Proc. Nat. Acad. Sci. USA
Vol. 69, No. 7, pp. 1948–1952, July 1972

Electron Microscope Studies of Heteroduplex DNA from a Deletion Mutant of Bacteriophage ϕX-174

(*E. coli*/phage lysozyme/length measurements/single-strand loops/DNA oligomers)

JUNG-SUH KIM, PHILLIP A. SHARP*, AND NORMAN DAVIDSON†

Department of Chemistry, California Institute of Technology, Pasadena, Calif. 91109

Contributed by Norman Davidson, December 30, 1971

ABSTRACT **A population of double-stranded replicative form of DNA molecules from bacteriophage ϕX-174 carrying a deletion of about 9% of the wild-type DNA has been discovered in a sample cultivated under conditions where the phage lysozyme gene is nonessential. The structures of deleted monomers, dimers, and trimers were studied by the electron microscope heteroduplex method. The dimers and trimers are head-to-tail repeats of the deleted monomers. Some interesting examples of the dynamical phenomenon of branch migration *in vitro* have been observed in heteroduplexes of deleted dimer and trimer strands with undeleted monomer viral strands from the wild-type phage.**

In our laboratory both single-stranded ϕX-174 DNA and the nicked double-stranded intracellular form, RF II, are used as internal standards for length measurements of other DNA molecules in the basic-protein film method of preparing samples for electron microscopy. The ϕX-174 DNA samples are prepared by members of the Sinsheimer group. We have observed that one of the RF-II ϕX DNA samples actually consists of a mixture of two kinds of molecules: in one kind, about 9% of the DNA is deleted; the other kind is of the normal size. We report here our electron microscope studies of the structure of this deleted DNA. Some interesting examples of the dynamical phenomenon of branch migration (1, 2) have been observed in heteroduplexes of a deleted dimer strand with undeleted, wild-type ϕX-174 DNA.

Our serendipitous discovery of a deletion mutant illustrates the effectiveness of electron microscopy as a tool for the study of DNA molecules. At the same time, Zuccarelli, Benbow, and Sinsheimer were independently conducting a systematic search for deletion mutants of ϕX. Their studies are reported in an accompanying publication [*Proc. Nat. Acad. Sci. USA* (1972) **69**, 1905–1910].

MATERIALS AND METHODS

DNA. A sample (sample **1**) of ϕX-174 RF(double-stranded) DNA was given to us by Dr. John Sedat. Its history appears to be as follows: Drs. C. Hutchison and M. Edgell grew a 300-liter batch of Amber 3 (lysis-defective) low-reverting phage, at a multiplicity of 1, using artificial lysis (3, 4). About 40% of the progeny were viable on a permissive host (Hfr 4714), indicating that the sample did not consist entirely of deletion mutants. About a year later, J. Sedat infected a 3-liter culture of the nonpermissive Hfr 4704 (*uvr*$^-$) host in the presence of 30 μg/ml of chloroamphenicol (which inhibits synthesis of the viral single-strands, but permits multiplication of the RF form) (5). After 90 min, the sample was lysed with Brij detergent and divided into a pellet and a supernatant fraction. The pellet contained 10% of all the RF DNA and was enriched in multiple-length DNA strands. We were given a sample of the supernatant fraction, which was relatively enriched in monomer forms. Sample **2**, a gift from R. Benbow, was similarly derived from the original Hutchison and Edgell culture, but had been enriched in multiple-length forms by sedimentation. It is probable that the deletion mutant was present and propagated in the original 300-liter growth.

Electron Microscopy. Our electron microscope techniques have been described (6). The essential point of the heteroduplex method is that if a duplex is formed by renaturation between a complete viral (plus) strand of wild-type ϕX DNA and a minus strand from a deletion mutant, the region present in the wild-type DNA, but deleted in the mutant, will form a single-strand loop in an otherwise duplex structure. By use of suitable concentrations of formamide in the spreading solution and in the hypophase for the basic-protein film technique, the single strand is displayed in an extended, measurable form

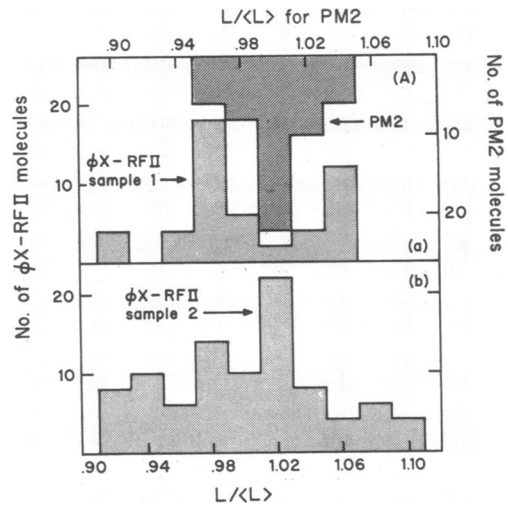

FIG. 1. Length distribution of PM2 DNA and of the ϕX-RFII DNA in each of the two samples. PM2 DNA and ϕX-RF DNA of sample 1 were spread on the same grid, from 40% formamide onto 10% formamide.

* Present address: Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. 11724.

† To whom reprint requests should be addressed.

(6, 7). Single-strand DNA looks a little thinner and more kinked than does double-strand DNA.

RESULTS

Existence of a Deletion. In a course of studies on heteroduplexes between T2 and T4 bacteriophage DNA, we frequently mounted both φX-RF (from sample 1) and λc_{26} DNA on the same grid as standards for measurements of double-strand length. The observed length ratio of φX-RFII to λDNA was 0.105(±0.002). Results of many other workers in this laboratory on other φX-RF samples had invariably given a ratio of 0.112(±0.002). Histograms of the length distribution of a homogeneous DNA (PM2) and of the present two φX-RF samples are shown in Fig. 1. The histogram and the length ratio to λ DNA suggest that sample 1 contains about 60% of a deleted form. To test this hypothesis, a mixture of the φX-RF DNA sample and a 10-fold excess of φX-174 single-strands (mostly intact) was denatured and renatured. When this mixture was examined in the electron microscope, there were a number of heteroduplexes showing a single, clear deletion loop (Fig. 2*a*).

Length measurements of the single-strand deletion loop give a size of 0.09(±0.02) relative to φX DNA. We take the molecular weight of double-stranded λ DNA as 30.8(±1) × 10^6 [46,500 nucleotide pairs (8)]; therefore, φX-174 DNA contains 5200 nucleotides, and the deletion consists of 470(±100) nucleotides. [The two main peaks in the histogram of Fig. 1*a* suggest a deletion of 0.07_5 (±0.02) φX units, in satisfactory agreement with the direct measurements.]

Sample 1 was denatured and renatured by itself. A deletion loop of the same size as seen above, due to a heteroduplex between deleted and nondeleted DNA in sample 1, was observed. For sample 2, deletion loops of size 490(±140) and 420(±110) were seen in the heteroduplex and self-renatured preparations, respectively. In sample 1 homoduplexes, no structure with substitution loops, two deletion-loops, or a

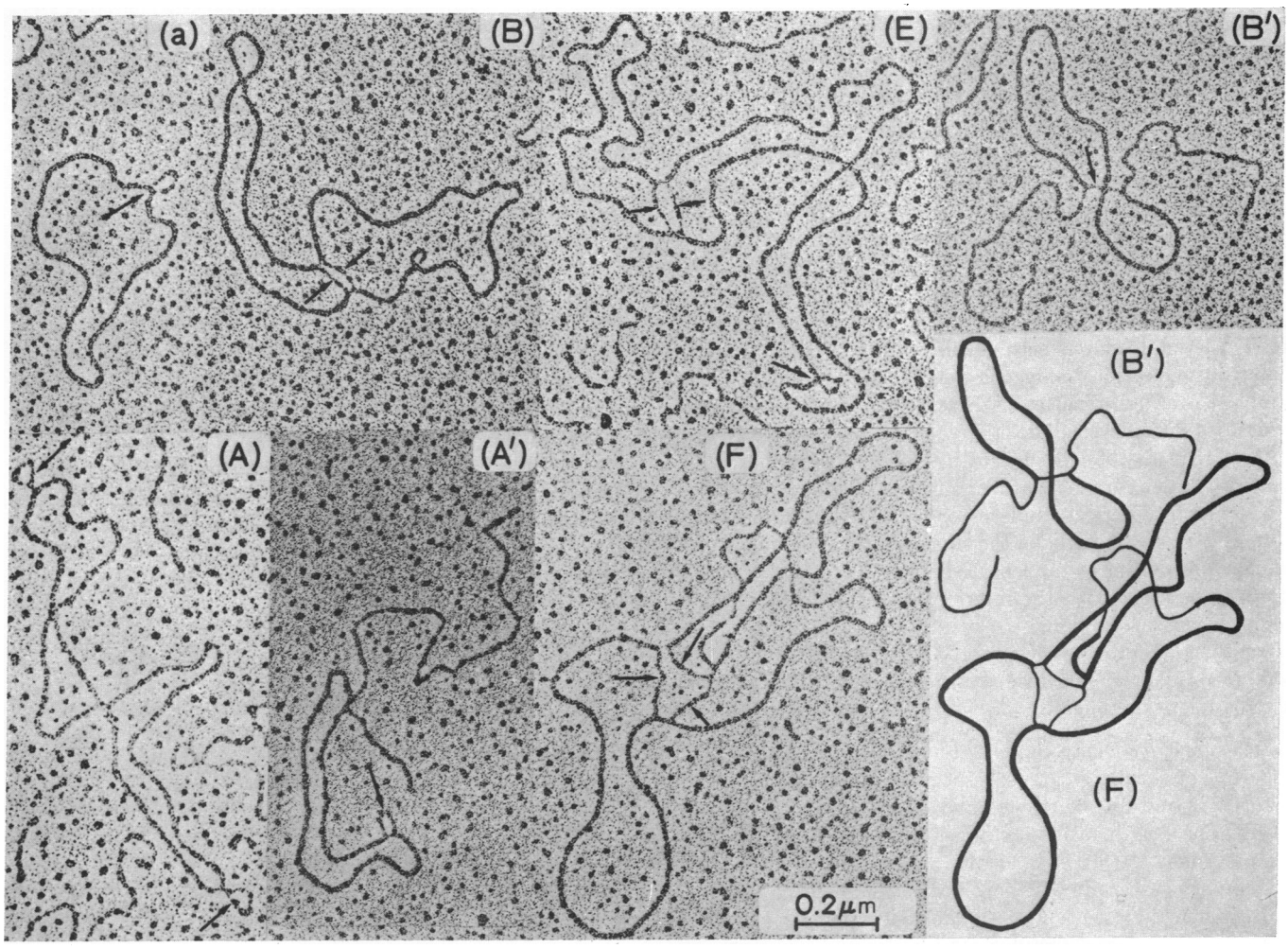

FIG. 2. (a) Electron micrograph of a monomer heteroduplex between a complete viral strand and a deleted complementary (minus) strand; the deletion loop is indicated by an *arrow*. (*A*) Heteroduplex type "structure *A*" between a deleted linear dimer strand and two circular viral strands. The two deletion loops are indicated by *arrows;* they are equidistant from the fusion point. A schematic representation is shown in Fig. 3*A*. (*B*) Heteroduplex of "structure *B*"; in this structure, the two deletion loops (*arrows*) have migrated to the fusion point; see Fig. 3*B*. (*A′* and *B′*) Heteroduplexes of a deleted linear dimer strand with only one circular viral strand, to give structures *A′* and *B′*, respectively; the deletion loop in *A′* and deletion strand in *B′* are indicated by *arrows*. Schematic representations are shown in Figs. 3*A′* and 3*B′*. A tracing of structure *B′* is shown at the *lower right* of Fig. 2; duplex regions are thick; single-strand regions are thinner. (*E* and *F*) Heteroduplexes of deleted trimer strands with three viral monomer circles. The deletion loops are indicated by *arrows*, schematics are in Fig. 3. A tracing of *F* is shown in the *lower right*. In this particular molecule, one of the duplex circles is incomplete because of single-strand breaks, as indicated in the tracing.

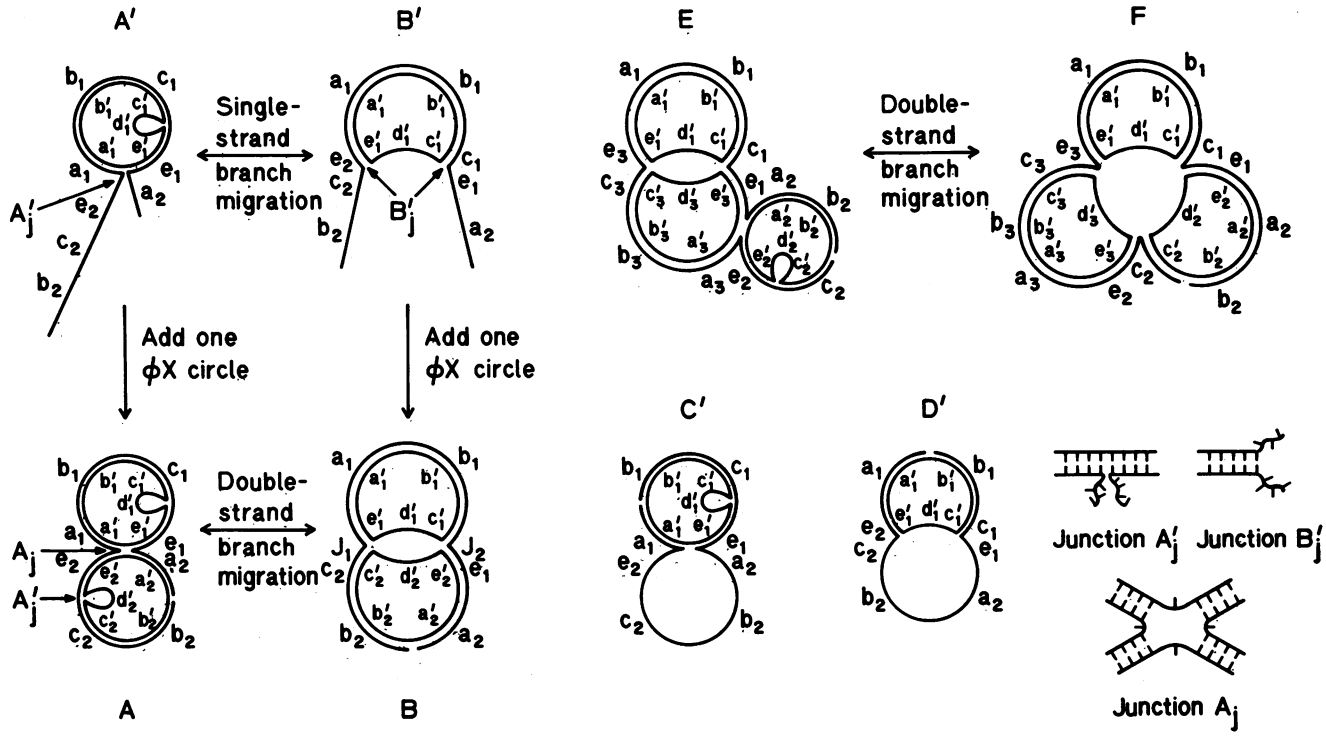

FIG. 3. Schematic representations of the several heteroduplex structures for which electron micrographs are shown in Fig. 2. The symbols are explained in the text. Micrographs of structures C' and D' are not included in Fig. 2. Schematic representations of the junctions A_j', B_j', and B, are also shown. The junctions labeled J_1 and J_2 in B have the same local stereochemistry as junction A_j'.

small deletion-loop was seen. Thus, within a resolution of about 50 nucleotides, only one deletion was present. When sample 2 was self-renatured, there was a low frequency of structures with two small deletion-loops, suggesting that there is a low frequency of molecules with a different nonoverlapping deletion. A denatured–renatured mixture of samples 1 and 2 showed mainly perfect duplexes or the single deletion-loop; samples 1 and 2 must have the same deletion. The observed standard deviation in the length distribution of the deletion loop is consistent with the interpretation that it is a single homogeneous deletion (6).

Another laboratory stock of φX-RF received earlier also contained about 1% of deleted molecules, as indicated by a heteroduplex experiment.

Structure of the Dimer. Samples 1 and 2 contained about 2% and 16% of the dimer forms, respectively, and slightly more than 0.2% and 3% of trimers and other multimers, respectively. The length ratios for dimer to monomer in the two samples were 2.04(±0.08) and 1.99(±0.11), respectively.

In order to ascertain whether the deleted dimer is a head-to-tail (tandem) duplication of the deleted monomer, we have observed the dimer structures after forming heteroduplexes from a sample with an 8-fold excess of circular φX viral strands. Photographs of the two structures observed are displayed in Fig. 2*A* and *B;* schematic explanations of the structures are given in Fig. 3*A* and *B*. (Sample 1 consisted mostly of RF-II nicked open-circular forms. In sample 2, about 60% of the molecules were nicked open-circles and 40% were RF-I supercoils. Supercoils do not participate in heteroduplex formation. Over 90% of the φX viral DNA was circular. Due to topological restrictions, only a linear strand can renature with a circular viral strand. Accordingly, in Fig. 3 we depict the heteroduplexes as having a nick in the deleted strand.) Structure *B* can only occur for a head-to-tail dimer. In any one heteroduplex of type *A*, the two distances from the fusion point of the two circles to the two deletion loops, shown as the sequences e_1 and e_2 in Fig. 3*A*, were equal. This distance was different in different heteroduplexes. The only reasonable interpretation of these observations is that the dimer is a head-to-tail duplication of the monomer.

Branch Migration. 150 Examples of structure *B* and 3 examples of structure *A* were observed in the sample scored. Renatured structures in which a linear dimer strand had associated with only one circular viral strand were also observed. The two structures mainly seen are displayed in Fig. 2*A'* and 2*B'*, and are schematically depicted in Fig. 3 as A' and *B'*. Observations on the relative numbers of the several structures after different renaturation times (Table 1) give a B' to A' ratio of about 2:1, independent of renaturation time.

The phenomenon of single-strand branch migration *in vitro* has been described (1). Its occurrence during *in vivo* recombination has been proposed (9). In structure *A'* in Fig. 3, there are two identical sequences, e_1 and e_2, complementary to the sequence e_1'. If the sequence e_1, shown mated to e_1', dissociates nucleotide by nucleotide and is replaced by the identical sequence e_2, there is no change in the number of base-pairs, but the deletion loop migrates with respect to the point where the two single-strand branches emerge from the duplex circle. This is branch migration. If migration occurs to the point where the sequence d' of the wild-type strand missing in the deleted DNA is between the two single-strand branches, structure *B'* results.

Similarly, in the duplex structure A, e_1 is mated to e_1' and e_2 to e_2'. By breaking and remaking base-pairs, one by one, e_1 can mate to e_2' and e_2 to e_1', resulting finally in structure B. This is the phenomenon of double-strand branch migration. Double-strand branch migration was clearly and explicitly described by Broker and Lehman (2), who proposed that it occurred *in vivo* as a part of the mechanism of genetic recombination between T-even phages.

If single-strand branch migration is slow compared to the rate of association of a linear, deleted dimer strand with a circular, φX viral strand, then the observed ratio of structure B' to A' of 2:1 is an indication of their relative rates of formation from the dissociated strands. If branch migration is fast, the ratio is primarily determined by the equilibrium constant of the reaction $B' \leftrightarrow A'$. It does not appear to us to be feasible, in the present system, to do decisive experiments to distinguish between the two hypotheses. We believe, however, that branch migration is sufficiently rapid so that B' and A' are at equilibrium with each other. This surmise is supported by two considerations: (*i*) the B' to A' ratio is constant with time; (*ii*) there is evidence that in the similar systems of renatured circularly-permuted DNAs from the phages of *Escherichia coli* 15 (1) and of T4 (unpublished data), branch migration is rapid.

Structure B' can react with a circular viral strand to give structure B. Similarly, A' gives A. We expect the rate-constants of the two reactions to be about the same. Since $B'/A' \approx 2$, the ratio B/A should be 2, unless there is interconversion between A and B by double-strand branch migration. Since we observe $B/A \approx 50$, we conclude that double-strand branch migration does occur under the renaturing conditions used, and that the equilibrium ratio of B to A under these conditions is about 50. We believe that this is rather direct experimental evidence for the occurrence of double-strand branch migration *in vitro*.

Several heteroduplexes of deleted trimers with three undeleted monomer circles were observed. Micrographs of the two kinds of structures seen are shown in Fig. 2*E* and *F*, and are schematically explained as structures E and F in Fig. 3. About equal numbers of structures E and F are seen. For a tandem trimer, it is a necessary condition in structures of type E that the distances e_1 and e_2 should be equal. This was observed to be the case. The three-leaf clover structure, F, is of course consistent with the head-to-tail tandem structure of the trimer.

If the equilibrium ratio of B to A is 50, we expect an equilibrium ratio of F to E of 50, contrary to observation. Thus, the indications are that the system is not at equilibrium, but that E may be formed by stepwise renaturation of φX circles to a trimer strand, and that the interconversion between structures F and E by double-strand branch migration is slow.

There is insufficient information as to the factors influencing the kinetics of branch migration. Several points, however, are sufficiently clear to merit discussion. Since branch migration processes involve successive opening and closing of base pairs, the rate should be faster under renaturing conditions than under conditions where base pairs are very stable. Indeed, the rate may increase as the solvent conditions approach T_m. There are topological constraints on the rate of branch migration in circular structures. In a branch migration reaction such as $A' \leftrightarrow B'$, one or the other of the dangling single-strands (presumably the shorter) must wind around through the circle once for every ten base-pairs exchanged by branch migration. This would slow down the rate for the present case, as compared to that observed in linear structures or in very large circular ones (1). On the other hand, in a double-strand branch migration such as $A \rightarrow B$, the two intersecting duplexes can rotate about their respective helix axes, and it is not necessary for one circle to rotate through the other. This is easy to see by construction of a model, but hard to depict in a drawing. However, in the branch migration of the trimer structure E to F, it is necessary for the circle with subscripts numbered 2 to rotate through circle 3 once for every ten base-pairs migrated. The duplex structures involved are more rigid and more extended in space than the single strands that wind through the circle for the $A' \rightarrow B'$ reaction. Therefore, we propose that branch migration from E to F is slow because of this topological constraint.

TABLE 1. *Numbers of dimer structures as a function of renaturation time**

Minutes of renaturation†	A'	B'	Uncertain A'‡	B
11	4	5	—	0
19	11	27	(26)	2
25	6	14	(24)	14
73	9	17	—	23

* These results, crude as they are, have been confirmed in repeat experiments with other samples.

† Sample 2 renatured with a 4-fold excess of viral strands, except that the 25-min point is with an 8-fold ratio of viral strands to RF II.

‡ These are ambiguous structures. Some are due to catenanes in the preparation, some are linear strands accidentally crossing duplex circles, and some are probably true A'.

Equilibrium Considerations. Several structural and statistical parameters influence the relative stabilities of structures A' and B' and of A and B. These parameters are not all known with sufficient accuracy to permit a theoretical prediction of the equilibrium ratios. However, by assuming that equilibrium obtains, we can compute a numerical value for the product of several unknown parameters that occur in the theory for use in interpreting future experiments on other systems.

There is a statistical factor, $g_1(A')/g_1(B')$, for structure A' relative to B', of 5000, because of the 5000 nucleotide sites at which the deletion loop can be situated around the circle in A'. A second statistical factor, $g_2(A')/g_2(B')$, favoring structure B', arises because of factors associated with the probability of ring closure. Structure A' consists of two closed-circles; a closed duplex loop of length $L_D = 5000$ nucleotide pairs and a closed single-strand loop of length $L_D = 500$ nucleotides, whereas structure B' contains only a single composite circle of duplex length, L_D, plus single-strand length, L_S. There are protruding single-strands in both A' and B'; these have the same partition functions. The two structures have the same number of base pairs. There is a third statistical weight ratio, $g_3(A')/g_3(B')$, because the two structures have different kinds of junctions between duplex regions and single-strand regions.

The quantity $g_2(A')/g_2(B')$ may be calculated according to the well known Jacobsen–Stockmayer (10) treatment for ring closure in random coils. The statistical weight of a circular duplex, relative to a linear duplex, is $(3/2\pi L_D b_D)^{3/2}\, dV_D$,

where b_D is the Kuhn statistical segment length of duplex DNA, and dV_D is a volume element for double-strand ring closure (11). Similarly, the statistical weight of a single-strand deletion loop of length L_S (relative to the linear strand of the same length) is $(3/2\pi L_S b_S)^{3/2}\, dV_S$, where b_S is the Kuhn statistical segment length of single-strand DNA under renaturing conditions. Structure B' contains a composite circle consisting of a duplex of length L_D and a single strand of length L_S. The statistical weight of this structure is $[3/2\pi(L_S b_S + L_D b_D)]^{3/2}\, dV_{DS}$, where dV_{DS} is the volume element for ring closure at a junction between double- and single-strand DNA. In the present instance, $L_D b_D \gg L_S b_S$; therefore,

$$\frac{g_2(B')}{g_2(A')} = \frac{[3/2\pi(L_S b_S + L_D b_D)]^{3/2}\, dV_{DS}}{(3/2\pi L_S b_S)^{3/2}\,(3/2\pi L_D b_D)^{3/2} dV_D dV_S} \approx \frac{1}{(3/2\pi L_S b_S)^{3/2} dV}, \quad [1]$$

where dV represents $dV_D\, dV_S/dV_{DS}$.

To apply Eq. [1], we take a contour length of 3.2_2 Å per nucleotide for both single- and double-strand DNA. Thus, for the ϕX deletion loop, $L_S = 0.16_1$ μm. The principal uncertainties are the values of dV and of b_S. We must estimate b_S for denatured DNA under renaturing conditions, where no hydrodynamic measurements are available. Arguments similar to those given by Wetmur and Davidson (12) lead to an estimate of 45 Å. We expect dV to be of the order of 10^{-21} cm^3, and write $dV = f\,10^{-21}$ cm^3, where f is hopefully of the order of unity. We then find that

$$[g_1(B)/g_1(A')][g_2(B)/g_2(A')] = 12/f \quad [2]$$

The factors affecting $g_3(B')/g_3(A')$ are the following: Structure A' contains two junctions of the type identified as A_j' in Fig. 3. Structure B' contains two junctions depicted as Bj' in Fig. 3. There may be some steric crowding in Aj', relative to Bj', but there is a favorable stacking interaction in Aj' that is lost in Bj'. We write $g_3(B')/g_3(A') = [g(Bj')/g(Aj')]^2$. If the equilibrium ratio, $g_1(B')g_2(B')g_3(B')/g_1(A')g_2(A')g_3(A')$, is 2, as suggested by the experiments, then from [2]

$$(1/f)[g(Bj')/g(Aj')]^2 = 2/12 = 1/6 \quad [3]$$

This numerical parameter, deduced from our experimental data, is available for prediction of equilibria in other systems.

Structures with the topology of C' and D' (Fig. 3) are seen at a frequency of about 1/10 A' and B', respectively, although they would be expected to be unstable. This frequency suggests that the relative rates of branch migration and of renaturation are such that the observed system is not quite at equilibrium.

The stabilities of the duplex structures A and B can be similarly discussed. Structure A involves two closed duplexes and two closed single-strand deletion loops. Structure B is more complicated. It can be shown that the statistical weight factor for the several ring closures is given approximately by

$$\frac{g_2(B)}{g_2(A)} = \frac{[3/2\pi(2L_S b_S)]^{3/2}\, dV}{\{[3/2\pi L_S b_S]^{3/2} dV\}^2} = \frac{1}{2^{3/2}(3/2\pi L_S b_S)^{3/2} dv} = \frac{1/3 \times 10^4}{f} \quad [4]$$

There is a statistical factor $g_1(B)/g_1(A) = 1/5000$. Structures A and B both contain two junctions of type A_j'. (These junctions are labeled J_1 and J_2 in Fig. 3*B*.) Structure A contains the additional junction labeled A_j in Fig. 3, where four duplex strands meet. There are presumably several base-pairs broken in this junction to accommodate the steric strains; thus, we expect $g_3(B)/g_3(A) = 1/g(Aj) > 1$. If the equilibrium ratio of B to A is 50:1, we find that

$$\frac{g_1(B)}{g_1(A)} \cdot \frac{g_2(B)}{g_2(A)} \cdot \frac{g_3(B)}{g_3(A)} = \frac{1.3 \times 10^4}{(5 \times 10^3) f g(A_j)} = 50 \quad [5]$$

or $fg(A_j) \approx 0.05$. This product is also available for application to other problems.

This research was supported by NIH Grant GM 10991. P. A. S. was supported by a USPHS fellowship. Our indebtedness to John Sedat is evident from the text. We wish also to acknowledge cordial interaction with A. Zuccarelli, R. L. Benbow, and R. L. Sinsheimer.

1. Lee, C. S., Davis, R. W. & Davidson, N. (1970) *J. Mol. Biol.* **48**, 1–22.
2. Broker, T. R. & Lehman, I. R. (1971) *J. Mol. Biol.* **60**, 131–149.
3. Hutchison, C. A. & Sinsheimer, R. L. (1966) *J. Mol. Biol* **18**, 429–447.
4. Eisenberg, M. G., Benbow, R. M. & Sinsheimer, R. L., *J. Mol. Biol.*, in press.
5. Komano, T. & Sinsheimer, R. L. (1968) *Biochim. Biophys. Acta* **155**, 295–298.
6. Davis, R. W., Simon, M. & Davidson, N. (1971) in *Methods in Enzymology*, ed. Grossman, L. (Academic Press, New York), Vol. XXID, pp. 413–428.
7. Westmoreland, B. C., Szybalski, W. & Ris, H. (1969) *Science* **163**, 1343–1348.
8. Davidson, N. & Szybalski, W. (1971) in *The Bacteriophage Lambda*, ed. Hershey, A. D. (Cold Spring Harbor Laboratory, New York), pp. 45–82.
9. Cassuto, E. & Radding, C. M. (1971) *Nature New Biol.* **229**, 13–16; **230**, 128–128.
10. Jacobsen, H. & Stockmayer, W. H. (1950) *J. Chem. Phys.* **18**, 1600–1606.
10. Jacobsen, H. & Stockmayer, W. H. (1950) *J. Chem. Phys.* **18**, 1600–1606.
11. Wang, J. C. & Davidson, N. (1966) *J. Mol. Biol.* **19**, 469–482.
12. Wetmur, J. G. & Davidson, N. (1968) *J. Mol. Biol.* **31**, 349–390.