequences relative to the chromosomal sequences in F13 and F42-1, any one heteroduplex is base

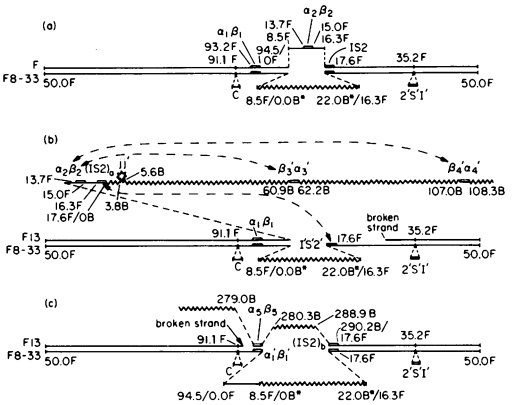

FIG. 4. *F13/F8-33 heteroduplexes. (a) Note that F8-33 has two insertions of length 1.3 kb, C at 91.1F and an inverted 16.3-17.6F (therefore 2′S′1′) at 35.2F (Sharp et al. [21]). In (b), $(IS2)_a$ of F13 pairs with IS2 (16.3-17.6F) of F8-33, confirming the presence of $(IS2)_a$ on F13. Since $\alpha_2\beta_2$ is not present on F8-33, $\alpha_2\beta_2$ of F13 is able to pair with $\beta'_3\alpha'_3$ or $\beta'_4\alpha'_4$ of F13. (c) is an out-of-register structure analogous to that of Fig. 3b showing that $\alpha_5\beta_5$ is present on F13.*

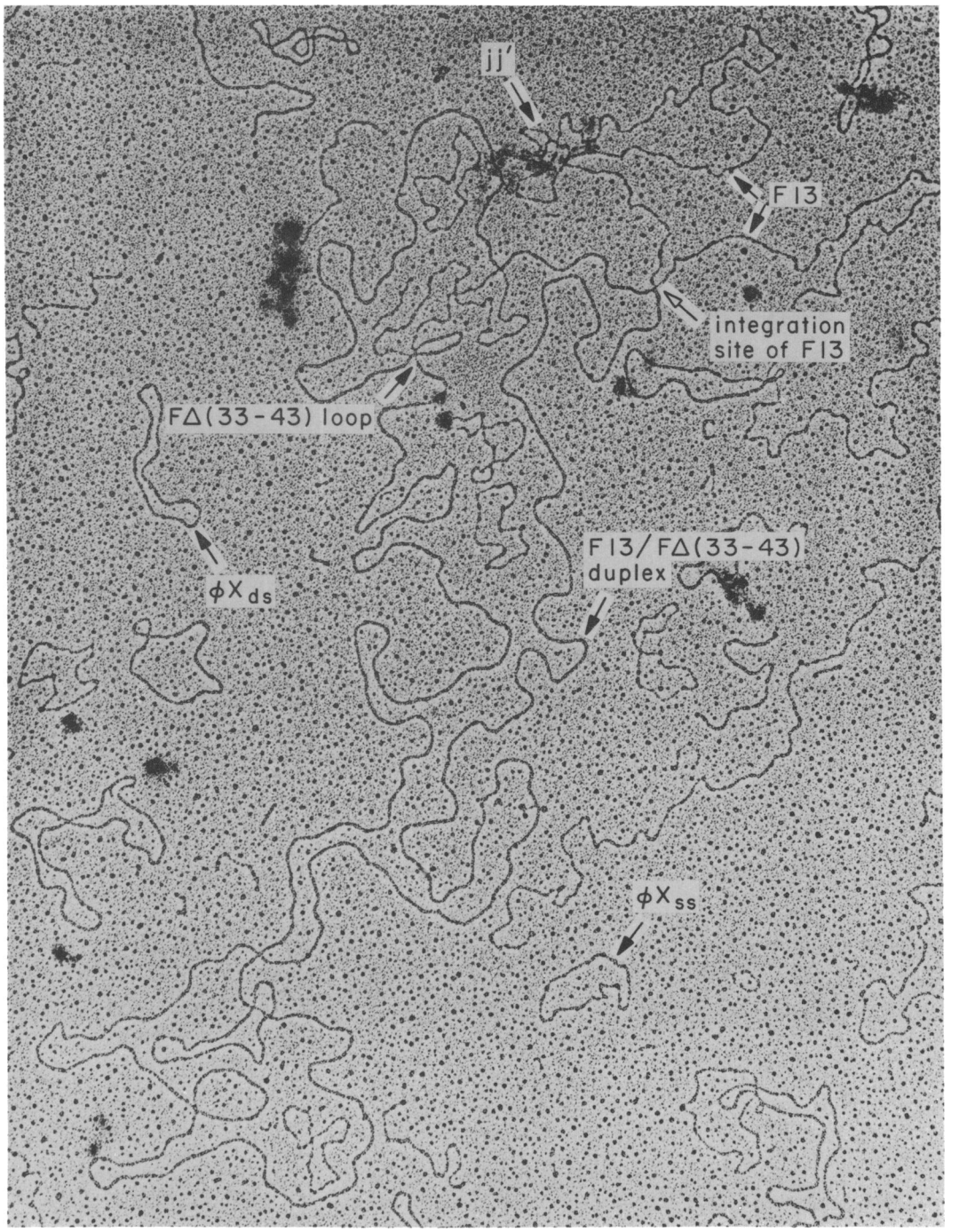

FIG. 5. *A heteroduplex molecule of F13 and FΔ(33-43). The structure includes a complete duplex circle of F sequences. At one point the Δ(33-43) single-strand loop leaves and returns to the duplex; at another point, marked in the figure as the integration site of F13, there are two single strands leaving the circle. These two single strands are part of a broken circle; they contain the chromosomal sequences of F13. One of the single strands carries the jj' secondary structure feature of F13.*

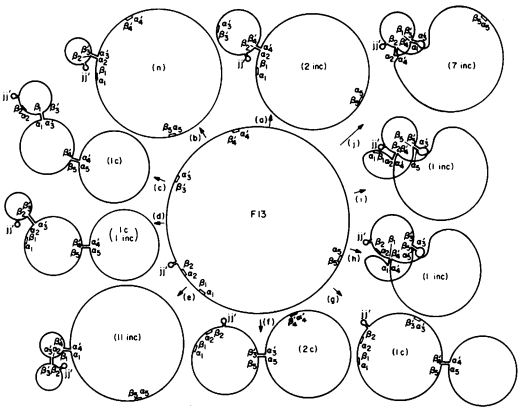

FIG. 6. *Possible structures due to reassociation of the αβ sequences on a single strand of F13. The number of times each structure was observed is indicated. It is probable that (b) was not observed because it readily forms (d) or (e). c, Complete structure; inc, incomplete structure; n, not observed.*

paired either in the F region or the chromosomal region, but not in both. Figure 10c is a diheteroduplex of one F42-1 with two F13-4 molecules, base paired in the F and chromosomal sequences, respectively.

A point of importance for the later discussion is the fact that the region of chromosomal DNA of F13, which is homologous with F42-1 DNA, has its left terminus at the $\alpha_3\beta_3$ sequence of F13 (Fig. 2). This is shown as follows. The structures in Fig. 9b are interpretable by assuming that the duplex regions are the chromosomal sequences of F13 and F42-1. It may be noted in Fig. 9b that the region of sequence complementarity with F42-1 of F13 begins at 8.2B, as determined by measuring the distance from the junction to the jj' feature. The average length of the common duplex sequences was measured as 54.8 ± 1.6 kb. This places the end of the homology region at 63.0B ± 1.6 on the F13, which is, within experimental error, the same as the coordinate of the left border of the segment $\alpha_3\beta_3$, namely, 62.2B.

The observed length of the chromosomal DNA of F42-1, as determined by heteroduplexes with F, is 52.9 ± 1.9 kb, which is a little shorter than the length of 54.8 ± 1.6 kb. Since the $\alpha_3\beta_3$ sequence occurs in F13 at 62.2–60.9B and the $\alpha_1\beta_1$ sequence occurs in F42-1 at 93.2F–94.5F/0B, that is, at the junction of F DNA with chromosome DNA, and since the sequences just beyond 62.2B on F13 are not homologous to F DNA, the observed structures can only be interpreted by the hypothesis that the F42-1 sequence in F13 ends precisely within the 60.9B to 62.2B $\beta_3\alpha_3$ region.

The $\alpha_4\beta_4$ (108.3–107.0B) in F13 frequently hybridizes with an $\alpha_2\beta_2$ (13.7–15.0F) of F42-1 (Fig. 9b). This again confirms the presence of $\alpha_4\beta_4$ in F13.

The several heteroduplex structures that bear on the relation of F13-4 to F42-1 are shown in Fig. 10a, b, and c. They are consistent with the structure assignments already given and require little comment. Figure 11 shows a complete molecule with the structure depicted in Fig. 10b. The diheteroduplex structure of Fig. 10c of an intact F42-1, base paired in the F region with an intact F13-4 and base paired in the chromosomal region with a broken strand of F13-4, shows that the F13-4 chromosomal sequences are complementary to F42-1 chromosomal sequences precisely up to the junction with F DNA of the latter. Thus, there are no translocated sequences from other parts of the *E. coli* genome or from elsewhere at this junction in F42-1.

F42-1 in AT3155 carries an additional insertion of the IS2 (εζ) sequence. The F42 derived from strain AT3155 (i.e., F42-1) carries a small insertion with the length of 1.3 kb at 22.4F, as shown in the summary graph in Fig. 2 and in the F13/F42-1 heteroduplex as a loop at 22.4F (see Fig. 9). This loop was not present in most of the F42 from JC1553, as described in our previous study (21). F13-4 is missing the F sequences from 17.6F to 37.8F. Therefore, the region around 22.4F of F42-1 is single stranded in the complete F13-4/F42-1 heteroduplex of Fig. 10a. However, in the incomplete structure of Fig. 10d, the insertion at 22.4F is hybridized to the one and only IS2 sequence of F13-4. Thus, the F42-1 insertion is an IS2. The position of the jj' loop contributes to the quantitative certainty of this identification.

F210 and F13. The heteroduplexes used to determine the structure of F210 are depicted in Fig. 12. The heteroduplex of Fig. 12c is noteworthy because it shows that the $\alpha\beta$ sequence occurs three times on F210, twice in its normal F positions and once corresponding to the $\alpha_4\beta_4$ sequence of F13. In Fig. 12b, F210 is base paired with F13 in the F sequences. In Fig. 12c, these two plasmids are base paired in their chromosomal sequences. Since *purE* has been identified to be carried in both F13 and F210, the location of *purE* must be located inside the overlapping chromosomal region, i.e., from 107.0B to 142.8B in F13 coordinates. From the final structure shown in Fig. 2, it may be inferred that F210 was formed from its parental Hfr by a type I excision between the points 12.0F and the right boundary of $\alpha_4\beta_4$.

F152-1 and the excision site of F13. In the

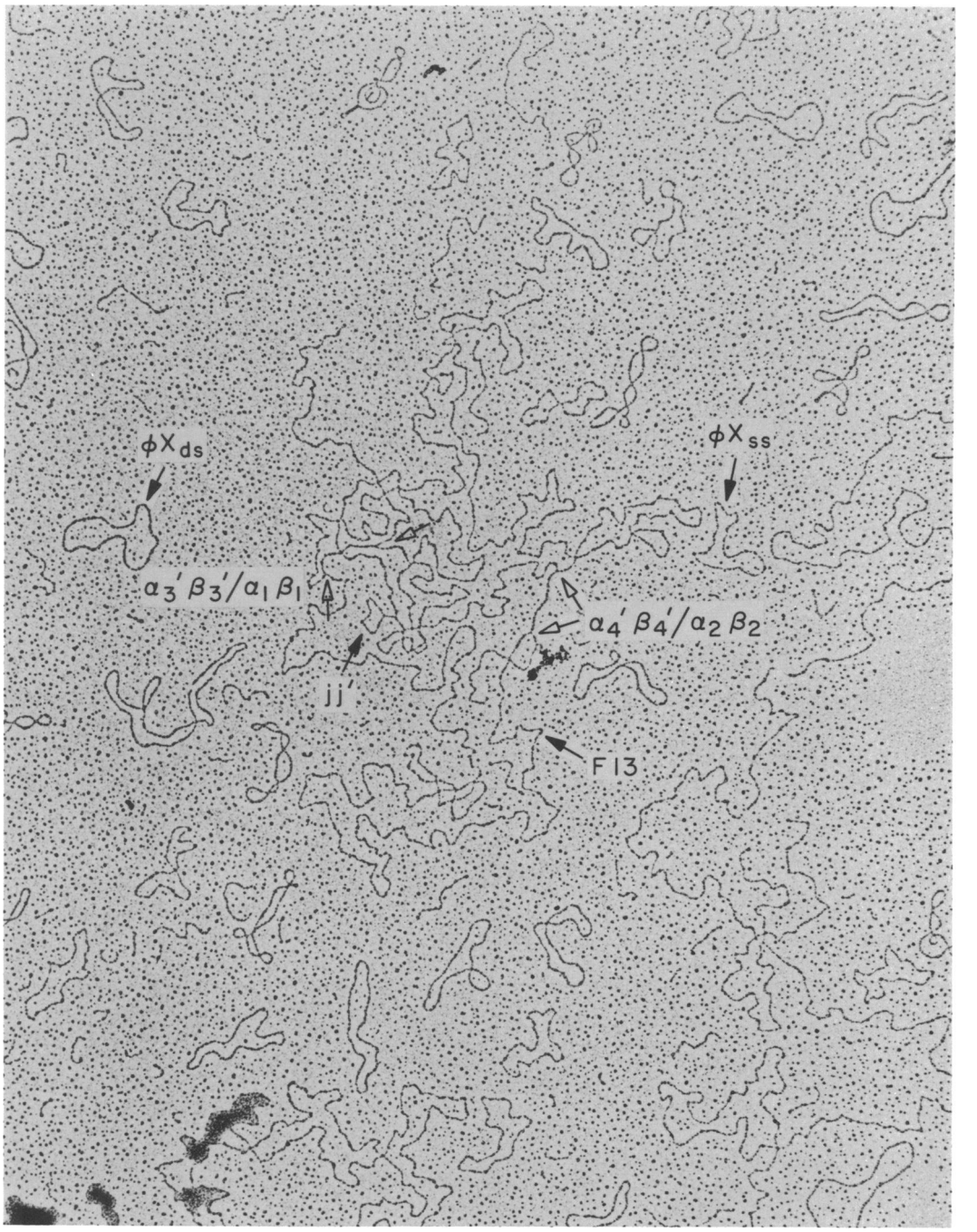

FIG. 7. *An incomplete single strand of F13 showing two short duplex regions, one formed by* $\alpha_1\beta_1/\alpha'_3\beta'_3$ *and the other by* $\alpha_2\beta_2/\alpha'_4\beta'_4$. *The jj′ feature is clearly observed in this molecule. This molecule has the structure corresponding to Fig. 6 (j).*

previous sections, the regions containing the chromosomal sequences of F42-1 (including *lac*) were mapped in the chromosomal DNA of F13 at 60.9B to 8.2B, whereas the sequences containing *purE* are in the F13 interval 142.8B to 107.0B. But *lac* and the other F42-1 sequences

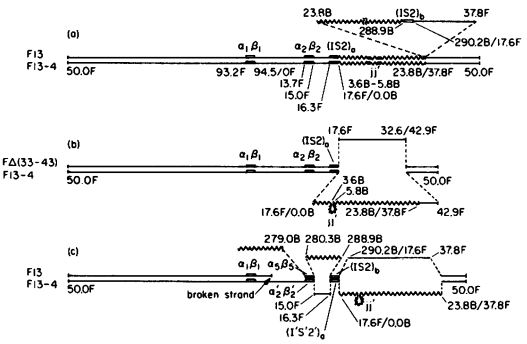

FIG. 8. *Heteroduplexes used to determine the structure of F13-4.*

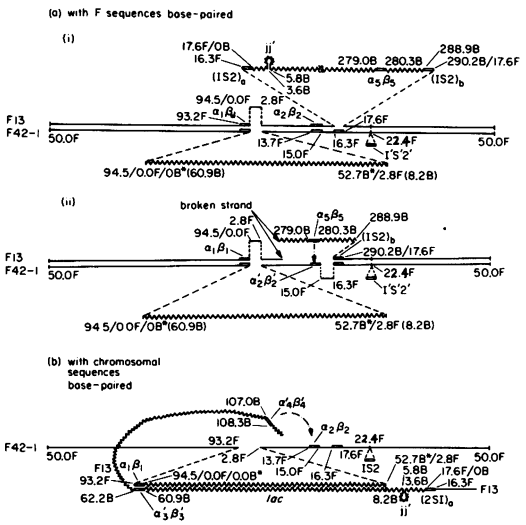

FIG. 9. *F13/F42-1 heteroduplexes. (a) The structure formed when the F sequences of these two F-primes are base paired. (b) The structure formed when their common chromosomal DNA sequences are base paired. The homologous segment on the chromosomal DNA contains the lac gene and ends at the $\alpha'_3\beta'_3$ sequence. In (b), the formation of a base-paired region between $\alpha'_4\beta'_4$ of F13 and $\alpha_2\beta_2$ of F42-1 indicates the presence of $\alpha'_4\beta'_4$ on F13. The uninterrupted homology of the duplex structure $\alpha_1\beta_1$ plus the chromosomal sequences of F42-1 with $\alpha_3\beta_3$ plus lac sequences of F13 is shown in structure (b).*

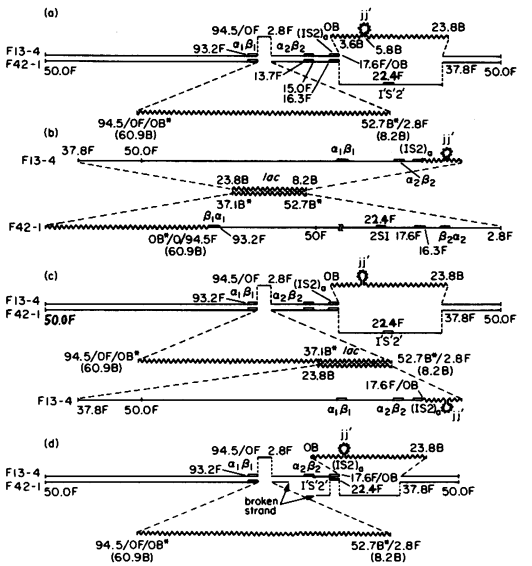

FIG. 10. *F13-4/F42-1 heteroduplexes. (a) and (b) depict structures with F sequences paired and B sequences paired, respectively. From (b) the position of the lac genes on F13 is bounded within the region 8.2–23.8B. The out-of-register structure of (d) shows that the insertion at 22.4F of F42-1 is IS2.*

are transferred as distal genes, whereas *purE* is a proximal gene on Hfr 13. Accordingly, the excision site for the formation of F13 from Hfr 13 must have F13 coordinates in the interval 60.9B to 107.0B.

The complete structure of F152-1 has been determined by Ohtsubo and Hsu (manuscript in preparation). Relevant features are shown in Fig. 2. These sequences start at a point to the right of *purE* (12.1 min). Proceeding clockwise, there is a segment of *E. coli* K-12 DNA of length 17.9 kb with no known markers, then a deletion spanning the genes *fep*, *lip* (14.6 min) through *nadA* (16.5 min), and then a segment of length 11.2 kb including *aroG* and *gal* (16.7 min). The junction of the counterclockwise end of the chromosomal sequences of F152-1 with F is at the $\alpha_1\beta_1$ sequence of F.

The heteroduplex structures illustrated in Fig. 13 show that the region of chromosomal homology of F13 with F152-1 has a length of 9.4 kb and extends from the $\alpha_4\beta_4$ sequence of F13 (108.3–107.0B on F13) to the right to the point 98.9B. Thus, 98.9B is the type II excision point of F13. On F152-1, this region of homology includes the $\alpha_1\beta_1$ sequence of F and 8.1 kb of chromosomal DNA. The $\alpha'_3\beta'_3$ sequence in F13 is frequently hybridized with the $\alpha_2\beta_2$ sequence in F152-1 forming the structures shown in Fig. 13 and 14. Such a feature was used to map the hybridized 9.4-kb chromosomal region in this heteroduplex.

As already discussed, the polarity of the F sequences relative to the chromosomal sequences on F152-1 is opposite to that of F13.

DISCUSSION

The structures of F13 and of the several F-primes studied here are shown in Fig. 2. The

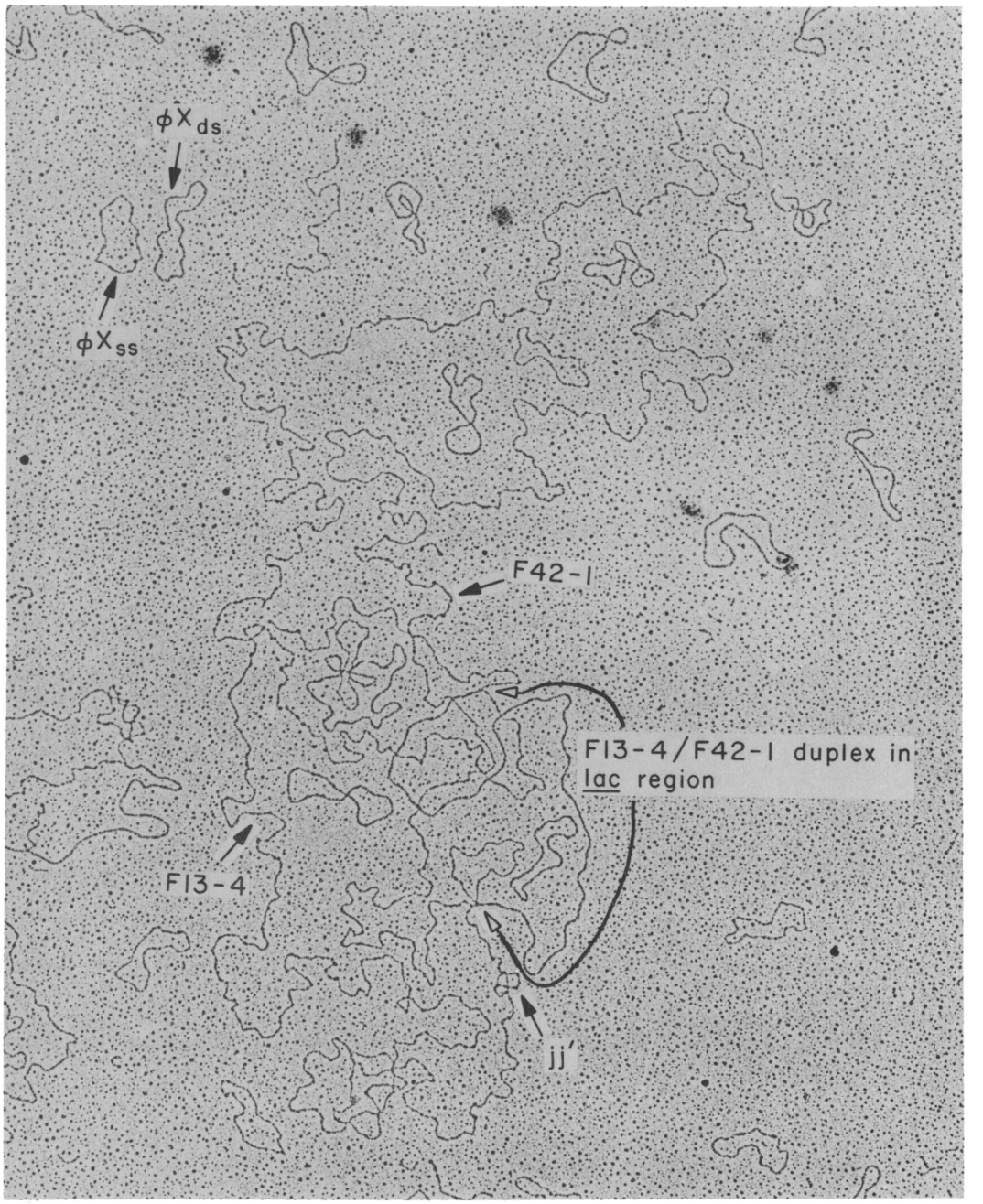

FIG. 11. *A heteroduplex molecule formed between F13-4 and F42-1. The duplex region in this molecule includes the lac genes present in both plasmids. The jj' feature in F13-4 is indicated. The F sequences in these two plasmids have the opposite polarity; therefore, in a heteroduplex in which the respective chromosomal sequences are paired the F-sequences are not complementary.*

positions of the *purE* and *lac* genes are assigned because the former is carried on F210 and the latter is on F42-1 and F13-4, as well as on F13. *proC* and *tsx* are known to lie between *lac* and *purE; proB* and *fep* are not carried on F13.

A detailed model for the integration of F to

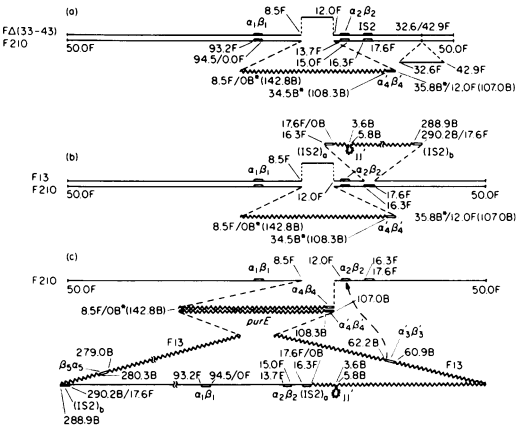

FIG. 12. *Heteroduplexes used in determining the structure of F210. As shown in Fig. 9c, we observe $\alpha'_3\beta'_3$ of F13 hybridized with $\alpha_2\beta_2$ of F210, but we do not observe $\alpha'_4\beta'_4$ of F13 so hybridized. This observation confirms the quantitative measurements in showing that $\alpha_4\beta_4$ occurs at the junction of chromosomal DNA and F DNA in F210.*

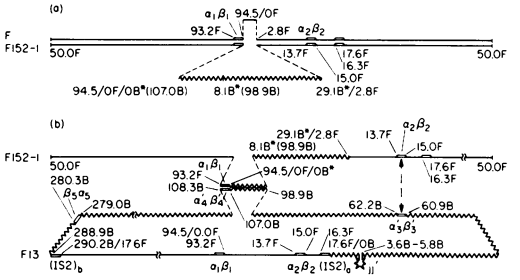

FIG. 13. *Heteroduplexes pertinent to the structure of F152-1; the relation of the chromosomal sequences of F13 and F152-1 is shown by the structure of (b).*

form an Hfr and the subsequent type II excision to give an F′ is shown in Fig. 15. This model is an amplification of that depicted in Fig. 1. We suppose there is a sequence g_1g_2 which occurs on both the chromosome of the F^+ cell and on F, and that integration occurs by reciprocal recombination between the two g_1g_2 sequences. This region then occurs twice on the Hfr at the two junctions between F DNA and chromosomal DNA. The g_1g_2 duplication is preserved in the F′ resulting from a type II excision.

Our results show that a particular sequence of F, the IS2 sequence (with coordinates 16.3–17.6F), occurs at the two junctions of chromosomal DNA with F DNA in F13. The model of Fig. 15 then suggests that Hfr 13 was formed by a reciprocal recombination between the IS2 sequence of F and a presumed IS2 sequence resident on the chromosome of the F^+ progenitor of Hfr 13.

Figure 16 shows a map of the relevant parts of the chromosome of *E. coli* K-12 inferred from the observed structure of F13. In accordance with the model of Fig. 15, it shows an IS2 at the position proposed for recombination to give Hfr 13. The map also shows the three $\alpha\beta$ sequences that occur on the chromosomal part of F13. We wish to suggest that these $\alpha\beta$ sequences are resident on the chromosome of many *E. coli* K-12 strains at the positions shown and that they are hot spots for recombination with F to give Hfr's. It is well known that there are hot spots for Hfr formation in *E. coli*, i.e., that several independently derived Hfr's have the same point of origin and direction of transfer within the resolution of genetic mapping. Of course, this resolution is usually not sufficient to show whether the sites of integration are precisely the same at the level of 100 to 1,000 nucleotides.

In Fig. 16 we show the points of origin and direction of transfer of 25 independently derived Hfr's with points of origin in that part of the *E. coli* chromosomes contained on F13 (3, 6; Fig. 3 of ref. 16). For each case, with the two exceptions of OR7 and OR54, the position of the point of origin, to the accuracy that it is known by genetic mapping, and the polarity of transfer can be predicted by assuming that integration takes place by recombination of an $\alpha\beta$ or IS2 sequence of F with the homologous sequence on the *E. coli* chromosome. (One of the exceptions, OR54, is an unstable Hfr.)

The plasmid F152-1 was derived by a type I excision from Hfr P3 (21; Ohtsubo and Hsu, manuscript in preparation). The $\alpha_1\beta_1$ sequence of F occurs in F152-1 at one junction with chromosomal DNA. This $\alpha_1\beta_1$ sequence and the adjoining chromosomal sequences of F152-1 are identical with the $\alpha_4\beta_4$ sequence and adjoining chromosomal sequences of F13. These facts and the available genetic data are completely consistent with the model that Hfr P3 was formed by integration of $\alpha_1\beta_1$ of F with the sequence $\alpha_4\beta_4$ on the *E. coli* chromosome, followed by a type I excision. The proposed structure of Hfr P3 and the excision event to give the parental F152 are shown in Fig. 16.

Similarly, the $\alpha_3\beta_3$ sequence and adjoining chromosomal sequence of F13 are identical to the $\alpha_1\beta_1$ sequence at the junction of F DNA with chromosomal DNA of F42 (F*lac*) and adjoining chromosomal sequences. These results imply that Hfr P804 (the parent of F42) was formed by integration of the $\alpha_1\beta_1$ sequence of F with the $\alpha_3\beta_3$ sequence on the *E. coli* K-12 chromosome, followed by a type I excision process (Fig. 16).

In general, it is not known whether F integra-

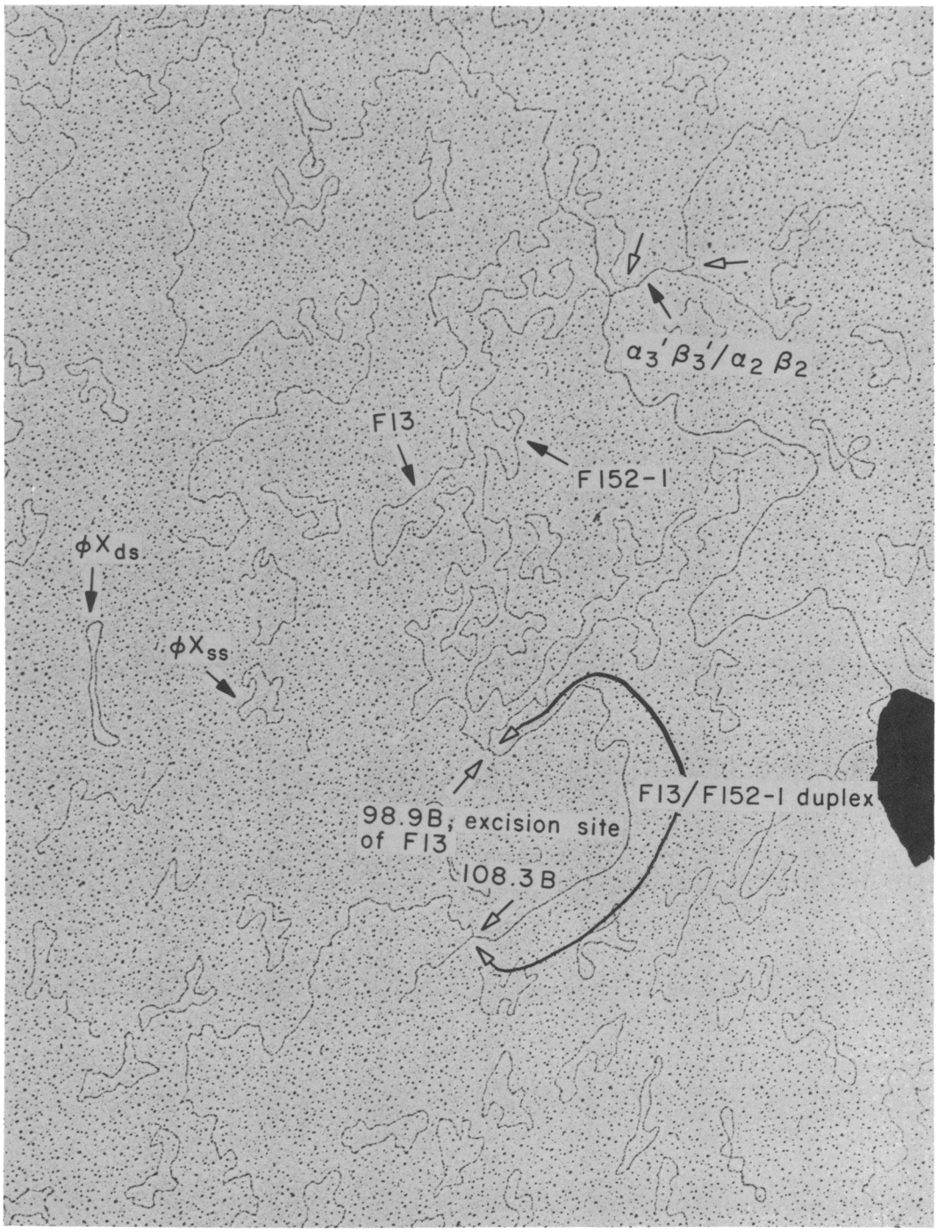

FIG. 14. *A heteroduplex molecule of F13 with F152-1. There is a duplex region from 98.9 to 108.3B representing the common chromosomal region in these two plasmids. There is an additional duplex region formed by pairing the* $\alpha'_3\beta'_3$ *sequence of F13 with the complementary sequence* $\alpha_2\beta_2$ *of F152-1.*

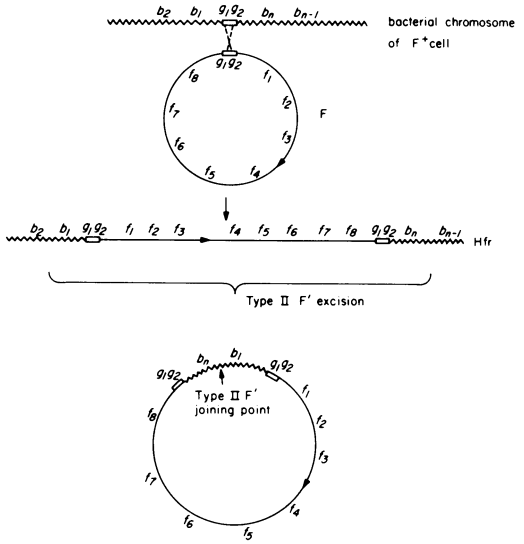

FIG. 15. *A model for integration of F into the bacterial chromosome by reciprocal recombination between the identical sequences g_1g_2, resident on the chromosome and on F. In the resulting Hfr chromosome and in the resulting type II F', the sequence g_1g_2 occurs, with the same polarity, at both junctions of F DNA with chromosomal DNA. The type II excision event depicted in the figure results in the joining of a point between b_n and b_{n-1} to a point between b_2 and b_1. Thus b_1 and b_n are joined in the plasmid.*

tion into the chromosome typically involves the joining between two points on the DNA molecules where there is little or no homology, as in illegitimate recombination (9) or as in site-specific λ recombination (10), or whether the joining takes place between two sequences which are homologous over a fairly large region. The discussion above suggests that for many cases of Hfr formation the reciprocal recombination occurs between special sequences 1.3 kb in length present on F and on the *E. coli* chromosome. It is clear that this general hypothesis can be tested by further heteroduplex studies of F' plasmids and/or segments of DNA prepared from Hfr chromosomes.

One qualification with regard to this general interpretation for the particular case of the formation of Hfr 13 should be noted. IS sequences seem to have the property of being readily translocated from one point in the *E. coli* chromosome to another (12, 13). Therefore, further work may show that IS2 and/or the several $\alpha\beta$ sequences that we propose were present on the chromosome of Hfr 13 are not present at the same positions on all *E. coli* K-12 strains, especially those that are only distantly related to Hfr 13. Indeed, Roy Curtiss (personal communication) has pointed out to us that Hfr 13, P3, and P804, Hfr's whose structures can be explained by our model, were all derived from the 58–161 subline of *E. coli* K-12, whereas Hfr OR7, which does not fit the model, was derived from the 112 subline (2).

Figure 16 also shows a proposed structure for Hfr OR7 and a model of the type I excision event giving F210. The type I joining takes place between 12.0F and point β_4. This single piece of evidence suggests that β may be a point which is active in half-site specific F recombination phenomena.

Two additional points deserve mention. F13 has a duplication of the IS2 sequence, one at each junction of chromosomal DNA with F DNA. F14 has a duplication of the 2.8–8.5F sequence, one at each junction of chromosomal DNA with F DNA (18). We may ask what is the frequency of a recombination event between the duplicated sequences on a single F' DNA molecule generating F and a circular molecule containing only the chromosomal sequences of the

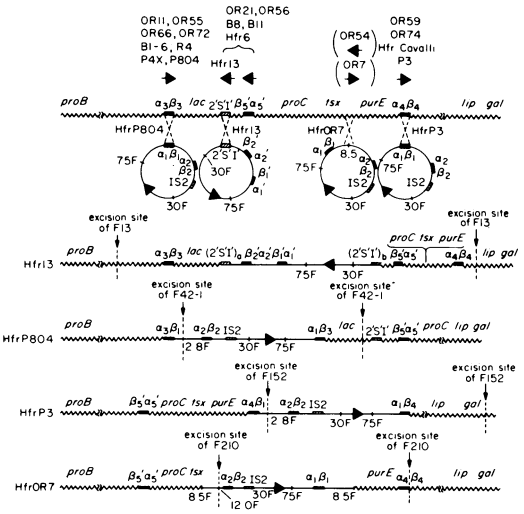

FIG. 16. *A model for the formation of Hfr's by reciprocal recombination between $\alpha\beta$ or IS2 on F with an $\alpha\beta$ or IS2 resident on the E. coli chromosome. The top of the figures is a map of the chromosome of E. coli K-12 in the region around lac. The positions of IS2 and $\alpha\beta$ deduced from the structure of F13 are shown. It is proposed that Hfr P3 and Hfr P804 were formed by recombination between $\alpha_1\beta_1$ of F with $\alpha_4\beta_4$ and $\alpha_3\beta_3$ of E. coli K-12, respectively. It is proposed that all the Hfr's OR11, . . ., P804, were formed by recombination of either $\alpha_1\beta_1$ or $\alpha_2\beta_2$ of F with $\alpha_3\beta_3$ of E. coli K-12. The Hfr's, OR21, . . ., Hfr 6, were formed by recombination either at 2'S'I' or at $\beta_5\alpha_5$, etc. The structures of several Hfr's and the excision events leading to the formation of certain F-primes are also depicted.*

F′ plus one copy of the duplicated sequence. Both F13 and F14 carry essential genes and in our case have been extracted from haploid hosts. Therefore, the segregation event would be lethal. Nevertheless, the resulting recombinant molecules are observed at a frequency of 5 to 10% in F14 extracted from the strain χ1254 (AB1206). There is evidence that this high frequency of segregation is *recA* independent (18). Hirota and Sneath (11) report that, when F13 is transferred into *lac*$^-$ recipients to give *lac*$^+$ bacteria, there are some later *lac*$^-$ segregants and a considerable fraction of these are F$^+$. However, we do not observe any circular DNA molecules of length 94.5 kb (F) and 290.1 kb (F13-F) in our preparations of F13. All this indicates that there is an observable rate of recombination between the IS2 sequences on F13, but that this recombination rate is lower than that of the two 2.8–8.5F sequences on F14 in AB1206. Thus, there are factors not yet understood which affect the recombination rate between repeated sequences on F in plasmids.

F13-4 is a fertile F-prime plasmid. It is deleted in the sequences of F extending from 17.6 to 37.8. This is additional evidence that this region of F is not essential for autonomous replication or for fertility (1, 7).

ACKNOWLEDGMENTS

We are grateful to our colleagues M.-T. Hsu and R. C. Deonier for advice and discussion.

This research was supported by Public Health Service grant GM 10991 from the National Institute of General Medical Sciences.

LITERATURE CITED

1. Anthony, W. N., R. C. Deonier, H. J. Lee, S. Hu, E. Ohtsubo, N. Davidson, and P. Broda. 1974. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. IX. Note on the deletion mutant of F, FΔ(33-43). J. Mol. Biol. **89**:647-650.
2. Bachmann, B. J. 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. **36**:525-557.
3. Broda, P. 1967. The formation of Hfr strains in *Escherichia coli* K-12. Genet. Res. **9**:35-47.
4. Broda, P., J. R. Beckwith, and J. Scaife. 1964. The characterization of a new type of F-prime factor in *Escherichia coli* K-12. Genet. Res. **5**:489-494.
5. Broda, P., and P. Meacock. 1971. Isolation and characterization of Hfr strains from a recombination-deficient strain of *Escherichia coli*. Mol. Gen. Genet. **113**:166-173.
6. Curtiss, R., III., and D. R. Stallions. 1969. Probability of F integration and frequency of stable Hfr donors in F$^+$ populations of *Escherichia coli* K-12. Genetics **63**:27-38.
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. Deonier, R. C., E. Ohtsubo, H. J. Lee, and N. Davidson. 1974. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. VII. Mapping the ribosomal RNA genes of plasmid F14. J. Mol. Biol. **89**:619-630.
9. Franklin, N. C. 1971. Illegitimate recombination, p. 175-194. *In* A. D. Hershey (ed.), The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
10. Gottesman, M. E., and R. A. Weisberg. 1971. Prophage insertion and excision, p. 113-138. *In* A. D. Hershey (ed.), The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. Hirota, Y., and P. H. A. Sneath. 1961. F′ and F-mediated transduction in *Escherichia coli* K-12. Jpn. J. Genet. **36**:307-318.
12. Hirsch, H. J., H. Saedler, and P. Starlinger. 1972. Insertion mutations in the control region of the galactose operon in *E. coli*. II. Physical characterization of the mutations. Mol. Gen. Genet. **115**:266-276.
13. Hirsch, H. J., P. Starlinger, and P. Brachet. 1972. Two kinds of insertions in bacterial genes. Mol. Gen. Genet. **119**:191-206.
14. Hsu, M. T., and N. Davidson. 1972. Structure of inserted bacteriophage Mu-1 DNA and physical mapping of bacterial genes by Mu-1 DNA insertion. Proc. Natl. Acad. Sci. U.S.A. **69**:2823-2827.
15. 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.
16. Low, K. B. 1972. *Escherichia coli* K-12 F-prime factors, old and new. Bacteriol. Rev. **36**:587-607.
17. Martuscelli, J., A. L. Taylor, D. J. Cummings, V. A. Chaplin, S. S. De Long, and L. Cãnedo. 1971. Electron microscopic evidence for linear insertion of bacteriophage Mu-1 in lysogenic bacteria. J. Virol. **8**:551-563.
18. 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**:565-584.
19. Scaife, J. 1967. Episomes. Annu. Rev. Microbiol. **21**:601-638.
20. Scaife, J., and A. P. Pekhof. 1964. Deletion of chromosomal markers in association with F-prime factor formation in *Escherichia coli* K-12. Genet. Res. **5**:495-498.
21. Sharp, P. A., M. T. Hsu, E. Ohtsubo, and N. Davidson. 1972. Electron microscope heteroduplex studies of sequence relations among plasmids of *E. coli*. I. Structure of F-prime factors. J. Mol. Biol. **71**:471-497.
22. Sugino, Y. 1966. Mutants of *Escherichia coli* sensitive to methylene blue and acridines. Genet. Res. **7**:1-11.
23. Taylor, A. L., and C. D. Trotter. 1972. Linkage map of *Escherichia coli* K-12. Bacteriol. Rev. **36**:504-524.
24. Willetts, M. S. 1972. Location of the origin of transfer of the sex factor F. J. Bacteriol. **112**:773-778.