SPECIFIC DNA RESTRICTION FRAGMENTS AS INTERNAL MARKERS IN THE ELECTRON MICROSCOPE: SUPERIMPOSING THE ϕX174 GENETIC MAP ON ϕX174/S13 DNA HETERODUPLEXES

Thesis by

John Lee Compton

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1975

(Submitted September 27, 1974)

To my Parents

for their love, encouragement and generosity

Acknowledgments

My primary acknowledgment is reserved for my advisor, Robert Sinsheimer, who has extended to me the hospitality of his laboratory for the past five years. Although he was often too quick in picking out the weak points in my arguments he was equally quick to suggest and help and he has taught me the art of science by his example. Only his minor flaws, such as being unable to appreciate the thrill of hitting the opposing scrum-half before he releases the ball or the agony of slicing out of bounds on the 18th, have lessened our relationship. I shall be indebted to him forever.

Secondly, I would like to thank Norman Davidson for his patient instruction in experimental design and data analysis. Whenever I would assume one thing he would assume the opposite and force me to prove him wrong. This was usually difficult, but always instructive; I assume I'll never regret it.

I would like to thank Jerome Vinograd, Bill Wood and Giuseppe Attardi, who rounded out my committee, for their patience and help and for treating me as a colleague as I learned.

Within our laboratory and throughout the Division there have been many people who have helped the years pass more productively and more happily. To Gloria, who could bring biochemical companies to their knees but who couldn't seem to figure out how to put ice in the seminar tea when it was 100°F outside; to Mrs. J., who was never daunted by absolutely incredible piles of glassware and who always had a smile and a laugh for me; to Jane and Bill and Ed, who endured those

4:58 PM buying sprees; to Polly for her greetings; to Eleanor for her monetary manipulations; and to Geraldine for the inside track to the inner sanctum, may I express my gratitude and best wishes.

And to my colleagues: Bob for his fellowship within our small room as well as for his decals and his pig; Lloyd for his unbounded generosity, humor and artistic advice; Paul for knowledge; Karel for pronouncing Budejovice for me; and Ali, Romillio, Amy, Pieter, Reinhardt, Lois, etc. for fun and for help.

I am pleased to express special thanks to Pat Koen, whom I could usually outdrive but seldomly outshoot, who kept me in the subbasement on too many beautiful days by always doing his best to keep the microscopes running.

And finally to Cathie who turned a practice golf ball into a dream come true.

Writers cramp is the crippling enemy of the thesis writer and I have strength left for only the most important acknowledgments. I especially wish to thank NIH for constant support which allowed me to purchase from Anheuser-Busch and Acushnet, without whose products I would have retained even less of my sanity.

ABSTRACT - CHAPTER I

ϕX174 RF II* molecules were constructed *in vitro* by annealing circular ϕX174 viral DNA and linear ϕX174 complementary strand DNA. When a specific fragment of ϕX174 RF purified from a digestion with the restriction endonuclease Hind from *Hemophilus influenzae* was included in the renaturation mixture ϕX174 RF II molecules containing a loop corresponding in size to the fragment were seen. This was confirmed for three different DNA fragments.

ϕX174 RF II molecules were constructed in the presence of two restriction fragments. In addition to molecules containing one loop corresponding in size to one of the fragments molecules containing two loops were seen. The two loops corresponded in size to the two restriction fragments and the distances between the two loops corresponded to the distances between the two fragments on the Hind cleavage map of ϕX174 RF.

By including two fragments of distinguishable size that map asymmetrically on the cleavage map in the renaturation mixture, ϕX174 RF II molecules were constructed which could be oriented on the cleavage map. Using the two loops as internal markers the ϕX174 genetic map was superimposed on the electron micrographs of these molecules.

ABSTRACT - CHAPTER II

ϕX174/S13 DNA heteroduplexes were constructed by annealing purified ϕX174 complementary strand DNA and S13 viral DNA. The heteroduplexes were mounted for electron microscopy under a series of increasingly denaturing conditions and photographed. The micrographs were measured and oriented to form a map showing the progressive denaturation of various regions of the molecule.

Such heteroduplex molecules were totally denatured and mixed with a denatured specific fragment of ϕX174 RF* purified from a digestion with the restriction endonuclease Hind of Hemophilus influenzae. After renaturation the molecules were mounted for electron microscopy under standard conditions. Each of three different ϕX174 RF fragments could be seen in a different, specific region of the heteroduplex and could be oriented with respect to the others to correspond to the known ϕX174 cleavage map.

The ϕX174 genetic map was thus superimposed on the heteroduplex denaturation map by using the included fragments as reference points. Under mildly denaturing conditions three regions of the heteroduplex are conserved as double-stranded: two regions of about 10% of the ϕX174 genome each in genes A and H and a region of about 25% of the ϕX174 genome covering gene E and most of gene F. Under increasingly denaturing conditions only two small duplex regions of 2 and 4% are conserved, one each in genes E and F.

TABLE OF CONTENTS

General Introduction

1. ϕX174 Background

ϕX174 is an icosahedral bacteriophage with a diameter of 250 Å which contains a circular, single-stranded DNA genome with a molecular weight of 1.7×10^6 daltons. ϕX174 contains nine identified cistrons.

(a) The structure of ϕX174

Sinsheimer (1959a) published an electron micrograph of ϕX174 and described the phage as a polyhedral object about 300 Å in diameter. Hall et al. (1959) described the structure of ϕX174 as twelve regularly arranged knobs in the shape of a dodecahedron. They measured the diameter of the particle as 250 Å and calculated that a sphere containing Sinsheimer's 4.6×10^6 daltons of protein at a density of 1.3 gm/cm^3 and 1.7×10^6 daltons of DNA at a density of 1.7 gm/cm^3 would have a diameter of 240 Å. Tromans & Horne (1961) confirmed the icosahedral structure of the phage and Brown et al. (1971) published an electron micrograph of a freeze-etch replica of ϕX174 attached to the surface of a bacterial cell which shows the structure very well.

Sinsheimer (1959b) observed a splitting in the sedimentation boundary of ϕX174 DNA under conditions which minimized intramolecular interactions and hydrogen bonding. The proportions of the two components varied in different preparations and no nucleotide composition differences could be detected. Maclean & Hall (1962) observed that the DNase sensitive strands extruded by partially decomposed phage particles did not branch and were quite rigid except for a short region near the tip which bent back on itself. Fiers & Sinsheimer (1962a) showed that ϕX174 DNA was resistant to exonucleases but were unable to discount the

possibility of 3'-terminal phosphate groups which would block the action of the exonucleases. They were, however, able to show (Fiers & Sinsheimer, 1962b) that random endonuclease hits converted the faster sedimenting DNA form into its slower sedimenting companion. They concluded that the genome of ϕX174 is circular and thus would have no free 3'-OH group susceptible to exonuclease attack. A circular DNA also explains the loops seen at the ends of the otherwise rigid DNA strands. The circularity of the double-stranded replicative form was subsequently confirmed by electron microscopy (Kleinschmidt et al., 1963).

With a genome of 1.7×10^6 daltons ϕX174 was the smallest known virus with a DNA genome. Sinsheimer (1959b) also reported a series of criteria by which the ϕX174 DNA differed from native double-stranded DNAs: the DNA bases reacted with formaldehyde; no hyperchromicity was observed over a wide range of probable melting temperatures; the ultraviolet absorption spectrum behaved as a function of temperature like denatured DNA and as a function of ionic strength at ionic strengths which did not affect native DNA; it was too flexible as determined by light scattering. He concluded that these and other observations could be explained if the DNA were single-stranded rather than double-stranded.

(b) The process of infection with ϕX174

The process of infection begins with the reversible attachment of the phage to a susceptible host cell wall. Phages eluted from cells by washing with borate-EDTA are still infective (Newbold & Sinsheimer, 1970). Electron micrographs show that attached phages are buried to

about one-half their diameter in the cell surface and have their five-fold axis of symmetry normal to the surface (Brown et al., 1971). Attachment is followed by eclipse, during which an irreversible conformational change occurs in the phages. Eclipsed particles which have either spontaneously detached or have been removed from the cell wall by vigorous washing have 40% or more of their DNA protruding from the protein coat and are uninfective (Newbold & Sinsheimer, 1970). In spite of this DNA release eclipsed particles remaining on the cell surface appear to be still associated with only an outer layer of the cell wall (Knippers et al., 1969). During the next stage the phage DNA penetrates the cell wall. The simultaneous synthesis of a complementary strand is not required (Knippers et al., 1969), thus eliminating the possibility that the DNA is pulled into the cell by the cell's active requirement for template to complete its participation in this step.

The synthesis of the complementary strand is accomplished by pre-existing host enzymes unless one or more of the ϕX174 capsid proteins enters the cell with the DNA and is necessary for its synthesis (Greenlee & Sinsheimer, 1968; see Jazwinski et al., 1973). Salivar & Sinsheimer (1969) and Yarus & Sinsheimer (1967) have shown that this parental replicative form (RF) occupies one of a limited number of special sites within the cell. The parental RF is transcribed and the phage-specific mRNA translated before RF replication begins because the phage gene A product is necessary for the formation of progeny RF molecules (Tessman, 1966; Lindqvist & Sinsheimer, 1967). Henry & Knippers (1974) have purified the gene A product and have shown that it is a DNA endonuclease which specifically cleaves ϕX174 viral DNA and

the viral strand of ϕX174 RF. The parental RF is replicated semi-conservatively (Denhardt & Sinsheimer, 1965) for the first 12 to 15 minutes after infection at which time RF and host DNA synthesis stop (Lindqvist & Sinsheimer, 1967; Hutchison & Sinsheimer, 1966). Henry and Knippers (1974) have shown that the ϕX174 gene A product is a DNA endonuclease which cleaves ϕX174 RF I and ϕX174 single-stranded circular DNA once, and propose that the site of action is within a self-complementary hairpin region of the type suggested by the work of Schekman and Ray (1971) and Francke and Ray (1971).

Release of the progeny phage occurs in a single short burst from individual cells, but the onset of lysis for different cells varies over a period of approximately one bacterial generation (Hutchison & Sinsheimer, 1963). Lysis requires the participation of

the proteins coded for by genes E, H and J (Hutchison and Sinsheimer, 1966; Benbow et al., 1972). Benbow et al. (1972) speculated that the gene J product is the smallest structural component of the virion and that it is required in large amounts to interact with the gene E product to produce normal lysis.

(c) The genetics of ϕX174

The ϕX174 genome contains enough DNA to code for approximately ten proteins. Benbow et al. (1971) mapped temperature sensitive and nonsense mutants in eight cistrons using two-, three- and four-factor crosses. Hayashi & Hayashi (1970) described a ninth cistron, I, but Benbow et al. (1974) reported that the statement of the existence of I had been retracted. Benbow et al. (1972) proposed a ninth cistron themselves, however, which they named J and described as being necessary for normal lysis of the host cell.

Benbow et al. (1971) reported unusually high recombination frequencies between certain mutants in cistron A. It is possible that such an abnormally high level of recombination could result from the frequent nicking of the DNA within the cistron A region. Henry and Knippers (1974) believe that the site of action of their cistron A endonuclease, described in the preceding section, is within cistron A, which would provide such a high frequency of nicking.

2. The Relationship Between ϕX174 and S13

ϕX174 and S13 are bacteriophages which infect the same host. They have each been extensively studied in different laboratories and show a wide variety of similarities and differences.

(a) Basic similarities

Zahler (1958) reported that anti-serum prepared against S13 was effective against ϕX174, although the kinetics of inactivation were different. Maclean & Hall (1962) looked at ϕX174 and S13 in the electron microscope and reported that the two were indistinguishable. Spencer et al. (1972) published a study of the physical properties of S13 viral and replicative form DNA. The values they obtained are consistent with those obtained for ϕX174 by other workers and are included in Table 1.

(b) Genetic similarities

Since the phage contain much of the information necessary for reproducing themselves similarities in the physical particles should reflect genetic similarities. Tessman & Schleser (1963) observed genetic recombination between the two phages and complementation in mixed infections (Hutchison, 1968; Jeng et al., 1970). Genetic mapping of the two phages has shown that the order of the cistrons is the same (Benbow et al., 1971; Baker & Tessman, 1967), and physical studies have shown that the direction of translation of the two phages is the same (Benbow et al., 1972; Vanderbilt et al., 1971; Tessman et al., 1967).

(c) Basic differences

In spite of the extent of similarity between ϕX174 and S13 it is quite clear that they are not the same phage. Zahler's antiserum experiments showed that ϕX174 was inactivated at an initial rate between 0.25 and 0.50 that of S13. This probably reflects a difference in antigens present on the phage protein coats. Maclean & Hall (1962)

reported that S13 was much less stable than ϕX174 with respect to structure and biological activity, presumably also a difference in the protein structure of the virus. Although the proteins translated from the two phage mRNAs had been considered to be indistinguishable in earlier experiments, Godson (1973) reported that only one of the S13 proteins comigrated with its ϕX174 equivalent on polyacrylamide gels while the other eight varied in size by as much as ten amino acids. He estimated that the S13 capsid protein (gene F) had a molecular weight of 51,000 daltons as compared to 50,000 daltons for the ϕX174 capsid protein. Poljak & Suruda (1969) reported that the amino acid compositions and tryptic peptide maps of the major capsid proteins from ϕX174 and S13 were indistinguishable.

If the differences observed above reflect actual differences in phage-coded proteins they would reflect differences in the DNAs of the two phages. Table 2 shows that S13 viral DNA is more pyrimidine rich than that of ϕX174. Darby *et al.* (1970) published a catalogue of the pyrimidine tracts found in the complementary strand of ϕX174*am*3 and compared it with a similar catalogue prepared for the complementary strand of S13*su*N15 by Cerny *et al.* (1969). The pyrimidine distributions are very similar in size, number and composition. Each phage complementary strand contains three undecanucleotides, no decanucleotides and similar numbers of smaller tracts. Although the numbers of the smaller tracts vary between the two phages there is a clear correlation between the two distributions; individual tracts are either common or rare in each phage. Darby *et al.* (1970) conclude that ϕX174*am*3 and S13*su*N15 are derived from a recent common ancestor.

Razin et al. (1970) report that ØX174 viral DNA contains one 5-methylcytosine residue and conclude that it is either p(5-m)CpTp or pTp(5-m)Cp flanked on either side by a purine. Cerny et al. (1969) observed a C_6T_3 tract in their depurination study which migrated in an abnormal manner. They suggested that it could contain a 5-methylcytosine residue. If this is true the region of DNA containing the 5-methylcytosine residue has undergone some divergence between the two phages, which suggests that it may not be an essential control sequence.

(d) The relationship between ØX174 and S13

ØX174 and S13 were independently isolated eight years apart (Sertic & Boulgakov, 1935; Burnet, 1927). The survey just completed raises questions about the relatedness of the phages. They seem to be distinct and yet by many criteria they are closely related. There are three possible relationships between the two phages that define two different ways of looking at them.

The first possibility is that ØX174 and S13 were separately isolated, were distinct phages and are available today in a form which reflects this separate origin. The second possibility is that the phages were separately isolated but that the two phages available today are descended from only one of the two isolations. The third possibility is that the two isolations were not independent.

The conclusion by Darby et al. (1970) that the two phages have a recent common ancestor is probably correct, but the question is whether "recent" would mean since 1927 or in the last few thousand (million) years. Spencer & Boshkov (1973) reported that the pyrimidine

tracts present in wild type S13 are indistinguishable from those of the previously characterized S13suN15 (Cerny et al., 1969). There are no transitions or point mutations in any of the longer tracts, specifically those seven to eleven in length. This study suffers from the same difficulty in quantitating the smaller tracts as described by Darby et al. (1970) but it does indicate that any changes that occurred during the derivation of S13suN15 from S13wt or that have occurred to either phage since the initial mutational event have occurred in both phages or have been of a minor nature compared to the differences between ϕX174 and S13.

Irrespective of any possible post-isolation relationship between ϕX174 and S13 it is clear that both of the present phage strains can differ from the phage originally isolated. Davis & Hyman (1971) reported the presence of a deletion in the genome of a T7L strain purchased from the American Type Culture Collection and Kim et al. (1972) reported the presence of a deletion in ϕX174am3 stocks which had been grown for several generations without cloning. These are clearly recent changes since T7 originated as a single plaque in 1947 and ϕX174am3 without the deletion is available in the same laboratory as is the deletion. Neither strain exhibited other detectable sequence heterogeneity such as was found in hybrids between T7 and T3 (Davis & Hyman, 1971). Characterization of the type and extent of sequence heterogeneity between ϕX174 and S13 would help establish the probable time of divergence from their common ancestor.

3. The Development of Formamide as a DNA Denaturing Solvent in Electron Microscopy

The process by which DNA denatures and renatures has been extensively studied since DNA was shown to be double-stranded. The development of conditions under which the double-stranded structure is melted into single-stranded loops in reproducible positions has provided a powerful technique for the physical mapping of DNA.

(a) Denaturation and renaturation of DNA

Double-stranded DNA has long been called "native DNA." This reflects the early findings of double-stranded DNA and the effort which resulted in the elucidation of the hydrogen bonded, double helical structure of double-stranded DNA. By implication single-stranded DNA is non-native, and double-stranded DNA which has been melted, i.e., which has had its two strands separated and is therefore no longer native, is said to be "denatured" and to be "renatured" when its component strands are once again joined.

Doty and his coworkers (see Rice & Doty, 1957; Marmur & Doty, 1959) extensively studied the thermal denaturation of double-stranded DNA. In 1960 Marmur & Lane, and Doty et al. published a joint study which examined the renaturation of thermally denatured DNA by slow cooling from the denaturing temperature. Doty et al. (1960) examined the physical properties of this renatured DNA and showed that it had the ordered characteristics of the original native DNA. Marmur & Lane (1960) looked at the ability of the renatured DNA to transform bacterial cells. By demonstrating that the renatured product was active

in a biological system they confirmed that the DNA was once again native in a biological as well as a physical sense.

The study of the kinetics of renaturation of denatured DNA became an important method of characterizing eukaryotic DNA which was far too large to be treated in the same way as viral or bacterial DNA. Britten & Kohne (1967) developed a means of plotting the renaturation of DNA relating initial concentration and time of renaturation that could be applied to any DNA sample. They showed that the use of hydroxyapatite chromatography to discriminate between single- and double-stranded DNA could be an important method in determining the organization and degree of repetition of DNA sequences in DNAs from different sources (see Britten & Kohne, 1968).

The complete thermal denaturation of DNA requires very high temperatures. Rice & Doty (1957) showed that even at temperatures well above the DNA melting temperature some segments of in-register base pairing remained intact in large DNAs. Prolonged exposure to high temperatures degraded the DNA and it became clear that the combination of necessary time and temperature to ensure the complete melting of DNA was too severe to maintain its integrity. An alternative appeared when Helmkamp & Ts'o (1961) and Ts'o et al. (1962) showed that double-stranded DNA in formamide behaved like single-stranded DNA with respect to optical rotatory dispersion and circular dichroism. They concluded that formamide removed all formal secondary structure from DNA.

Bonner et al. (1967) reported the serendipitous discovery that by adjusting the salt concentration RNA could be made to hybridize to DNA in formamide at temperatures ranging from 0°C to 24°C. They

reported that the use of the lower temperatures led to lower backgrounds and less degradation than the normal hybridization at about 66°C. Formamide is a denaturing solvent and can be used in place of heat to melt DNA. The thermal melting temperature (T_m) of DNA is a function of salt concentration and the stability of double-stranded nucleic acids in formamide was clearly affected by the ionic conditions. Careful standardization of formamide-ionic strength effects on the T_m of nucleic acids would permit the substitution of formamide and lower temperatures for the previously necessary high temperatures in dissociation and reassociation/hybridization reactions. Tibbetts et al. (1973) have reported the use of formamide to change the elution properties of a hydroxyapatite column.

(b) Electron microscopy of DNA

The development of the basic protein film method for visualizing DNA (Kleinschmidt & Zahn, 1959) allowed people to see the molecules they were working with in the electron microscope. The technique proved its usefulness right away by confirming the circular nature of ϕX174 RF DNA (Kleinschmidt et al., 1963). Its use was restricted to native, double-stranded DNA, however, because the large intramolecular forces present in single-stranded DNA caused it to form a compact clump. Davis & Davidson (1968) utilized the single-strand clumps formed in hetero-duplex DNA molecules made from different mutants of λ to map deletions and substitutions. DNA regions without corresponding sequences on the other strand remained single-stranded and collapsed. For the first

time physical distances measured on the DNA molecules could be correlated with map distances determined by genetic crosses.

Westmoreland et al. (1969) combined the DNA secondary structure relaxation properties of formamide with the Kleinschmidt basic film DNA spreading technique and published electron micrographs of heteroduplex λ DNA molecules similar to those of Davis & Davidson (1968) with the single strand DNA regions spread out rather than clumped. Davis et al. (1971) studied the effects of various agents on stabilization of the single-stranded structures and on the contrast between the DNA and the supporting film, and introduced the use of single- and double-stranded ϕX174 DNA standards to calculate the statistics of DNA molecule length measurements. Progress in this technique has made the visualization of single-stranded DNA in the electron microscope a routine procedure.

Marmur & Doty (1959) showed that the temperature at which double-stranded DNA separated into its component strands was a function of its base composition. Under their ionic conditions the melting temperature ranged from 69°C for poly A-T to 110°C for poly G-C. The difference in stability between an A-T base pair and a G-C base pair presented another method for mapping DNA. Increasing pH mimics the effect of increasing temperature and DNA regions which would be expected to separate at a particular temperature would be expected to separate at a corresponding pH. Inman & Schnös (1970) prepared electron microscope grids of λ DNA under conditions of different pH. They found specific, reproducible single-stranded regions and were able to divide the molecules into an A-T rich half, which contained unstable regions, and a G-C rich half which did not.

Davis & Hyman (1971) substituted formamide concentration for pH as a probe of DNA stability and looked at heteroduplex molecules formed by combining DNA strands from two sources. They were able to show that hybrid molecules formed between bacteriophages T7 and T3 DNA strands contained regions of DNA which denatured under conditions which would not separate homologous DNA strands. As the formamide concentration increased these regions grew and new denatured regions appeared. They related the unusually low melting temperatures of regions of the hybrids to the presence of mismatched DNA sequences and were able to construct a homology map between the two phages.

4. The Specific Fractionation of DNA with Restriction Endonucleases

The fractionation of DNA into unique fragments by incubation with restriction endonucleases provides specific DNA fragments of convenient sizes. The fragments are useful in mapping the DNAs from which they are made and in comparing DNAs from different sources. They are an important first step in the sequencing of large DNAs.

(a) Bacterial restriction enzymes

Meselson & Yuan (1968) discovered an ATP-dependent endonucleolytic activity in extracts of r_k bacteria. DNA from λ grown in m_k^- host cells was specifically degraded by this endonucleolytic activity. Klein & Sauerbier (1965) showed that a cleavage site could be rendered unrecognizable by methylation of a base within the cleavage site. Since DNA from λ grown in m_k cells was not degraded by the r_k extract the m_k system conferred an immunity to endonucleolytic degradation on the λ

DNA, probably by coding for a specific methylase.

Smith & Wilcox (1970) reported the isolation of a restriction endonuclease from Hemophilus influenzae and Kelly & Smith (1970) sequenced the hexanucleotide recognition sequence of the enzyme. The enzyme degraded T7 and P22 DNA to fragments with average lengths of 1000 and 1300 nucleotides respectively, but had no effect on H. influenzae DNA.

(b) Mapping with endo R

Danna & Nathans (1971) and Edgell et al. (1972) reported the fractionation of SV-40 and ϕX174 DNA, respectively, by cleavage with the restriction enzyme endo R (now known as Hind) from H. influenzae. The fragments could be separated by electrophoresis on polyacrylamide gels. Digestion of ϕX174 RF produced 13 fragments which electrophoresed in ten separable bands. The bands were named R1 through R10 in order of increasing mobility and decreasing size. The fragments were ordered onto a cleavage map by isolating partial digestion products and by digestion with a series of different restriction enzymes (Middleton et al., 1972; Lee & Sinsheimer, 1974a). Middleton et al. (1972) were able to orient the ϕX174 genetic map on the Hind cleavage map by infecting bacterial spheroplasts with defective single-stranded ϕX174 molecules containing an annealed R fragment. Whenever the defect in the single-stranded molecule occurred within the annealed R fragment the R fragment supplied the missing genetic information and the infection was productive.

Specific fragments isolated from a Hind digestion of ϕX174 RF contain unique regions of the ϕX174 genome. Lee & Sinsheimer (1974b) have developed a continuous electrophoresis apparatus that permits the recovery of R fragments in optical quantities. With the proper manipulation it should be possible to insert R fragments into molecules containing ϕX174 DNA. Since the fragments can be identified as to genetic function an inserted fragment would serve as an internal reference point in mapping a DNA molecule seen in the electron microscope.

Table 1

Comparison of the DNAs of ØX174 and S13

	ØX174		S13	
Single-stranded DNA contour length	1.77 ± 0.13 μm	7	1.90 ± 0.05 μm	1
Replicative Form contour length	1.64 ± 0.11 μm 1.89 ± 0.04 μm 1.77 ± 0.08 μm	2 3 5	1.83 ± 0.02 μm	1
Single-stranded DNA molecular weight	1.7×10^6	4	1.8×10^6	1
Replicative Form molecular weight	3.2×10^6 3.6×10^6	2 1	3.5×10^6	1
Single-stranded DNA buoyant density	1.72	4	1.726	1
Replicative Form buoyant density	1.709	6	1.710	1

References:
1. Spencer et al., 1972.
2. Kleinschmidt et al., 1963.
3. Chandler et al., 1964.
4. Sinsheimer, 1959b.
5. Roth & Hayashi, 1966.
6. Sinsheimer, 1968.
7. Freifelder et al., 1964.

Table 2

Molar Ratios of the DNA Bases

	A	G	T	C	Ref.
ØX174	1.00	0.98	1.33	0.75	4
S13	1.00	0.98	1.42	0.86	1

References: Same as in Table 1.

REFERENCES

Baker, R. & Tessman, I. (1967). Proc. Nat. Acad. Sci. USA 58, 1438-1445.

Benbow, R. M., Hutchison, C. A., Fabricant, J. D. & Sinsheimer, R. L. (1971). J. Virol. 7, 549-558.

Benbow, R. M., Mayol, R. F., Picchi, J. C. & Sinsheimer, R. L. (1972). J. Virol. 10, 99-114.

Benbow, R. M., Zuccarelli, A. J., Davis, G. C. & Sinsheimer, R. L. (1974). J. Virol. 13, 898-907.

Bonner, J., Kung, G. & Bekhor, I. (1967). Biochemistry 6, 3650-3653.

Britten, R. J. & Kohne, D. E. (1967). Carnegie Inst. Wash. Yearbook 65, 78-106.

Britten, R. J. & Kohne, D. E. (1968). Science 161, 529-540.

Brown, D. T., MacKenzie, J. M. & Boyer, M. E. (1971). J. Virol. 7, 836-846.

Burnet, F. M. (1927). Brit. J. Exp. Path. 8, 121-129.

Cerny, R., Cerna E. & Spencer, J. H. (1969). J. Mol. Biol. 46, 145-156.

Chandler, B., Hayashi, M., Hayashi, M. N. & Spiegelman, S. (1964). Science 143, 47-49.

Danna, K. J. & Nathans, D. (1971). Proc. Nat. Acad. Sci.USA 68, 2913-2917.

Darby, G., Dumas, L. B. & Sinsheimer, R. L. (1970). J. Mol. Biol. 52, 227-238.

Davis, R. W. & Davidson, N. (1968). Proc. Nat. Acad. Sci. USA 60, 243-250.

Davis, R. W. & Hyman, R. W. (1971). J. Mol. Biol. 62, 287-301.

Davis, R. W., Simon, M. N. & Davidson, N. (1971). In Methods in Enzymology, L. Grossman & K. Moldave, eds. Vol. XXI, part D, pp. 413-428, Academic Press, New York.

Denhardt, D. T. & Sinsheimer, R. L. (1965). J. Mol. Biol. 12, 647-662.

Doty, P., Marmur, J., Eigner, J. & Schildkraut, C. (1960). Proc. Nat. Acad. Sci. USA 46, 461-476.

Edgell, M. H., Hutchison, C. A., III, & Sclair, M. (1972). J. Virol. 9, 574-582.

Fiers, W. & Sinsheimer, R. L. (1962a). J. Mol. Biol. 5, 409-419.

Fiers, W. & Sinsheimer, R. L. (1962b). J. Mol. Biol. 5, 424-434.

Francke, B. & Ray, D. S. (1971). J. Mol. Biol. 61, 565-586.

Freifelder, D., Kleinschmidt, A. K. & Sinsheimer, R. L. (1964). Science 146, 254-255.

Godson, G. N. (1973). J. Mol. Biol. 77, 467-477.

Greenlee, L. L. & Sinsheimer, R. L. (1968). J. Mol. Biol. 32, 303-320.

Hall, C. E., Maclean, E. C. & Tessman, I. (1959). J. Mol. Biol. 1, 192-194.

Hayashi, Y. & Hayashi, M. (1970). Cold Spring Harbor Symp. Quant. Biol. 35, 171-177.

Helmkamp, B. K. & Ts'o, P. O. P. (1961). J. Amer. Chem. Soc. 83, 138-142.

Henry, T. J. & Knippers, R. (1974). Proc. Nat. Acad. Sci. USA 71, 1549-1553.

Hutchison, C. A. (1968). Ph.D. Thesis, California Institute of Technology.

Hutchison, C. A. & Sinsheimer, R. L. (1963). J. Mol. Biol. 7, 206-208.

Hutchison, C. A. & Sinsheimer, R. L. (1966). J. Mol. Biol. 18, 429-447.

Inman, R. B. & Schnos, M. (1970). J. Mol. Biol. 49, 93-98.

Jazwinski, S. M., Marco, R. & Kornberg. A. (1973). Proc. Nat. Acad. Sci. USA 70, 205-209.

Jeng, Y., Gelfand, M., Hayashi, M., Schleser, R. & Tessman, E. S. (1970). J. Mol. Biol. 49, 521-526.

Kelly, T. J., Jr. & Smith, H. O. (1970). J. Mol. Biol. 51, 393-409.

Kim, J.-S., Sharp, P. & Davidson, N. (1972). Proc. Nat. Acad. Sci. USA 69, 1948-1952.

Klein, A. & Sauerbier, W. (1965). Biochem. Biophys. Res. Commun. 18, 440-445.

Kleinschmidt, A. K., Burton, A. & Sinsheimer, R. L. (1963). Science 142, 961.

Kleinschmidt, A. K. & Zahn, R. K. (1959). Z. Naturforsch. 14b, 770.

Knippers, R., Salivar, W. O., Newbold, J. E. & Sinsheimer, R. L. (1969). J. Mol. Biol. 39, 641-654.

Lee, A. S. & Sinsheimer, R. L. (1974a). Proc. Nat. Acad. Sci. USA 71, 2882-2886.

Lee, A. S. & Sinsheimer, R. L. (1974b). Analyt. Biochem. 60, 640-644.

Lindqvist, R. H. & Sinsheimer, R. L. (1967). J. Mol. Biol. 30, 69-80.

Maclean, E. C. & Hall, C. E. (1962). J. Mol. Biol. 4, 173-178.

Marmur, J. & Doty, P. (1959). Nature 183, 1427-1429.

Marmur, J. & Lane, D. (1960). Proc. Nat. Acad. Sci. USA 46, 453-461.

Meselson, M. & Yuan, R. (1965). Nature 217, 1110-1114.

Middleton, J. H., Edgell, M. H. & Hutchison, C. A., III (1972). J. Virol. 10, 42-50.

Miller, L. K. & Sinsheimer, R. L. (1974). J. Virol., in press.

Newbold, J. E. & Sinsheimer, R. L. (1970). J. Mol. Biol. 49, 49-66.

Poljak, R. J. & Suruda, A. J. (1969). Virology 39, 145-147.

Razin, A. & Sedat, J. W. (1970). J. Mol. Biol. 53, 251-259.

Rice, S. A. & Doty, P. (1957). J. Amer. Chem. Soc. 79, 3937-3947.

Roth, T. F. & Hayashi, M. (1966). Science 154, 658-660.

Salivar, W. O. & Sinsheimer, R. L. (1969). J. Mol. Biol. 41, 39-65.

Schekman, R. W. & Ray, D. S. (1971). Nature New Biol. 231, 170-173.

Sertic, V. & Boulgakov, N. (1935). Compt. Rend. Soc. Biol. 119, 1270-1282.

Sinsheimer, R. L. (1959a). J. Mol. Biol. 1, 37-43.

Sinsheimer, R. L. (1959b). J. Mol. Biol. 1, 44-53.

Sinsheimer, R. L. (1968). Prog. Biophys. Nucleic Acid Res. 8, 115-169.

Smith, H. O. & Wilcox, K. W. (1970). J. Mol. Biol. 51, 379-391.

Spencer, J. H. & Boshkov, L. (1973). Can. J. Biochem. 51, 1206-1211.

Spencer, J. H., Cerny, R., Cerna, E. & Delaney, A. D. (1972). J. Virol. 10, 134-141.

Tessman, E. S. (1966). J. Mol. Biol. 17, 218-236.

Tessman, E. S. & Schleser, R. (1963). Virology 19, 239-240.

Tessman, I., Kumar, S. & Tessman, E. S. (1961). Science 158, 207-208.

Tibbets, C., Johansson, K. & Philipson, L. (1973). J. Virol. 12, 218-225.

Tromans, W. J. & Horne, R. W. (1961). Virology 15, 1-7.

Ts'o, P. O. P., Helmkamp, G. K. & Sander, C. (1962). Biochim. Biophys. Acta 55, 584-600.

Vanderbilt, A. S., Borras, M. T. & Tessman, E. S. (1971) Virology 43, 352-355.

Westmoreland, B. C., Szybalski, W. & Ris, H. (1969). Science 163, 1343-1348

Yarus, M. J. & Sinsheimer, R. L. (1967). J. Virol. 1, 135-144.

Zahler, S. A. (1958). J. Bacteriology 75, 310-315.

Chapter I

Specific ϕX174 DNA Fragments as Internal Markers for the Mapping of ϕX174 DNA Molecules in the Electron Microscope

1. Introduction

Skillful and imaginative manipulations during the preparation of DNA have greatly increased the range and subtlety of questions which can be answered with the electron microscope. Techniques which selectively denature regions of DNA rich in A-T base pairing (Inman & Schnös, 1970) or which denature partially homologous regions in heteroduplex molecules constructed from related DNA molecules (Davis et al., 1971; Davis & Hyman, 1971) have been used to study the structure and relationship of bacteriophage DNAs. However, a correlation between the physical map, as constructed in the electron microscope, and the genetic map of the virus has been possible only where the DNA is linear and has unique, distinguishable ends. There has been no way to correlate such maps for a circular molecule or for a linear one which is circularly permuted.

The introduction of restriction endonucleases has provided a means of cleaving DNA into reproducible, specific fragments which can be linearly ordered (Middleton et al., 1972; Danna et al., 1973). The restriction fragments produced from φX174 RF* by a series of endonucleases have been so ordered onto cleavage maps and the cleavage maps have been superimposed on the genetic map (Lee and Sinsheimer, 1974a). DNA fragments whose positions on the genetic map are known are thus available for hybridization reactions.

---

*Abbreviations used: RF, double-stranded circular replicative form DNA of φX174; RF I, supercoiled RF DNA with both strands covalently closed; RF II, RF DNA with one or more single-stranded nicks in either strand.

This study describes the use of such specific ϕX174 DNA restriction fragments as internal markers for molecules seen in the electron microscope. Using these fragments the genetic map can be superimposed on molecules in electron micrographs and specific regions of the observed molecules can be identified with their corresponding genetic function.

2. Materials and Methods

(a) Bacterial and bacteriophage strains

(i) Escherichia coli C and E. coli HF4714 (amber suppressor, rec-13$^+$) were used as host strains in all experiments.

(ii) ϕX174am3, lysis defective, was used as the source of all DNA.

(b) Media and solutions

(i) Bacteria were grown in tryptone broth supplemented with 1 mM-$CaCl_2$.

(ii) Tris/EDTA is 0.05 M-Tris, 0.005 M-EDTA, pH 8.5. 2X Tris/EDTA is 0.1 M-Tris, 0.01 M-EDTA, pH 8.5.

(c) Preparation of ϕX174

Phage stocks were prepared from single plaques. Pure, high titer stocks were prepared by banding phage in CsCl in a Spinco type 50 rotor at 30,000 revs/min for 48 h at 5°C. The wild type reversion frequency was less than 10^{-5}.

(d) Preparation of ϕX174 viral DNA

ϕX174 was extracted four times with phenol saturated with Tris/EDTA and adjusted to pH 8.5. The DNA was precipitated from the aqueous phase by adding 0.1 vol 3 M-sodium acetate (pH 5.5), 2 vol isopropanol and storing at -20°C overnight. The DNA was collected by centrifugation, resuspended in Tris/EDTA and dialyzed briefly. The ratio A_{260}/A_{280} was greater than 1.8, and the proportion of circular molecules was greater than 95% as determined by electron microscopy. The DNA was stored frozen until its use.

(e) Preparation of RF I

ϕX174 RF I was prepared as described by Johnson & Sinsheimer (1974). The RF was greater than 95% RF I as determined by electron microscopy.

(f) Preparation of ϕX174 complementary strand DNA

Complementary strand DNA was prepared from RF I by the method of Baas & Jansz (1972). 200 µg RF I in 0.27 ml 2X-Tris/EDTA was exposed to 800 roentgens of X-rays. 0.15 ml poly U,G (Miles Laboratories) at 1 mg/ml which had been partially degraded by heating to 80°C for 5 min in Tris/EDTA, and 8.0 ml 2X-Tris/EDTA were added to the DNA, and the solution was separated into two tubes. The DNA mixture was heated in water at 100°C for 3 min and was quickly cooled by swirling in an ice bath for 20 sec. 8.35 gm CsCl (dessicated) and 2 ml cold 2X-Tris/EDTA were added to each tube to complete the cooling. The densities of the solutions were adjusted to 1.745 as necessary and the solutions were put in polyallomer centrifuge tubes. 30 µl 30% (w/v) sarkosyl (Geigy) was

added to each tube to prevent DNA from adhering to the walls, and the solution was centrifuged to equilibrium in a Spinco type 50 rotor at 40,000 revs/min for 50 h at 5°C. 15-drop fractions were collected starting from the bottom of the tube through a 10 µl micropipette inserted from above and the A_{260} measured for each fraction. Four absorbing regions were observed. A region of high absorption at the bottom of the tube was assumed to be poly U,G and a peak near the top at a density of 1.700 was assumed to be a mixture of RF I and RF II. Two intermediate peaks at densities of 1.76-1.77 and 1.725 were assumed to be complementary strand-poly U,G hybrid and free viral strand, respectively. The pooled fractions containing the complementary strand DNA-poly U,G hybrid (from both centrifuge tubes) were adjusted to 0.2 M-NaOH, 0.01 M-EDTA and incubated at 37°C for 1 h to degrade the poly U,G. The NaOH was removed by dialysis against Tris/EDTA and the DNA precipitated by adding 0.1 vol 3 M-sodium acetate (pH 5.5), 2 vol isopropanol and storing at -20°C overnight. The DNA was collected by centrifugation, resuspended in 2X-Tris/EDTA and dialyzed briefly. The total yield was about 10 µg. The DNA was stored frozen until used.

(g) Source of restriction fragments

The restriction enzyme fragments were isolated from a digestion of φX174 RF I by the restriction endonuclease Hind from *Hemophilus influenzae* and were the generous gift of L. H. Smith.

(h) Preparation of hybrids

0.25 µg φX174 viral strand and 0.25 µg φX174 complementary strand were combined with 0.25 µg of each restriction fragment (when

used) in a volume of 0.1 ml. The mixture was dialyzed against 90% formamide (Matheson, Coleman, Bell, 99%), 0.01 M-EDTA at 22° for 2 h to denature the double-stranded fragments and against 0.2 M-Tris, 0.02 M-EDTA, pH 8.5, 60% formamide at 22°C for 1 h to renature the DNA. Formamide was removed by dialyzing overnight against 2X-Tris/EDTA at 4°C. Renatured mixtures were stored frozen until use.

(i) Mounting of DNA for electron microscopy

DNA was prepared for electron microscopy by the method of Davis et al. (1971) as standardized by Davis & Hyman (1971). DNA at a concentration of 1.0 µg/ml in 2X-Tris/EDTA, 50% formamide and 40 µg/ml cytochrome C was layered onto a hypophase of 20% formamide in 0.01 M-Tris, 0.001 M-EDTA, pH 8.5 by allowing it to run down a microscope slide inserted into the hypophase. The DNA/cytochrome film was picked up onto parlodion coated grids and was stained with uranyl-acetate and shadowed with platinum-palladium.

(j) Electron microscopy and molecule measuring

Molecules were photographed in a Phillips EM 300 electron microscope. The molecules were projected onto the measuring surface of a Hewlett-Packard Model 9864A Digitizer and the different regions of the molecules traced and measured. For molecules containing loops the proportion of the genome involved in a loop was expressed as the ratio of the longest arm of the loop to the total molecule length, including the longest arm of each loop.

3. Results

(a) Renaturation of ϕX174 viral and complementary DNA

The complementary strand preparation contained 8.6% intact circular molecules, 47.0% full length linear molecules and 44.4% linear fragments. After extensive self-renaturation fewer than 1% of the full length molecules were double-stranded and it was concluded that the complementary strand preparation was nearly free of viral DNA molecules.

Renaturation of complementary and circular viral DNA greatly increased the number of double-stranded molecules. The proportion of double-stranded molecules increased with time of renaturation, but renaturations longer than one to two hours were not useful because of the large number of multiple renaturation forms. An estimated 60% of the molecules from a five hour renaturation were associated with large, interwoven clumps of DNA. A one hour renaturation time was adopted as a compromise between a lower yield of double-stranded molecules and excessive aggregation.

Viral and complementary strand DNA renatured together under standard conditions contained a mixture of single and double-stranded circles, single-stranded linears and fragments, and two infrequent circular forms. One of the circular forms looked like two partially interwoven circles and appeared in a wide range of contortions and degrees of double-strandedness. These were classified as abortive attempts to renature by a circular viral strand and a circular complementary strand. Circular molecules of the second type contained a loop, such as seen in replicating intermediates (Kasamatsu *et al.*,

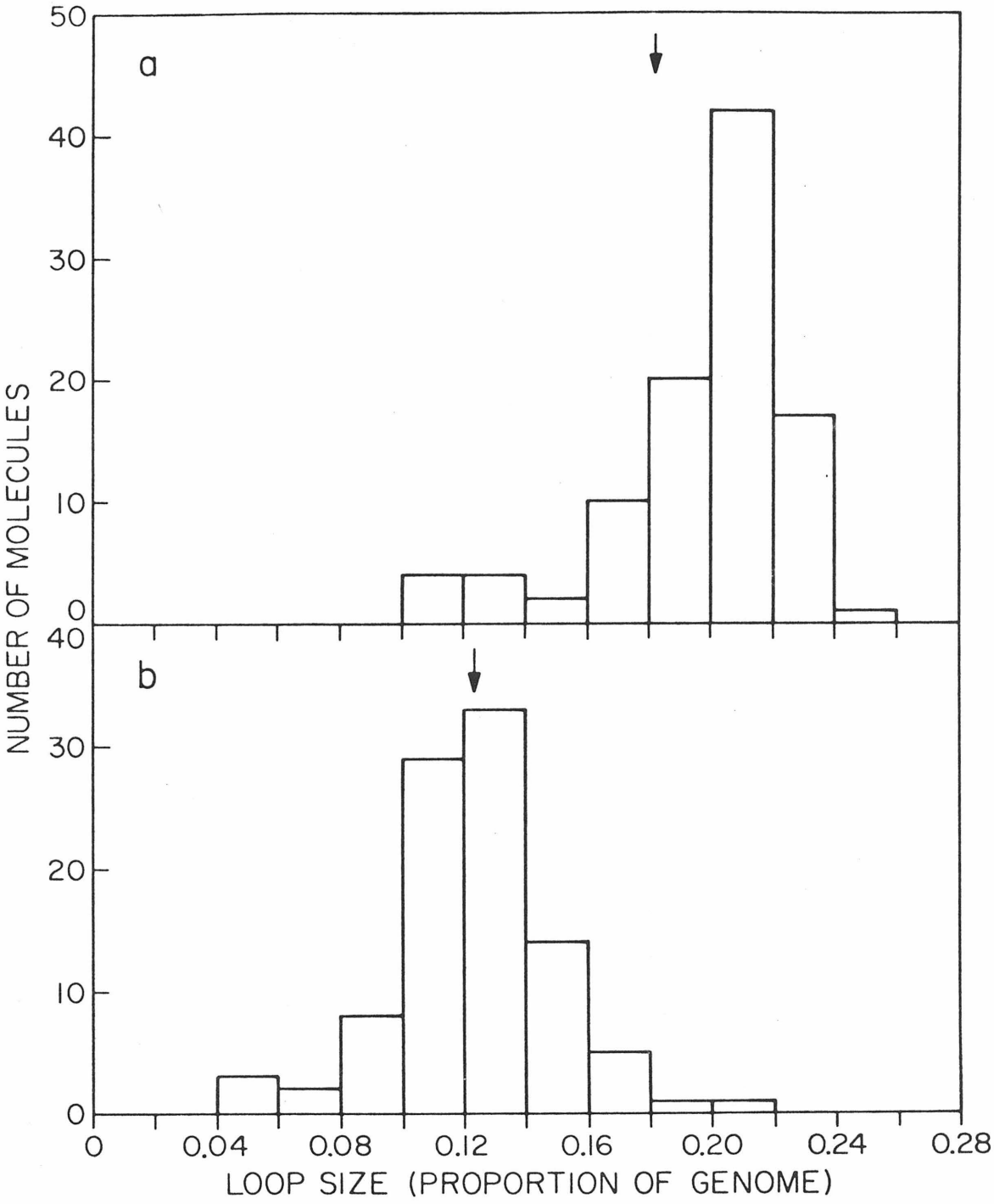
1971), and were considered to contain an inserted DNA fragment. The molecules containing an insert loop were seen at a frequency of 0.3 to 0.5% of the total circular molecules. The size of the insert loop varied widely, reflecting the variety in the size of the DNA fragments in the complementary strand preparation.

(b) The insertion of specific DNA fragments into ϕX174 RF II made in vitro

Hemophilus influenzae restriction endonuclease Hind cleaves ϕX174 RF into 13 specific fragments which electrophorese in ten separable bands (Edgell et al., 1972), called R1 to R10 in order of increasing mobility and decreasing size. The fragments have been oriented onto a cleavage map of the genome by analyses of products of partial digestions and by sequential use of two or three restriction enzymes (Middleton et al., 1972; Lee and Sinsheimer, 1974a), and the cleavage map has been oriented on the ϕX174 genetic map by infecting bacterial spheroplasts with defective single strand molecules carrying an annealed wild type restriction fragment (Middleton et al., 1972). The fragments can be purified in optical quantities by continuous gel electrophoresis (Lee & Sinsheimer, 1974b) and are thus ideally suited for investigations requiring specific DNA fragments.

When either fragment R1 or R3 was included in a renaturation of ϕX174 viral and complementary DNA the proportion of circular molecules containing an insert loop increased five to ten fold. Figure 1 shows histograms of loop size distributions for two separate renaturations including R1 and R3, respectively. Each distribution has a

FIG. 1. Histograms of the loop sizes seen in renaturations including ϕX174 viral and complementary DNA and a specific restriction fragment from a Hind digestion of ϕX174 RF: (a) fragment R1, (b) fragment R3. In each case the arrow indicates the expected loop size based on the estimates of fragment size in Lee & Sinsheimer (1974a). The loops at values away from the maxima are attributed to randomly sized DNA fragments contained in the complementary strand preparation.


a
b
NUMBER OF MOLECULES
LOOP SIZE (PROPORTION OF GENOME)
0
10
20
30
40
50
0.04
0.08
0.12
0.16
0.20
0.24
0.28

distinct maximum reflecting the size of the fragment used and only a background level of randomly sized loops in the size range in which loops formed by the other fragment would appear. In order to confirm the independence of the two distinct loop sizes R1 and R3 were included in the same renaturation. Figure 2 shows that the loop size distribution for this two fragment renaturation has two maxima, one corresponding to each of the maxima in Figure 1. The addition of a specific DNA fragment to a renaturation mixture causes the appearance of loops that correspond in size to that fragment, independent of the presence of a second specific fragment. We conclude that a single-stranded DNA fragment binds to one of the full length strands, and that this hybrid then renatures with its full length counterpart (which may also carry a fragment). The presence of the fragment displaces the opposite strand and generates the loop structure seen in the electron microscope.

The independence and reproducibility of the loops suggest that their specific origin is the R fragment included in the renaturation; however, the loop attributed to R1 is larger than would be expected on the basis of current estimates of the size of R1. A renaturation including fragments R1 and R4 was prepared in order to test the reproducibility of the R1 loop and to see what size loop a smaller fragment would produce. Figure 3 is a histogram of the loop size distribution for this renaturation. There are two maxima, one corresponding to the R1 peak in Figures 1 and 2, and a new peak at a smaller loop size. The two loops attributed to R3 and R4 are the sizes expected for these fragments, but the R1 loop is larger than would be expected for the R1 fragment by 150 nucleotides. During the last two years the estimated

FIG. 2. Histogram of loop sizes seen in a renaturation including both fragments R1 and R3 in addition to viral and complementary DNA. The two peaks are in the same positions as the two individual peaks in Figure 1 and indicate that the specific sizes of the loops are not altered by the presence of a second fragment. The arrows indicate the expected sizes as indicated in Figure 1.

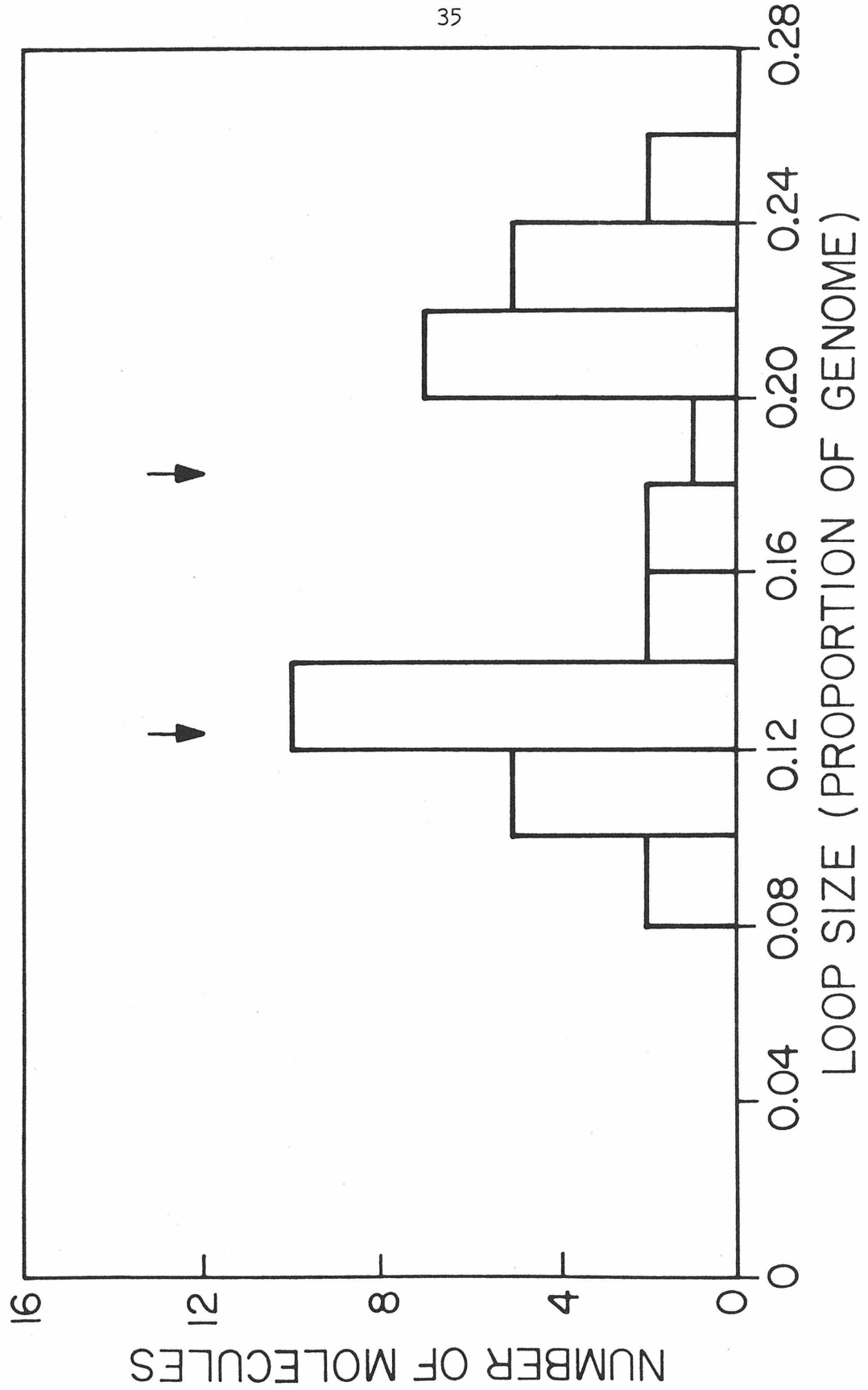
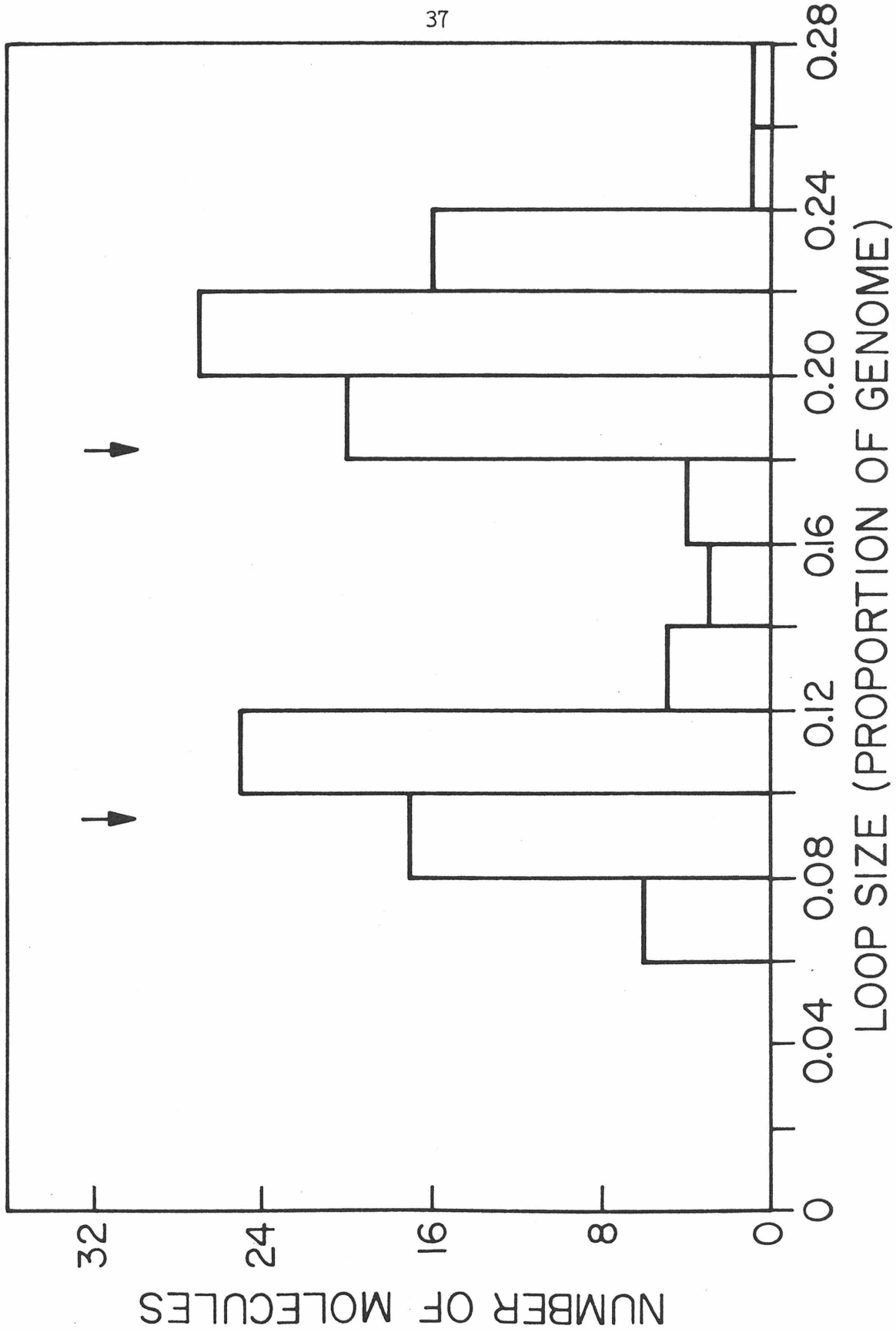

NUMBER OF MOLECULES
16
12
8
4
0
0
0.04
0.08
0.12
0.16
0.20
0.24
0.28
LOOP SIZE (PROPORTION OF GENOME)

FIG. 3. Histogram of loop sizes seen in a renaturation including fragments R1 and R4. The arrows indicate the expected sizes as indicated in Figure 1. The peak representing the larger loops is at the same position as the large loop peaks in Figures 1 and 2 while the small loop peak is at a smaller value than the R3 peak in Figures 1 and 2. The three restriction fragments used each cause the formation of a class of molecules containing an insert loop of a specific, distinguishable size.


NUMBER OF MOLECULES
32
24
16
8
0
0
0.04
0.08
0.12
0.16
0.20
0.24
0.28
LOOP SIZE (PROPORTION OF GENOME)

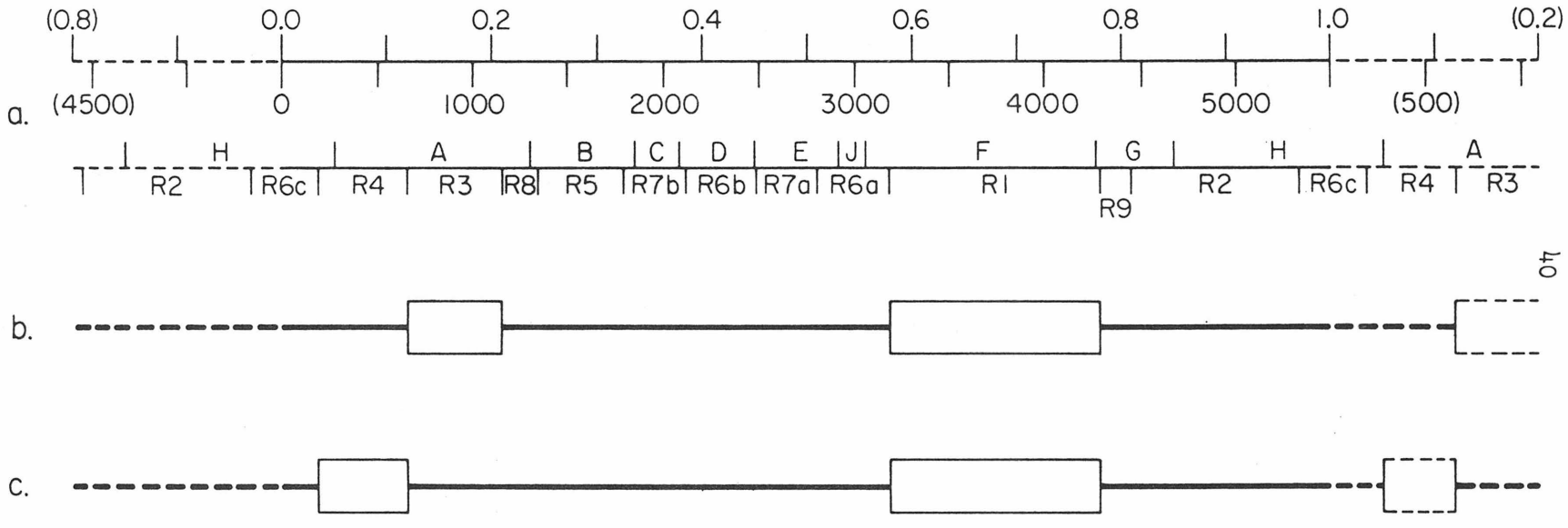
size of R1 has been reduced from 1400 base pairs (Edgell et al., 1972) to 1000 base pairs (Lee and Sinsheimer, 1974c). Unless the presence of the R1 fragment causes the reproducible melting of an adjacent duplex region it is likely that 1000 base pairs is an underestimate of its size.

(c) The use of restriction fragments to orient molecules to an external map

Before the insertion of restriction fragments can be used as a mapping technique the binding sites must be shown to be unique. Figure 4a is the Hind cleavage map of φX174 superimposed on the φX174 genetic map (adapted from Lee & Sinsheimer, 1974c). Fragments R1 and R3 were initially chosen because they were large fragments, maximally separated on the cleavage map. Figure 4b is a molecular profile of the loops expected in a molecule containing both an R1 and an R3 fragment. Twenty-seven loop-containing molecules were seen in the renaturation including both R1 and R3 fragments. Four of these molecules contained two loops, one corresponding in size to each fragment and in each case the two loops were on opposite sides of the molecule, as expected from the cleavage map. Molecules containing R1 and R3 loops are shown in Plates I and IIa, and a molecule containing both R1 and R3 loops is shown in Plate II

The Hind fragments therefore produce insert loops reflecting their size and bind to specific sites on the genome. The proper choice of fragments should allow a molecule containing two loops to be superimposed on the cleavage map and on the genetic map. The molecule profile in Figure 4c shows that a molecule containing R1 and R4 insert loops would have two distinguishable loops arranged asymmetrically on

FIG. 4. (a) A Hind cleavage map of ϕX174 RF and the ϕX174 genetic map matched to a linear scale of 5500 nucleotides. Adapted from Lee & Sinsheimer (1974c). (b) The expected profile for a molecule containing fragments R1 and R3. (c) The expected profile for a molecule containing fragment R1 and R4.


(0.8)
0.0
0.2
0.4
0.6
0.8
1.0
(0.2)
a.
(4500)
0
1000
2000
3000
4000
5000
(500)
H
A
B
C
D
E
J
F
G
H
A
R2
R6c
R4
R3
R8
R5
R7b
R6b
R7a
R6a
R1
R9
R2
R6c
R4
R3
b.
c.

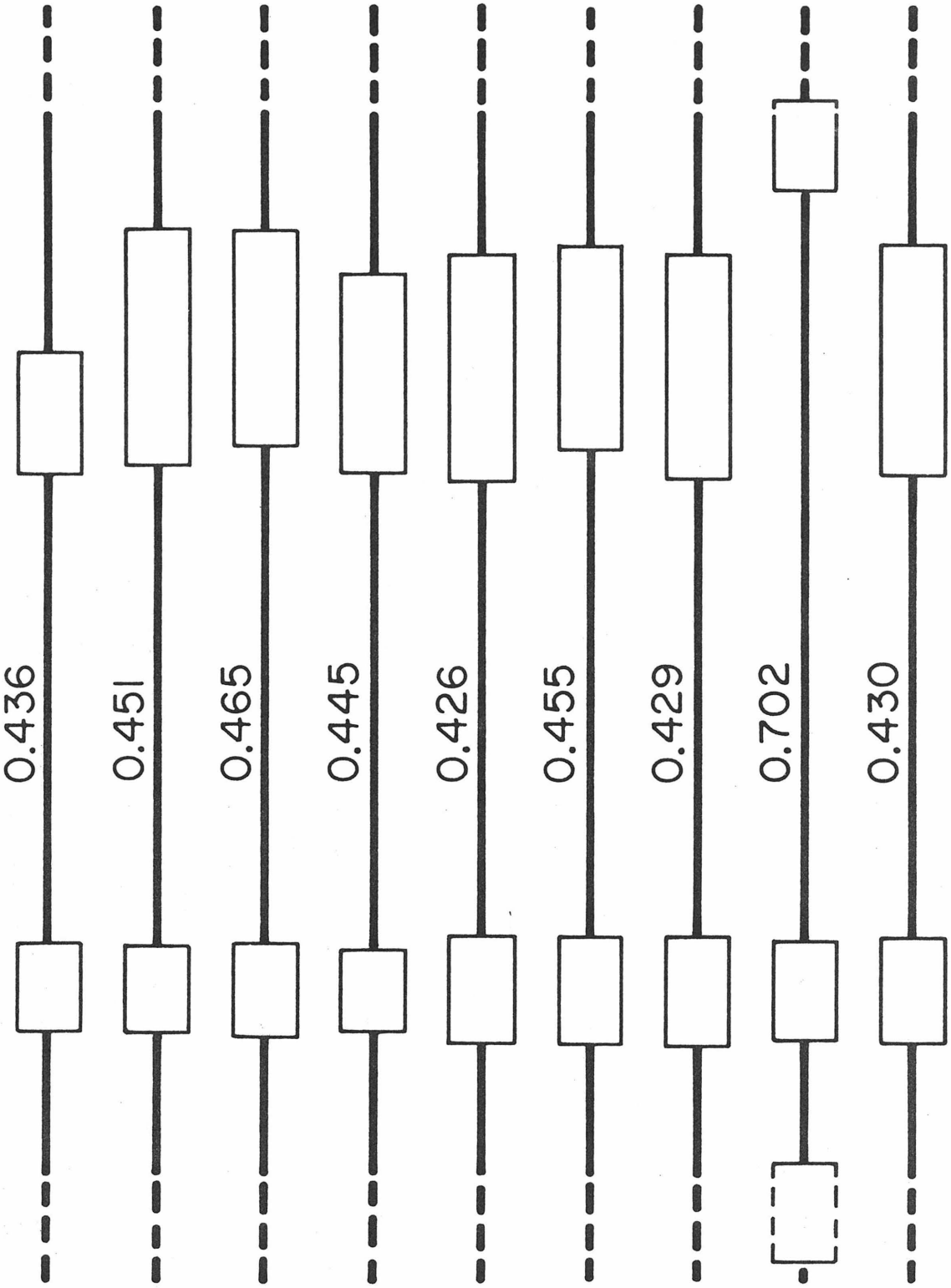
the genome. Nine of the 120 loop-containing molecules seen in the renaturation including both R1 and R4 had two insert loops. Their profiles are shown in Figure 5. Seven of these molecules are of the type predicted in Figure 4c and are considered to have insert loops formed by an R1 and an R4 fragment. Plate IIb shows a molecule containing an R4 loop and Plate IIIb shows a molecule containing both R1 and R4 loops. The molecules with both R1 and R4 insert loops fulfill the necessary requirements and thus can be superimposed on the genetic map. Plate IV is a micrograph of a molecule containing R1 and R4 loops and a line drawing of the molecule with the ϕX174 genetic map superimposed.

4. Discussion

The specificity of the loop size in molecules containing an insert loop and the reproducible relationship of the loops in molecules containing two loops indicate that specific DNA fragments can be inserted into ϕX174 RF II molecules, made *in vitro*, at specific sites determined by their base sequence. The most obvious use of this technique is the mapping of restriction fragments relative to each other. The process of comparing two-loop molecules to an existing cleavage map, as just described, could be reversed to construct a cleavage map from the insert loops. Although the larger fragments have been used in this study the technique is applicable to smaller fragments. The accompanying paper (Compton & Sinsheimer, 1975) describes the use of R8 which is estimated to be 180 base pairs.

In combination with the technique of Griffith *et al.* (1971) the insert loops could be used to map the location of specific single

FIG. 5. Molecular profiles for the two-loop molecules from the renaturation with fragments R1 and R4. The centers of the loop closest in size to R4 in each molecule have been aligned. Seven of the molecules exhibit the profile expected for R1 and R4 loops. One contains two loops correctly oriented but with a smaller R1 loop, and one contains a loop corresponding to R4 and a second apparently random loop. The number indicates the length of the double-stranded region separating the two loops expressed as proportion of the entire molecule.


0.436
0.451
0.465
0.445
0.426
0.455
0.429
0.702
0.430

strand nicks or specific gaps by binding DNA polymerase I to the hybrid. The same technique will be useful as methods are developed to stabilize other enzymes at their recognition sites on DNA, especially if they can be labeled with an electron-opaque marker such as ferritin (Wu & Davidson, 1971). The technique has already been used to map the DNA regions conserved between two independently isolated bacteriophages (Compton & Sinsheimer, 1975).

It will be necessary to increase the yield of molecules containing two insert loops in order to make this a practical mapping procedure. The low level of insert loops in spite of the 6 to 10 fold excess of the restriction fragment is attributed to the displacement of the fragment by the opposite strand. The proportion of circles containing loops seems to be a function of both their formation and decay with increasing time of renaturation. The formation of a covalently bonded hybrid between fragments and one full length strand before the addition of the opposite strand would prevent the displacement of the fragment. Psoralen has been described as a drug which intercalates between adjacent bases in duplex DNA and forms covalent cross linkages when exposed to ultraviolet light (Cole, 1971). It may be possible to insert psoralen into duplex DNA under conditions which minimize base stacking in single-stranded DNA. A restriction fragment could then be covalently bonded to its full length complementary strand with a minimal risk of any damage to the single-stranded DNA which would prevent its reannealing and formation of the double fragment hybrid.

REFERENCES

Baas, P. D. & Jansz, H. S. (1972). J. Mol. Biol. 63, 557-568.

Cole, R. S. (1971). Biochim. Biophys. Acta 254, 30-39.

Compton, J. L. and Sinsheimer, R. L. (1975). J. Mol. Biol.

Danna, K. J., Sack, G. H., Jr. & Nathans, D. (1973). J. Mol. Biol. 78, 363-376.

Davis, R. W. & Hyman, R. W. (1971). J. Mol. Biol. 62, 287-301.

Davis, R. W., Simon, M. N. & Davidson, N. (1971). In Methods in Enzymology (Grossman, L. & Moldave, K., eds.), volume XXI, part D, pp. 413-428, Academic Press, New York.

Edgell, M. H., Hutchison, C. A., III & Sclair, M. (1972). J. Virol. 9, 574-582.

Griffith, J., Huberman, J. A. & Kornberg, A. (1971). J. Mol. Biol. 55, 209-214.

Inman, R. B. & Schnös, M. (1970). J. Mol. Biol. 49, 93-98.

Johnson, P. H. & Sinsheimer, R. L. (1974). J. Mol. Biol. 83, 47-61.

Kasamatsu, H., Robberson, D. L. & Vinograd, J. (1971). Proc. Nat. Acad. Sci. USA 68, 2252-2257.

Lee, A. S. (1974). Ph.D. Thesis, California Institute of Technology.

Lee, A. S. & Sinsheimer, R. L. (1974a). Proc. Nat. Acad. Sci. USA 71, 2882-2886.

Lee, A. S. & Sinsheimer, R. L. (1974b). Analyt. Biochem. 60, 640-644.

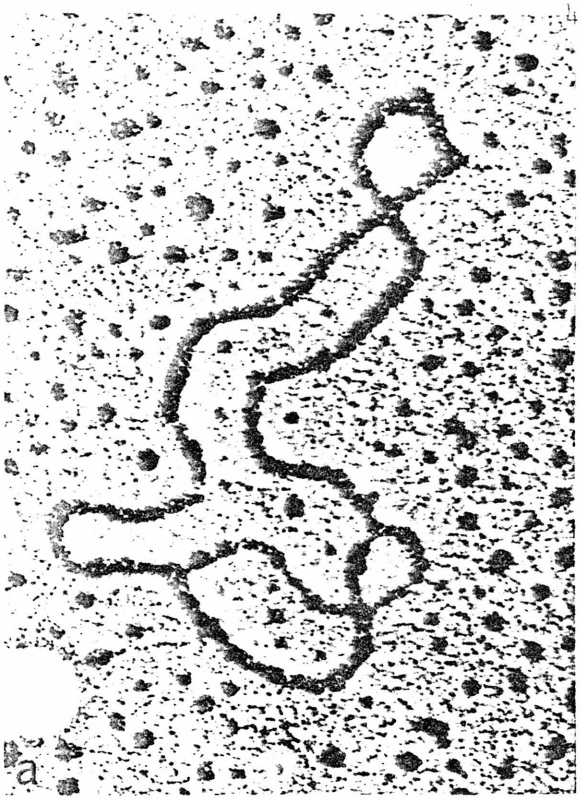
Lee, A. S. & Sinsheimer, R. L. (1974c). J. Virol., In Press.

Middleton, J. H., Edgell, M. H. & Hutchison, C. A., III (1972). J. Virol. 10, 42-50.

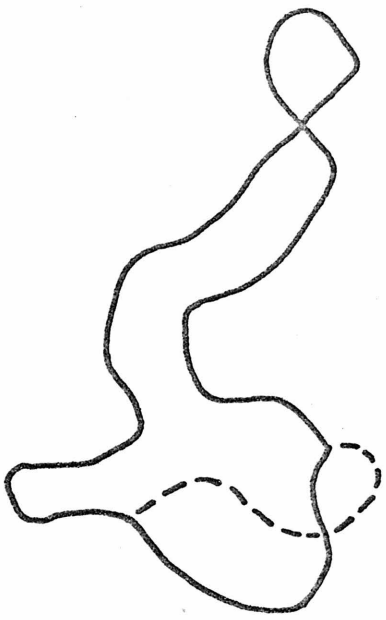
Sinsheimer, R. L. (1959). J. Mol. Biol. 1, 43-53.

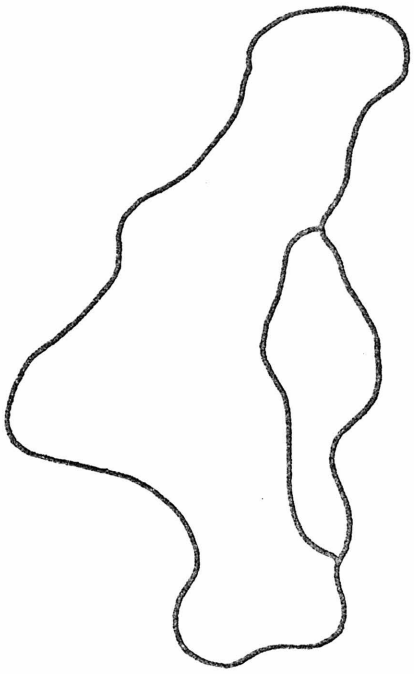
Wu, M. & Davidson, N. (1971). Fed. Proc. 30, 1054.

PLATE I: ØX174 RF II molecules containing R1 loops. (a) An RF II molecule containing one R1 fragment strand annealed to its complementary full length strand. The displaced single-stranded segment is indicated by the dashed line in the interpretive drawing. (b) An RF II molecule which contains an R1 fragment strand annealed to each of the full length strands.



a





b

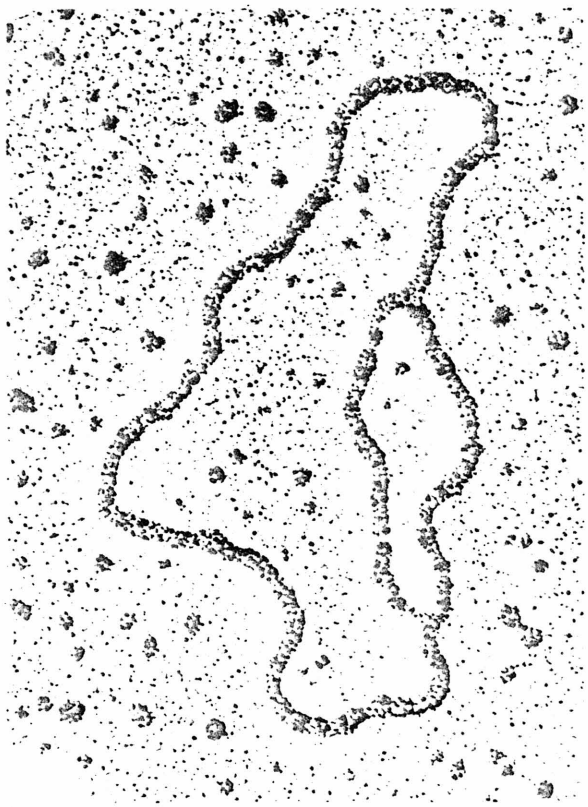
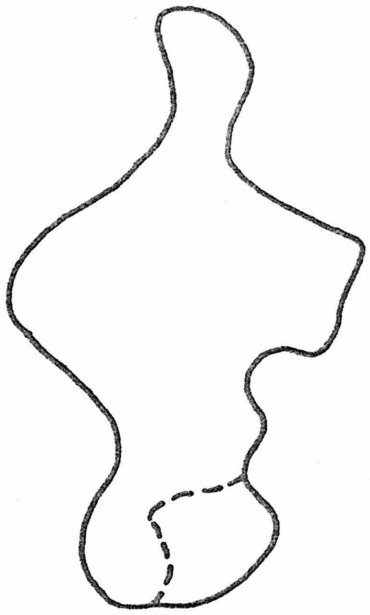


PLATE II: ϕX174 RF II molecules containing insert loops formed by (a) R3 and (b) R4. The probable single-stranded regions are indicated by dashed lines. The fragment in the center of the molecule in (b) is a double-stranded R4 fragment.


50
a



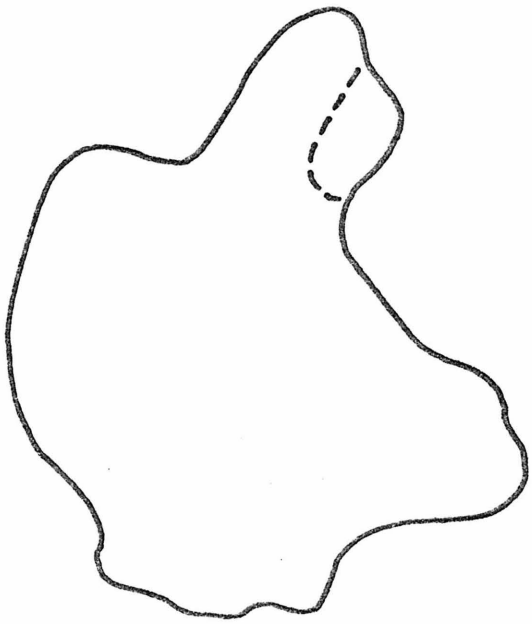
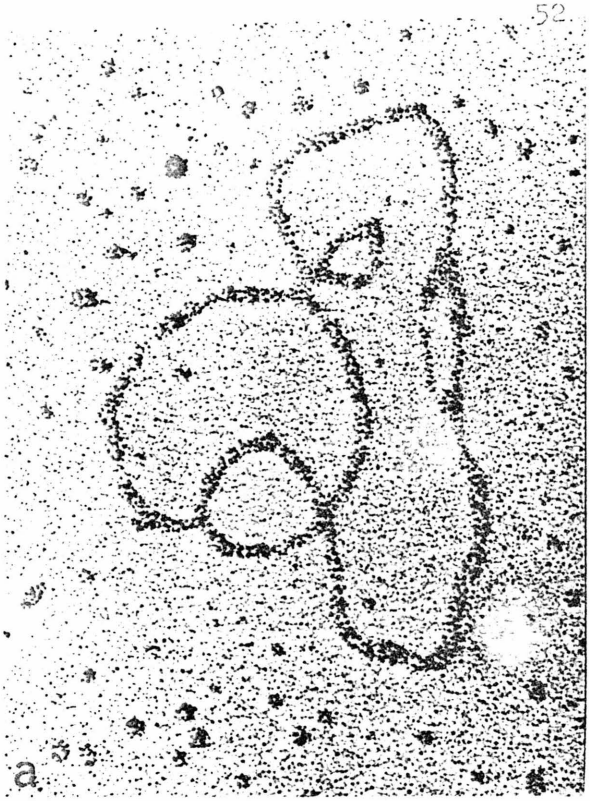
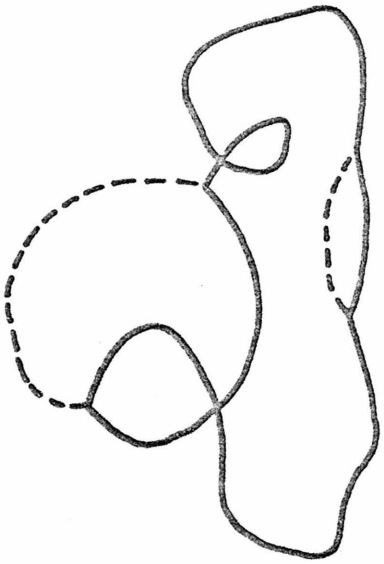
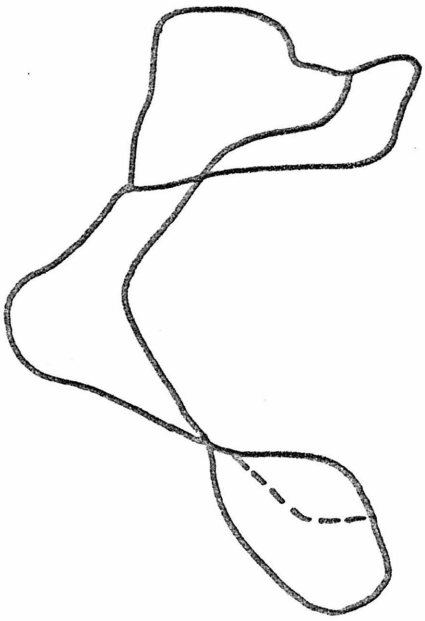

b

PLATE III: ØX174 RF II molecules containing two insert loops. (a) A molecule containing loops formed by R1 (left) and R3 (right). (b) A molecule containing loops formed by R1 (top) and R4 (bottom). Dashed lines indicate probable single-stranded regions.



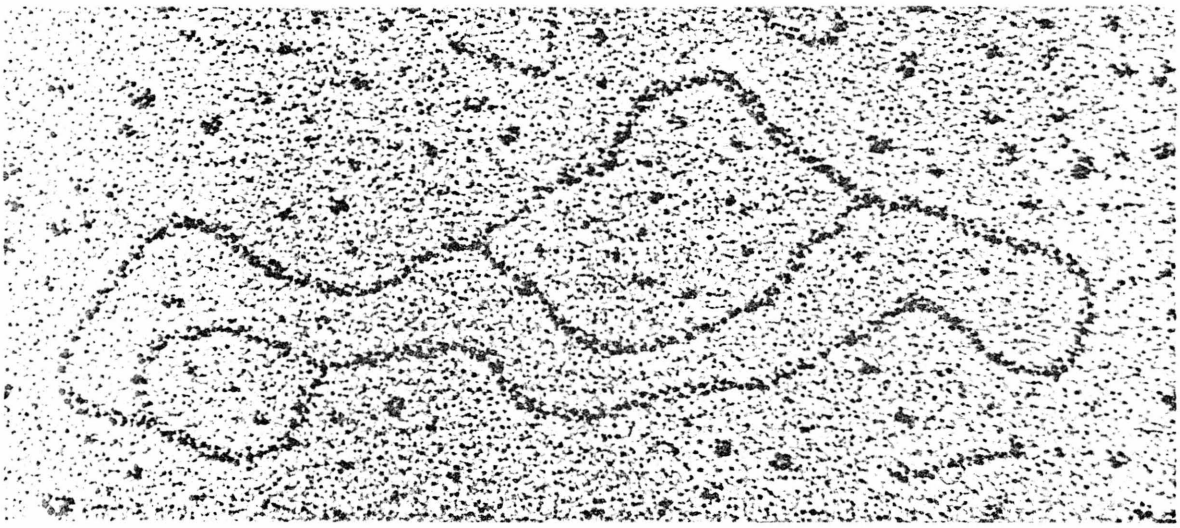
a

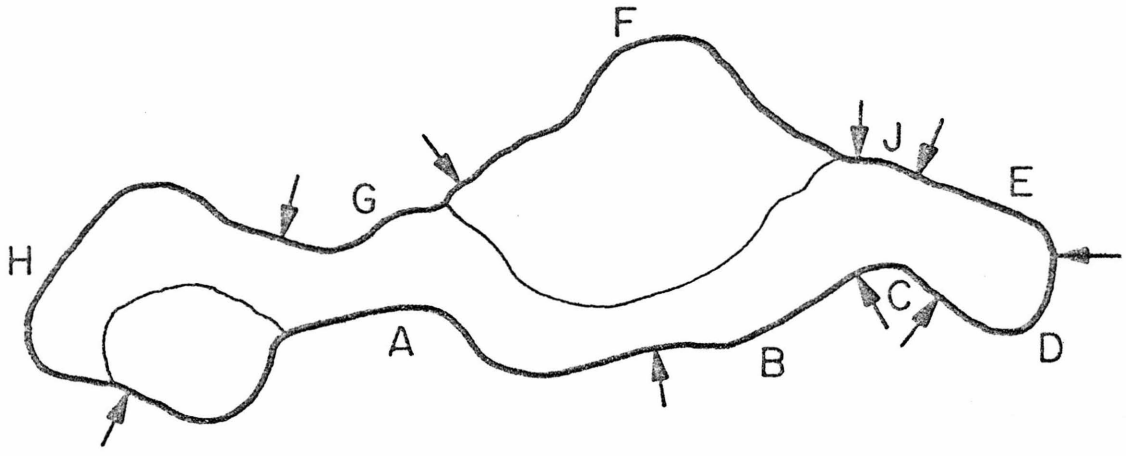




b

PLATE IV: A micrograph of a ϕX174 RF II molecule containing insert loops formed by fragments R1 and R4. The ϕX174 genetic map has been superimposed on the drawing and the boundaries between adjacent genes are denoted by arrows. The probable single-stranded regions are denoted by thin lines.




F
J
E
G
H
C
A
B
D

Chapter II

Aligning the ϕX174 Genetic Map and the ϕX174/S13 Heteroduplex Denaturation Map

1. Introduction

The conservation of nucleotide sequence between two related DNAs can be examined by preparing heteroduplex molecules containing one DNA strand from each source. When prepared for electron microscopy, regions of the molecules in which the base pairing is conserved appear as double-stranded segments while regions which have been destabilized by the preparation conditions appear as single-stranded loops or tails. The most common technique uses formamide both to stabilize single-stranded DNA in an extended form and to provide a range of conditions which increasingly destabilizes partially homologous duplex regions. The formamide technique has been used to examine the relationships between the lambdoid coliphages (Westmoreland et al., 1969; Davis et al., 1971; Simon et al., 1971) and the coliphage pairs T3 and T7 (Davis & Hyman, 1971) and T2 and T4 (Beckendorf et al., 1973). In these studies the genetic maps of the DNAs could be superimposed on the heteroduplex denaturation maps and the significance of the physical differences described in terms of the genetic maps. The ability to superimpose the genetic map on the heteroduplex maps, however, depended on the use of linear DNAs with distinguishable ends and the availability of specific deletions whose positions on the genetic maps were known.

Godson (1973) reported that the small single-stranded DNA containing coliphages ϕX174 and S13 contained only one small region which remained double-stranded when heteroduplexes constructed from

the two RFs* were mounted under conditions close to the melting temperature of the homoduplexes. This evidence for such extensive sequence divergence in spite of immunological, physical and genetic similarities (Sinsheimer, 1968) could be used together with the well-characterized genetic maps of the two phages (Benbow et al., 1974; Baker & Tessman, 1967) to investigate the tolerance of a genetic system for DNA sequence divergence. Since the phage DNAs are circular and without distinguishing characteristics, however, Godson was unable to align the denaturation patterns seen under different conditions or to superimpose the genetic map on his heteroduplexes.

In the accompanying paper Compton & Sinsheimer (1975) report that specific DNA fragments purified from the digestion of ϕX174 RF by restriction endonucleases can be used to provide internal reference points for otherwise featureless, circular DNA molecules. Some of the ϕX174 RF II molecules made in vitro in the presence of two restriction fragments contain two insert loops caused by the displacement of the opposite strand by the annealed DNA fragments. The two loops are positioned with respect to each other as the fragments are on the ϕX174 cleavage map and can be used as internal reference points to superimpose the ϕX174 genetic map on the molecule.

This report describes the application of this technique to superimpose the ϕX174 genetic map on a denaturation map constructed for the ϕX174/S13 heteroduplex.

---

*Abbreviations used: RF, double-stranded circular replicative form DNA; RF I, supercoiled RF DNA with both strands covalently closed; RF II, RF DNA with one or more single-stranded nicks in either strand.

2. Materials and Methods

(a) Bacterial and bacteriophage strains

(i) *Escherichia coli* C and *E. coli* HF4714 (amber suppressor, rec-13$^+$) were used as host strains in all experiments.

(ii) ϕX174*am*3, lysis defective, and S13 wild type were used as the sources of DNA.

(b) Media and solutions

Bacteria were grown in tryptone broth supplemented with 1 mM-$CaCl_2$.

(c) Preparation of ϕX174 and S13

Phage stocks were prepared from single plaques. Pure, high titer stocks were prepared by centrifuging to equilibrium in CsCl in a Spinco type 50 rotor at 30,000 revs/min for 48 h at 5°C. The ϕX174 *am*3 wild type reversion frequency was less than 10^{-5}.

(d) Preparation of S13 viral DNA

S13 viral DNA was prepared as described for ϕX174 viral DNA by Compton & Sinsheimer (1975). The ratio A_{260}/A_{280} was greater than 1.8, and the proportion of circular molecules was greater than 95% as determined by electron microscopy. The DNA was stored frozen until its use.

(e) Preparation of ϕX174 RF I

ϕX174 RF I was prepared as described by Johnson & Sinsheimer (1974). The RF was greater than 95% RF I as determined by electron

microscopy. The RF was stored frozen until its use.

(f) Preparation of ϕX174 complementary strand DNA

ϕX174 complementary strand DNA was prepared by the method of Baas & Jansz (1972) as described by Compton & Sinsheimer (1975).

(g) Source of restriction fragments

The restriction enzyme fragments were isolated from a digestion of ϕX174 RF I by the restriction endonuclease Hind from Hemophilus influenzae and were the generous gift of L. H. Smith.

(h) Preparation of heteroduplexes

0.25 µg ϕX174 complementary strand DNA and 0.25 µg S13 viral DNA were combined with 0.25 µg of a restriction fragment (when used) in a volume of 0.1 ml. The mixture was dialyzed against 90% formamide (Matheson, Coleman, Bell, 99%), 0.01 M-EDTA at 22°C for 2 h to denature the double-stranded fragment, and against 100 ml 60% formamide, 0.2 M-Tris, 0.02 M-EDTA, pH 8.5, at 22°C for 1 h to facilitate the renaturation of the ϕX174 complementary strand and the viral strand of the restriction fragment. 300 ml 0.2 M-Tris, 0.02 M-EDTA, pH 8.5, was added to the dialysate at a rate of 100 ml/h to provide a temporal gradient of formamide concentration to promote the annealing of the ϕX174 and S13 DNAs. Formamide was removed from the sample by dialyzing overnight against 0.1 M-Tris, 0.01 M-EDTA, pH 8.5 at 4°C. Renatured mixtures were stored frozen until their use.

(i) Mounting of DNA for electron microscopy

DNA was prepared for electron microscopy by the method of Davis et al. (1971) as standardized by Davis & Hyman (1971). DNA at a concentration of 1.0 µg/ml in 0.1 M-Tris, 0.01 M-EDTA, pH 8.5, 40 µg/ml cytochrome c and formamide was layered onto a formamide hypophase in 0.01 M-Tris, 0.001 M-EDTA, pH 8.5 by allowing it to run down a microscope slide inserted into the hypophase. DNA in 40 to 80% formamide was layered onto a hypophase 30% lower in formamide, and DNA in 30% formamide was layered onto a hypophase 5% in formamide. The DNA/cytochrome film was picked up onto parlodion-coated grids and was stained with uranyl-acetate and shadowed with platinum-palladium.

(j) Electron microsocpy and molecule measuring

Molecules were photographed in a Phillips EM 300 electron microscope. The molecules were projected onto the measuring surface of a Hewlett-Packard Model 9864A Digitizer and the different regions of the molecules traced and measured. Single-stranded regions were normalized to double-stranded lengths by conversion factors determined in a series of mountings of ϕX174 single-stranded DNA and ϕX174 RF. Although this method of normalization is not as good as one employing internal standards no significant variation was seen in the ratio of double-stranded and single-stranded DNA in a series of mountings under the same conditions.

3. Results

(a) Renaturation of ϕX174 complementary and S13 viral DNA

Wetmur & Davidson (1968) showed that the maximal rate of renaturation of denatured DNA occurs under conditions equivalent to 25°C below the melting temperature of the DNA. Since the melting temperature of the ϕX174/S13 heteroduplex varies widely in different parts of the molecule it is not possible to select one set of renaturation conditions corresponding to $(T_m - 25)°C$. Under the conditions used to renature ϕX174 RF II the rate of heteroduplex formation is slow, reflecting the small proportion of each of the molecules containing highly conserved sequences. Under less denaturing conditions large multiple renaturation masses formed quickly. The gradient of formamide concentration in the renaturation mixture described in Materials and Methods was an attempt to provide conditions under which a large portion of each heteroduplex could renature. The time of renaturation was a compromise between lower yield of heteroduplexes and excessive aggregation. Under these conditions between 10 and 25% of the circular molecules were observed as heteroduplexes.

A spreading solution of 40% formamide and a hypophase of 10% formamide were adopted as standard conditions for the purpose of comparing different heteroduplex preparations. Under these conditions four different types of heteroduplex structures were seen. An example of the most common heteroduplex, type I, is shown in Plate Ia along with an interpretive drawing. There are three single-stranded loops separated by two short, double-stranded regions and one long conserved

stretch accounting for approximately 25% of the molecule. The two loops on either side of the long duplex region are of two consistently different sizes and the type I molecules can be ordered as a set. The smaller of the two loops adjacent to the long duplex region will be called L2 and the larger L3; the loop between L2 and L3 will be called L1. In addition to these three basic loops type II molecules contain a small loop adjacent to L1 on the side toward L2, such as shown in Plate Ib. Loop L2 contains some DNA sequences which can occaisionally form a stable duplex region under the standard conditions. Type III molecules have two small loops in place of the L2 loop; an example is shown in Plate IIa. The fourth type of heteroduplex seen combines the characteristics of type II and type III molecules and has a small loop adjacent to loop L1 and two small loops in place of loop L2. A type IV molecule is shown in Plate IIb.

An occasional molecule mounted under standard conditions will show a small loop in the long double-stranded region between loops L2 and L3. The loop occurs at either of two places in the double-stranded region and its appearance is independent of the small loop found in type II or type IV molecules and of the appearance of one or two loops in the L2 region. For purposes of classification molecules containing such a loop are typed according to their other features and then are noted as having the extra loop. Plate IIIa shows a type I molecule containing such an extra loop in the region between L2 and L3. Under higher formamide conditions, as will be discussed in a later section, the two loops in the region between loops L2 and L3 join to form a larger loop that separates the two most highly conserved regions in the

heteroduplex. Plate IIIb shows a heteroduplex between a circular S13 viral strand and a linear ϕX174 complementary strand fragment spread from 50% formamide. The two small loops between loops L2 and L3 have each formed without melting the short duplex segment between them. The occasional appearance of these loops under standard conditions has allowed the orientation of the denaturation patterns seen at higher formamide concentration with respect to the pattern seen under standard conditions.

(b) The length of the two DNAs

Although single-stranded DNA is expected to vary in length the two single-stranded arms of loop L1 show a consistent difference in length that is greater than the differences between the arms of the other two loops. The measured length of a DNA region contains errors associated with the stretching of the DNA during preparation and with the measurement itself. Since the heteroduplexes that have been photographed and measured are a sampling from a population and since the errors in the sampling can be expressed, the difference between two loop arms can be tested for its significance. Our null hypothesis is that the two arms of the loop are the same length. To test the hypothesis we want to compare the actual difference with the expected difference, zero. We calculate a chi-square value using

$$\frac{\sum_{i=1}^{N} (D_{1_i} - D_{2_i})^2}{S^2}$$

where $D_{1_i} \geq D_{2_i}$ are the two loop arm values in a sample of N pairs and

S^2 is the variance of the sample of 2N values. For loops L2 and L3 the chi-square value gives a confidence level of 0.35, indicating that the probability of a series of larger differences occurring in a sample where the two arms are the same length is 0.35. This indicates that the arms of loops L2 and L3 are probably of equal length. The confidence level for loop L1, however, is less than 0.001 and indicates that the deviation from our null hypothesis is significant. We conclude that the two DNAs are of unequal length and that this difference is localized in the region of loop L1.

The S13 DNA used to form the heteroduplexes was greater than 95% circular and the ϕX174 complementary strand was less than 10% circular. As will be described in a later section, heteroduplexes mounted under higher formamide concentrations contain decreasing amounts of duplex DNA. As the amount of duplex DNA decreases the probability of finding a heteroduplex with the nick in the ϕX174 complementary strand in a duplex region also decreases. Most molecules therefore have single-stranded linear tails (see Plate VI for two examples). Because of the relative proportions of circles and linears in the two DNA preparations a heteroduplex formed between a circular ϕX174 complementary strand and a linear S13 strand should occur with a frequency of only about 1%. Therefore most of the tailed portions will be the ϕX174 strand.

In each of the heteroduplexes measured the circular strand was longer than the linear. Although the ϕX174 complementary strand preparation contained some DNA fragments, the longest linear strands in the heteroduplexes were shorter than the circular strands by the

same difference as observed between the L1 arms. We conclude that the S13 viral DNA used in these experiments was 8% longer than the ϕX174 complementary strands. Spencer et al. (1972) found values of contour length for S13 single-stranded DNA and RF which are larger than most of the values reported for ϕX174. Although contour length varies with the conditions of preparation their results are consistent with the difference in length we see.

(c) The calculation of average molecular profiles

In order to describe the changes observed within the heteroduplex denaturation pattern with increasing formamide concentration it is desirable to be able to describe the heteroduplex molecules in an average molecular profile. The heteroduplexes were classified as to type and an average value determined for each region within each type. All single-stranded loops, with the exception of L1, were treated by averaging the two loop arms, normalized to double-stranded length. The two arms of L1 were normalized and averaged and half of the average difference between the two arms was subtracted to estimate the length of the ϕX174 arm of the loop. A total contour length was calculated for each molecule using the ϕX174 value for L1. The molecule was then normalized to unit length and each region was expressed as the proportion of the ϕX174 length. The average molecular profile was calculated from the average size of each region of the molecule.

Although this method of calculating an average does not show variation in loop size we feel that it is more helpful in indicating the appearance of a typical molecule than a pure probability distri-

bution. Loops which are common to all molecules are represented by their average size and the difference in the arms of L1 is shown by a loop with two different lengths above and below the central line representing duplex DNA. The loops and duplex regions which are not common to all molecules are represented by lines closer to the central line indicating duplex DNA. The separation of the loop arms from the central duplex line represents the probability of finding that loop in a molecule. The size and position of these occasional features on the average molecular profile represents the average size and position. Figure 1 shows molecular profiles for the five heteroduplexes used as examples of the standard condition molecules and shown in Plates I, II and IIIa. Figure 1f is the average molecular profile for heteroduplexes prepared under standard conditions.

(d) The molecular profile as a function of formamide concentration

Helmkamp & Ts'o (1961) and Ts'o et al. (1962) showed that formamide removed all formal secondary structure from DNA. McConaughy et al. (1969) reported that a 1% increase in formamide concentration lowers the melting temperature of DNA by 0.72°C. Although the exact value is in doubt (Bluthmann et al., 1973; Tibbets et al., 1973) it is clear that formamide can denature DNA. Heteroduplexes were mounted under a series of formamide concentrations in order to determine the average molecular profile as a function of the approach to the melting point of ϕX174 RF II. Figure 2 shows average molecular profiles for heteroduplexes mounted under each of the conditions.

FIG. 1. Molecular profiles of the ϕX174/S13 heteroduplexes shown as representative of the molecules seen under standard conditions. In each profile the heavy, central line represents double-stranded DNA and the light, separated lines represent single-stranded loops. All single-stranded DNA was normalized to double-stranded lengths and the two arms of each loop, except for L1, were averaged. The length of each heteroduplex is normalized to unit length including the shorter arm of L1, and the longer arm of L1 is shown as a trapezoid above the central line. The end of loop L2 on the L1 side is normalized at 0.265 genome lengths for each profile. (a) The type I molecule from Plate Ia. (b) The type II molecule from Plate Ib. (c) the type III molecule from Plate IIa. (d) The type IV molecule from Plate IIb. (e) The type II molecule containing a loop in the long duplex region between L2 and L3 shown in Plate IIIa. (f) The average molecular profile for all molecules seen under standard conditions. The calculation and interpretation of the average profile are described in the text.

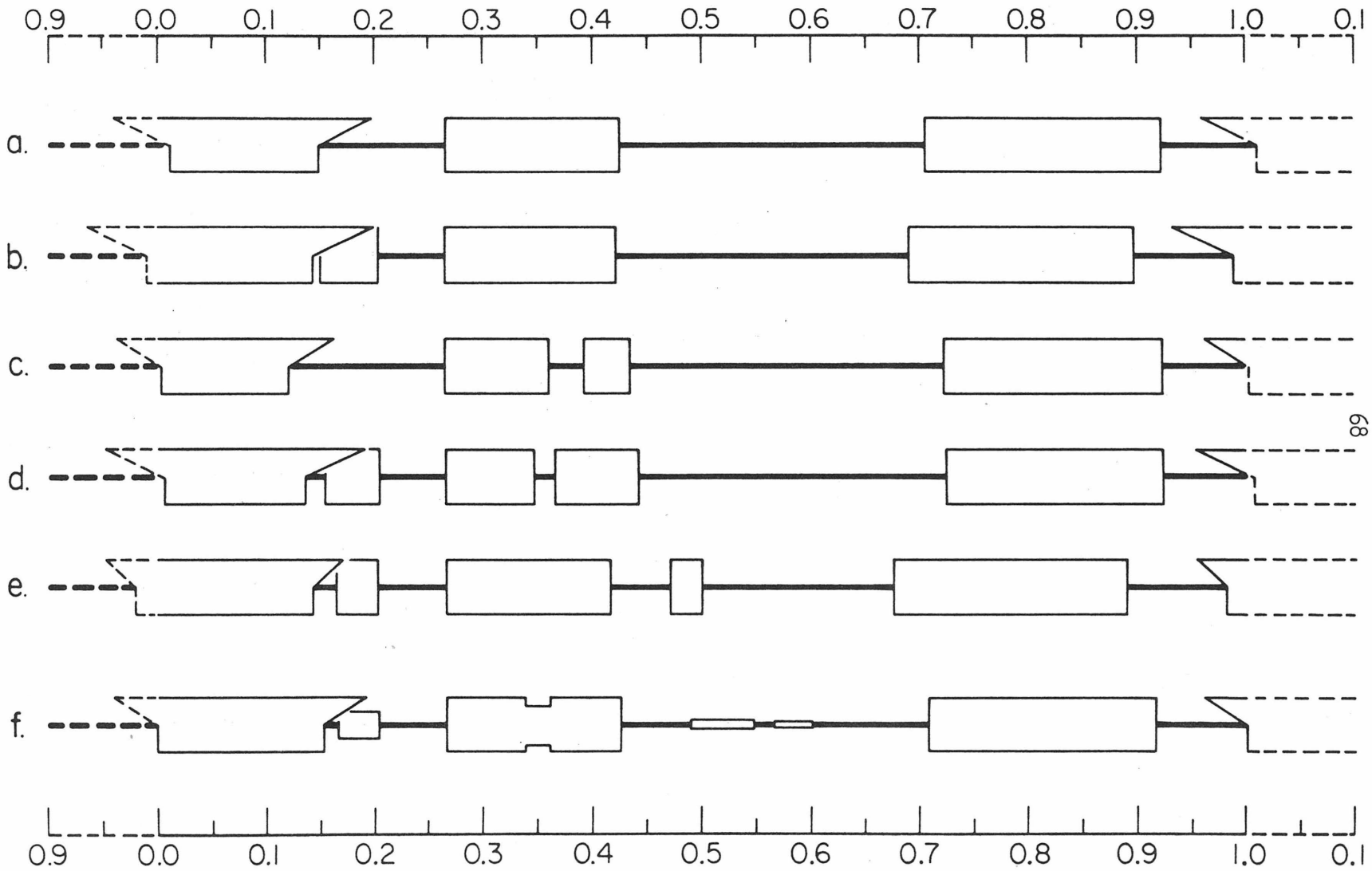

0.9
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.1
a.
b.
c.
d.
e.
f.
0.9
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.1

FIG. 2. The average molecular profiles of the ϕX174/S13 heteroduplexes seen under the following spreading conditions. (a) 30% formamide. (b) 40% formamide (standard conditions). (c) 50% formamide. (d) 60% formamide. (e) 70% formamide. (f) 80% formamide.

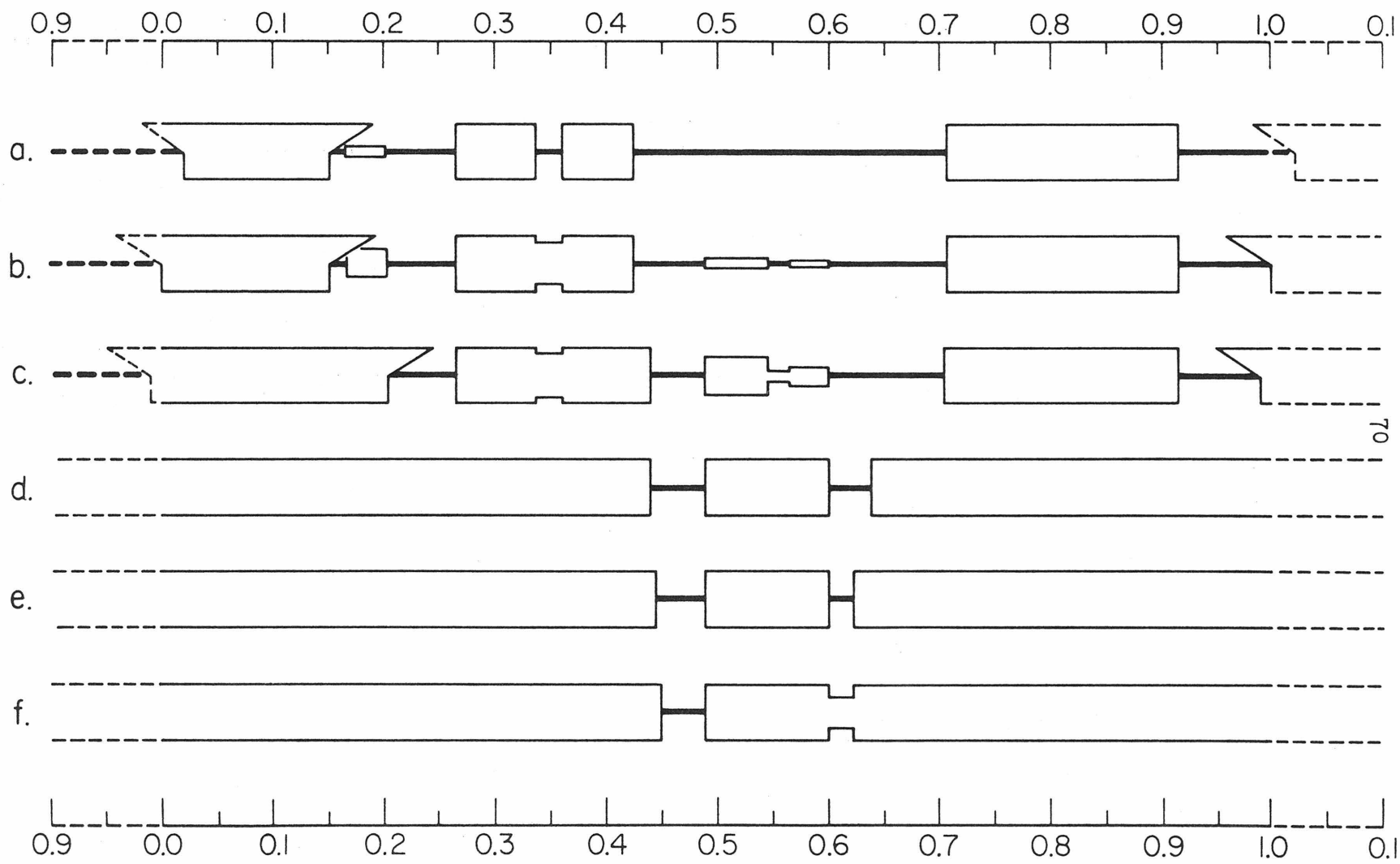

0.9
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.1
a.
b.
c.
d.
e.
f.
70

Plate IVa shows a heteroduplex mounted from 30% formamide. All of the molecules mounted under these conditions had two small loops in the L2 loop region which confirms that the standard conditions are close to the denaturation point for the duplex region between the two small loops. Loop L3 remained the same size as under standard conditions; loop L1 was smaller, fewer molecules contained the small loop adjacent to loop L1 and none had a loop in the long duplex region between loops L2 and L3. Plate IVb shows a heteroduplex molecule mounted under standard conditions for comparison.

Heteroduplexes spread from 50% formamide contained an L1 loop that grew in size largely by including the small loop which was adjacent to it at 30 and 40% formamide. Fewer than 25% of the molecules had two separate loops in the L2 loop region and the L3 loop increased very slightly in size. The proportion of molecules with one or both of the loops in the region between loops L2 and L3 increased and a few of the molecules contained one large loop in place of the two smaller ones. Plate Va shows a rare molecule from 50% formamide with the duplex region between loops L1 and L2 melted, producing one large loop.

Heteroduplexes mounted from 60 and 70% formamide were similar in appearance. Two short duplex regions bounding a loop corresponding in size to the sum of the two occasional loops found in the region between loops L2 and L3 at lower formamide concentrations were the only conserved regions. The profiles in Figure 2d and Figure 2e show that part of the duplex found at 60% formamide melts in 70% formamide. There is no way to be certain if the smaller duplex remaining in 70% formamide corresponds to the smaller duplex region in

60% formamide although the similarity in size suggests the progression shown in Figure 2. Plate Vb shows a heteroduplex molecule mounted from 60% formamide; molecules mounted from 70% formamide are indistinguishable except by careful measuring.

Approximately half of the molecules mounted from 80% formamide were of the type seen in 70% formamide and half contained only one duplex region. Figure 2f shows the average molecular profile and Plate VI shows molecules of each type.

(e) The insertion of specific ϕX174 DNA fragments into ϕX174/S13 heteroduplex molecules

R1 is the largest fragment produced by cleavage of ϕX174 RF with the restriction endonuclease Hind from Hemophilus influenzae. It probably is between 1000 and 1150 base pairs in length (for a discussion see Compton & Sinsheimer, 1975) and includes almost all of gene F. R1 was included in a renaturation between ϕX174 complementary strand DNA and S13 viral DNA in an attempt to construct a heteroduplex molecule containing an internal marker. The rate of heteroduplex formation in the initial renaturation in 60% formamide is low, but the rate of renaturation between the ϕX174 complementary strand and the viral strand of the R1 fragment is normal. The pre-incubation in 60%

formamide should increase the number of ϕX174 complementary strands carrying an annealed R1 strand and thus increase the yield of heteroduplexes containing the R1 strand.

Heteroduplexes prepared in the presence of R1 and mounted under standard conditions were of the types described in section (a). In more than half of the heteroduplexes the double-stranded region between loops L2 and L3 was only about half as long as in the heteroduplexes without R1 and L3 was correspondingly larger. The increase in size of L3 was specific and reproducible and a region of double-stranded DNA could be distinguished in one of the loop arms. We conclude that R1 specifically anneals to the ϕX174 complementary strand in the region adjacent to, and partially within, loop L3 in the ϕX174/S13 heteroduplex. Although the complementary strand of the R1 fragment could bind to the S13 viral strand only that part of the fragment which corresponded to the duplex region adjacent to L3 would form a duplex hybrid; the part of R1 which corresponded to the ϕX174 arm of the L3 loop would form a single-stranded tail. A ϕX174/S13 heteroduplex molecule containing an R1 fragment bound to the S13 strand would therefore have half of the R1 strand bound to the S13 arm of the enlarged L3 and half as single-stranded tail within L3. No such structures were seen. Plate VIIb is an example of a heteroduplex containing an R1 strand.

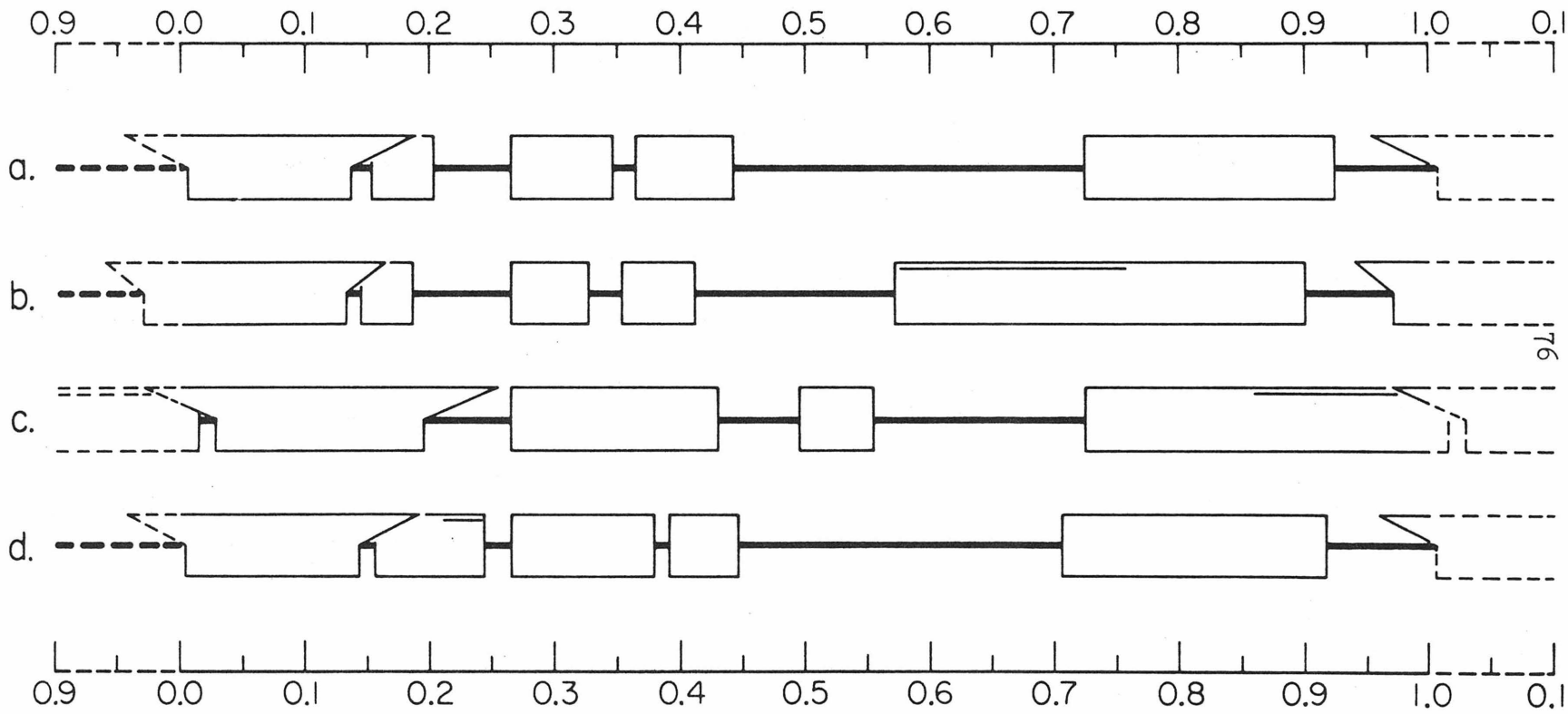
Since the ϕX174/S13 heteroduplex is circular only the point in the center of the fragment and the point on the opposite side of the heteroduplex ring can be located on the ϕX174 genetic map. Heteroduplexes containing R2 were prepared in order to provide a second

marker. The heteroduplexes containing R2 contained an L3 loop enlarged at the end opposite to that affected by R1 and a correspondingly shorter duplex region between loops L1 and L3. Again, no structures which could be interpreted as the homologous portion of R2 annealed to an S13 strand in a heteroduplex were seen. Plate VIIIa shows a ϕX174/S13 heteroduplex molecule containing an R2 fragment.

Using fragments R1 and R2 as markers the Hind cleavage map of ϕX174 can be superimposed on the ϕX174/S13 heteroduplex profile. The positioning of the cleavage map indicates that R8 should then lie in the region between L1 and L2, at or near the site of the small occasional loop adjacent to L1. Heteroduplexes containing R8 were prepared and Plate VIIIb shows a typical molecule. The loop between L1 and L2 is about twice as large as one that would be expected to be formed by R8, but it is also larger than was ever seen in renaturations without R8. We conclude that R8 anneals at a site adjacent to the occasional loop between L1 and L2 and produces one loop the size of both the occasional loop and the R8 fragment by destabilizing the heteroduplex in the region of the occasional loop.

Figure 3 shows molecular profiles for the four molecules in Plates VII and VIII. Each of the three fragments produces a unique, visible change in the heteroduplex molecular profile in regions that map in relation to each other just as would be expected from the Hind cleavage map of ϕX174 RF.

FIG. 3. Molecular profiles of the ØX174/S13 heteroduplexes shown in Plate VII and Plate VIII. (a) The type IV molecule without R fragment from Plate VIIa. (b) The molecule containing an R1 fragment from Plate VIIb. (c) The molecule containing an R2 fragment from Plate VIIIa. (d) The molecule containing an R8 fragment from Plate VIIIb. In each case the presence of a fragment is shown by an extra line within the extended loop.

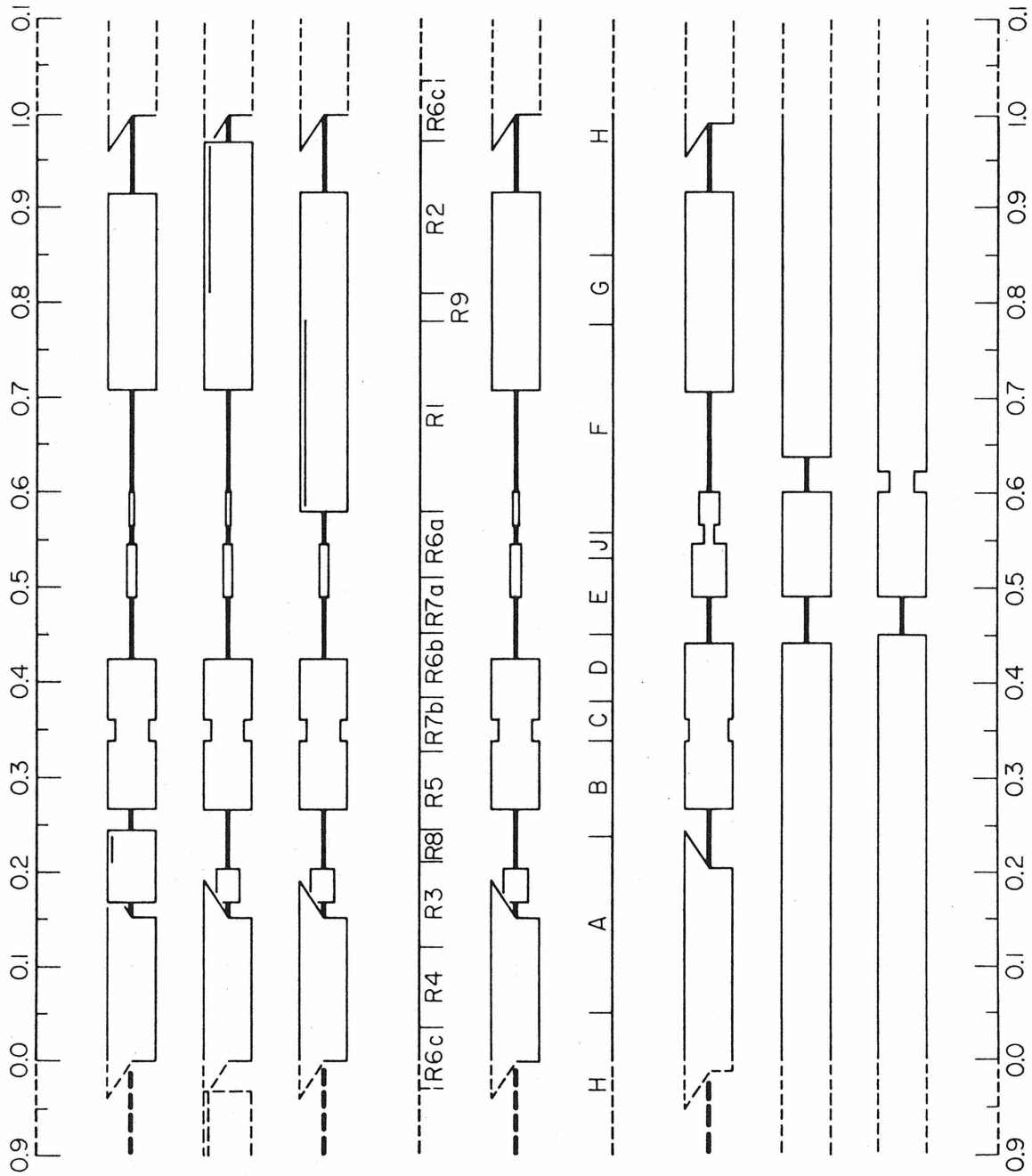

0.9
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.1
a.
b.
c.
d.
0.9
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.1

(f) The relationship between the ϕX174 genetic map and the ϕX174/S13 heteroduplex map

The fragments produced from ϕX174 RF by Hind have been ordered onto a cleavage map of the genome by analyses of products of partial digestions and by sequential use of two or three restriction enzymes (Middleton et al., 1972; Lee & Sinsheimer, 1974), and the cleavage map has been oriented on the ϕX174 genetic map by infecting bacterial spheroplasts with defective single-stranded molecules carrying an annealed wild type restriction fragment (Middleton et al., 1972). Compton & Sinsheimer (1975) have shown that the ϕX174 genetic map can be superimposed on RF II molecules containing two insert loops produced by annealed R fragments. In the same manner the ϕX174 genetic map can be superimposed on the heteroduplex molecules containing R fragments, and consequently on the ϕX174/S13 denaturation map.

The molecular profile in the center of Figure 4 is the ϕX174/S13 average profile under standard conditions. The Hind cleavage map of ϕX174 is immediately above the standard profile and the ϕX174 genetic map is immediately below. The three profiles above the center are the average molecular profiles for heteroduplexes containing, from the center, R1, R2 and R8, and show how the inserted fragments correspond to their positions on the cleavage map. The three profiles below the center are the average molecular profiles of the ϕX174/S13 heteroduplex in, from the center, 50, 60 and 80% formamide and show the relationship betwe conservation of duplex DNA and the ϕX174 genetic map. This relationship enables us to say that the most highly conserved regions between ϕX174 and S13 are in genes E and F.

FIG. 4. The relationships between the average molecular profiles of the ϕX174/S13 heteroduplexes, the Hind cleavage map of ϕX174 RF and the ϕX174 genetic map. The profile in the center is the average molecular profile for ϕX174/S13 heteroduplexes seen under standard conditions. The Hind cleavage map is immediately above the standard profile and the ϕX174 genetic map is immediately below. The three profiles above the center are the average molecular profiles for heteroduplexes containing fragments R1, R2 and R8, respectively. The three profiles below the center are the average molecular profiles for heteroduplexes seen in 50%, 60% and 80% formamide, respectively.


0.9
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.1
R6c
R4
R3
R8
R5
R7b
R6b
R7a
R6a
R1
R9
R2
R6c
H
A
B
C
D
E
J
F
G
H

4. Discussion

It is clear from the relative stabilities of the phage particles (Maclean & Hall, 1962), their different sensitivities to the same antiserum (Zahler, 1958), and the pyrimidine tracts of the DNAs (Darby et al., 1970; Cerny et al., 1969) that ϕX174 and S13 are different phages. At the same time, however, they are clearly closely related. The two phages can complement in mixed infections in all but one cistron (Jeng et al., 1970) and there is a possibility that even this one qualification results from the use of a phage carrying two mutations in different cistrons. The relationship between ϕX174 and S13 was extended by the observation of a low level of recombination between them (Tessman & Schleser, 1963). ϕX174 and S13 were isolated several years and several hundred miles apart (Sertic & Boulgakov, 1935; Burnet, 1927) and excluding the possibility that S13 was carried to the site of the isolation of ϕX174 and rediscovered the two phages are most likely descended from a common ancestor. It was the opportunity to examine the evolutionary divergence within such a simple well-characterized virus that initially led to this project. Because of the extensive similarities between the two phages the degree of sequence divergence indicated by the heteroduplex map was unexpected.

The ϕX174/S13 heteroduplex denaturation map described in this report is similar to the one described by Godson (1973) but differs in two important points. We see less homology under mildly denaturing conditions and more homology under conditions which should be more stringent than Godson's most severe conditions. The denaturation of a heteroduplex molecule is a function of the ionic strength of the

spreading solution and hypophase and of the temperature during mounting. The ionic strength of the solution used in the DNA preparation should be identical or very similar since we have both used the spreading conditions of Davis & Hyman (1971). Our heteroduplexes were uniformly prepared at 22°C whereas Godson reported that he used 20°C. It is possible that this 2°C difference during mounting is responsible for the different profile we see in 30% formamide if the melting temperature of the regions which we see as single-stranded is less than 2°C lower than the temperature simulated by 30% formamide. If the difference between our work is based on this 2°C, however, then our results are even more divergent under highly denaturing conditions. Our preparation in 80% formamide is both 5% higher in formamide and mounted at a temperature 2°C higher, and yet we see more homology between the two DNAs.

Godson (1973) reported that he saw no difference in the lengths of the two DNAs, while we see a consistent difference of about 8% localized in the region of genes A and H. Although a judgment must be made from a large sample rather than from a few isolated molecules the examples of heteroduplex structure in Godson (1973) each show a difference in strand length. The two circular molecules in his Plate Ic differ by 7% and the lower loop of the molecule in his Plate Ib has arms that differ by 24% of the total contour length of the molecule if we accept his interpretation of single- and double-stranded DNA. The length of single-stranded DNA varies greatly and these differences may not be significant, but all three of the molecules in his paper which could be measured showed loop arm differences

greater than the average difference we find for loops which we believe have equal length arms. If there is a length difference of 8% in his samples then heteroduplexes which were completely double-stranded (which Godson observed in 30% formamide) would have contained a deletion loop of the type seen by Zuccarelli et al. (1972 and Kim et al. (1972), unless the length difference is so distributed that any loops would be too small to be seen. This question cannot be resolved at this time.

Under the least stringent conditions used in this study, 30% formamide, all of gene G (spike) and most of genes A (RF replication), B (structural; single-strand replication), D (single-strand replication) and H (spike) are in non-homologous regions of the heteroduplex. 30% formamide lowers the T_m of DNA by 21.6°C (using the value of ΔT_m=-0.72°C/1% formamide from McConaughy et al. (1969)) or by 16.8°C (using the value of ΔT_m=-0.56°C/1% formamide from Tibbets et al. (1973)). Since the T_m of ϕX174 RF II under these ionic conditions is 82°C in the absence of formamide (Godson 1973) the T_m of ϕX174 RF II in 30% formamide would be lowered to either 60.4 or 65.2°C, which are 38.4 and 43.2°C above the mounting temperature, respectively. Laird et al. (1969) reported that a 1.5% base mismatch lowers the melting temperature of DNA by 1°C. Thus in a heteroduplex DNA region which is single-stranded in 30% formamide the minimum base mismatch must be between 57.6 and 64.8%, depending on the value chosen for the formamide lowering of the T_m. For a region which remains duplex in 30% formamide but is melted in 40% formamide the minimum base mismatch is between 46.8 and 56.4% and the maximum is between 57.6 and 64.8%. Similarly, the range of base mismatch is 36.0 to 56.4% for a region which first melts in 50% formamide, 25.2 to 48.0% for a region

which first melts in 60% formamide, 14.4 to 39.6% for a region which first melts in 70% formamide and 3.6 to 31.2% for a region which first melts in 80% formamide.

The double-stranded region which occurs within the L2 region under mild conditions is contained within gene C (single-strand replication), and the conserved region between the two small loops in the long duplex region between loops L2 and L3 is within gene J (structural ?). The most highly conserved regions fall in genes E (lysis) and F (capsid). Poljak and Suruda (1969) reported that they could detect no differences in the amino acid compositions of the major capsid proteins (gene F) of the two phages, although they were unable to rule out the substitution of non-polar residues. The heteroduplex map, however, shows that about one quarter of gene F nearest gene G contains at least 57.6 to 64.8% base mismatch, the quarter adjacent to gene J between 36.0 and 56.4%, and all but a small region on the J side of the center at least 25.2 to 39.6% base mismatch. In spite of the fact that it is partially conserved in 80% formamide the small duplex region in gene F can still contain between 3.6 and 31.2% base mismatch. The half of gene E adjacent to J contains between 36.0 and 56.4% base mismatch, but the half adjacent to D remains duplex in 80% formamide and should contain less than 22.8% base mismatch, and perhaps less than 3.6%.

Gene E is the one function which is not necessary to yield mature, infective phages from an infection and it is interesting that it should be so highly conserved. Since the gene products of the cistrons which contain extensive base mismatches still complement in a mixed infection, the DNAs must code for interchangeable products in each case. The

redundancy within the genetic code can account for more than 33% base mismatch without changing the amino acid sequence of the gene product, depending on the amino acids involved. The substitution of amino acids which do not alter the function of the gene product can extend the degree of acceptable base mismatch. The lack of divergence of the DNA sequences coding for the gene E product suggests either that the protein contains amino acids which show less redundancy in the genetic code or residues for which substitutions cannot be tolerated. In each of the ϕX174 and S13 genes A, B, D, E, F and H about one-half shows a high degree of base mismatch while the other half is more highly conserved. This repetitive pattern suggests that in each case about half of the gene has resisted divergence and may reflect the presence of an active site or structurally important region in the protein product.

ϕX174 and S13 are closely related, and yet have widely differing DNA sequences. Besides the opportunity to study the divergence between related organisms the system offers the possibility of examining the relationship between DNA sequence and protein function. As DNA and protein sequences become available a comparison of the two phages will indicate the degree to which amino acid sequence is independent of DNA sequence and the degree to which protein function is independent of amino acid sequence. The preliminary work reported here indicates that the phage-host system is very tolerant of diversity.

REFERENCES

Baas, P. D. & Jansz, H. S. (1972). J. Mol. Biol. 63, 557-568.

Baker, R. & Tessman, I. (1967). Proc. Nat. Acad. Sci. USA 58, 1438-1445.

Beckendorf, S. K., Kim, J. H. & Lielausis, I. (1973). J. Mol. Biol. 73, 17-35.

Benbow, R. M., Zuccarelli, A. J., Davis, G. & Sinsheimer, R. L. (1974). J. Virol. 13, 898-907.

Blüthmann, H., Brück, D., Hübner, L. & Schöffski, A. (1973). Biochem. Biophys. Res. Commun. 50, 91-97.

Burnet, F. M. (1927). Brit. J. Exp. Path. 8, 121-129.

Cerny, R., Cerna, E. & Spencer, J. H. (1969). J. Mol. Biol. 46, 145-156.

Compton, J. L. & Sinsheimer, R. L. (1975). J. Mol. Biol.

Darby, G., Dumas, L. B. & Sinsheimer, R. L. (1970). J. Mol. Biol. 52, 227-238.

Davis, R. W. & Hyman, R. W. (1971). J. Mol. Biol. 62, 287-301.

Davis, R. W., Simon, M. N. & Davidson, N. (1971). In Methods in Enzymology, L. Grossman & K. Moldave, eds. Vol. XXI, part D, pp. 413-428, Academic Press, New York.

Godson, G. N. (1973). J. Mol. Biol. 77, 467-477.

Helmkamp, B. K. & Ts'o, P. O. P. (1961). J. Amer. Chem. Soc. 83, 138-142.

Jeng, Y., Gelfand, M., Hayashi, M., Schleser, R. & Tessman, E. S. (1970). J. Mol. Biol. 49, 521-526.

Johnson, P. H. & Sinsheimer, R. L. (1974). J. Mol. Biol. 83, 47-61.

Kim, J.-S., Sharp, P. A. & Davidson, N. (1972). Proc. Nat. Acad. Sci. 69, 1948-1952.

Laird, C. D., McConaughy, B. L. & McCarthy, B. J. (1969). Nature 224, 149-154.

Lee, A. S. & Sinsheimer, R. L. (1974a). Proc. Nat. Acad. Sci. USA 71, 2882-2886.

Maclean, E. C. & Hall, C. E. (1962). J. Mol. Biol. 4, 173-178.

McConaughy, B. L., Laird, C. D. & McCarthy, B. J. (1969). Biochemistry 8, 3289-3295.

Middleton, J. H., Edgell, M. H. & Hutchison, C. A., III (1972). J. Virol. 10, 42-50.

Poljak, R. J. & Suruda, A. J. (1969). Virology 39, 145-147.

Sertic, V. & Boulgakov, N. (1935). Compt. Rend. Soc. Biol. 119, 1270-1282.

Simon, M. N., Davis, R. W. & Davidson, N. (1971). In The Bacteriophage Lambda, A. D. Hershey, ed., pp. 313-328. Cold Spring Harb. Lab. Quant. Biol.

Sinsheimer, R. L. (1968). Prog. Biophys. Nucleic Acid Res. 8, 115-169.

Spencer, J. H., Cerny, R., Cerna, E. & Delaney, A. D. (1972). J. Virol. 10, 134-141.

Tessman, E. S. & Schleser, R. (1963). Virology 19, 239-240.

Tibbets, C., Johansson, K. & Philipson, L. (1973). J. Virol. 12, 218-225.

Ts'o, P. O. P., Helmkamp, G. K. & Sander, C. (1962). Biochim. Biophys. Acta 55, 584-600.

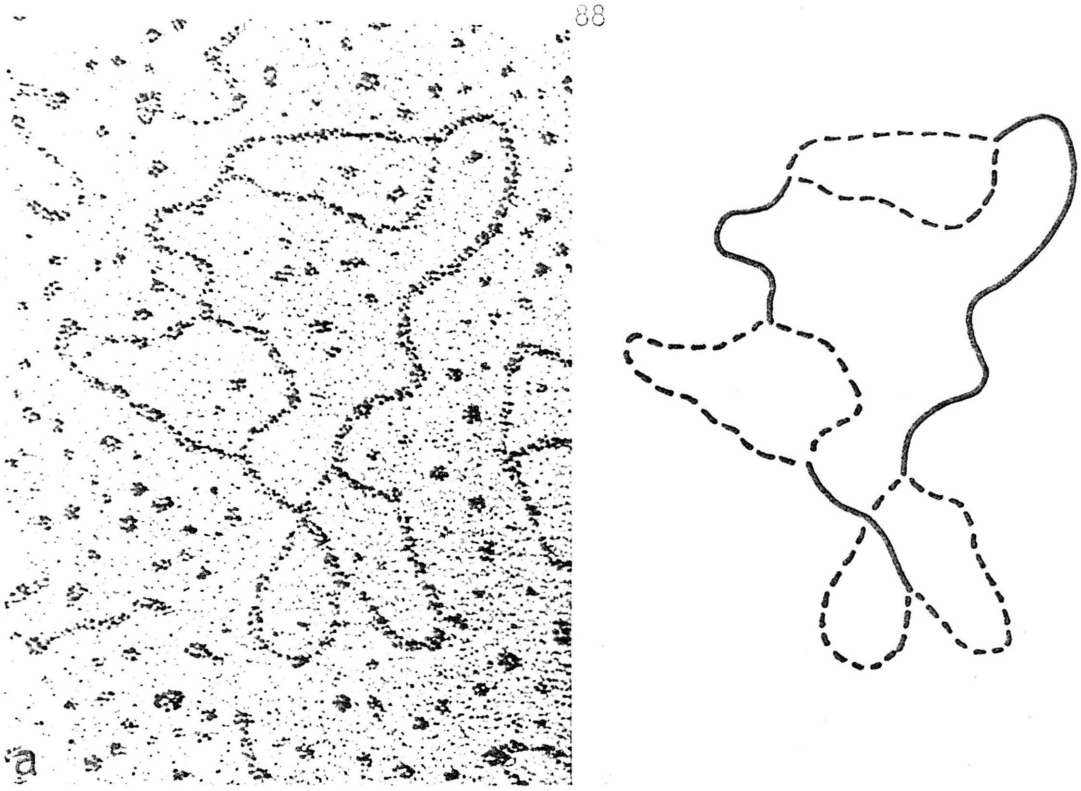
Westmoreland, B. C., Szybalski, W. & Ris, H. (1969). Science 163, 1343-1348.

Wetmur, J. G. & Davidson, N. (1968). J. Mol. Biol. 31, 349-370.

Zahler, S. A. (1958). J. Bacteriology 75, 310-315.

Zuccarelli, A. J., Benbow, R. M., Davis, G. & Sinsheimer, R. L. (1974). J. Virol., in press.

PLATE I: Micrographs and interpretive drawings of ϕX174/S13 heteroduplex molecules seen under standard conditions. Solid lines in the drawing represent double-stranded DNA and dashed lines represent single-stranded DNA. (a) A type I molecule. Loop L1 is in the center on the left, L2 is at the top and L3 is at the bottom. The long duplex region between L2 and L3 runs vertically on the right side of the molecule. (b) A type II molecule. The small loop which appears between L1 and L2 is just above center on the left.


a

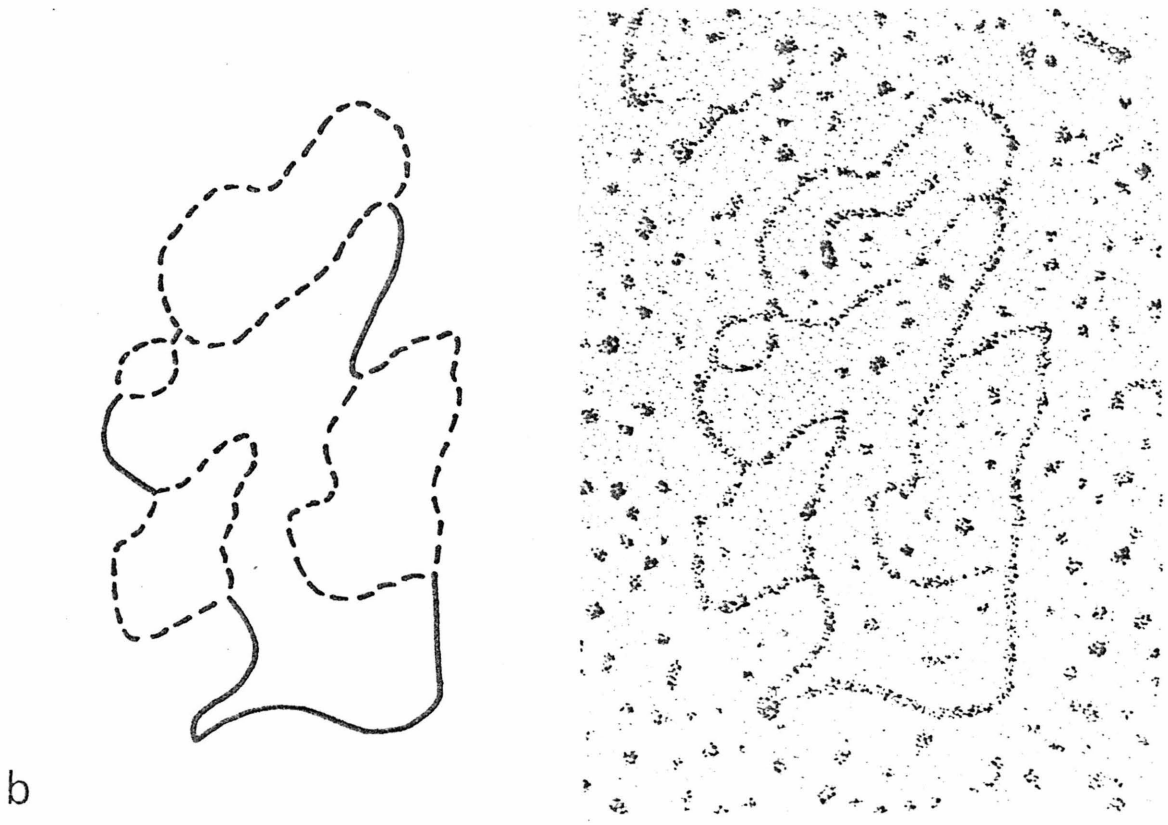

b

PLATE II: Standard heteroduplexes (continued). (a) A type III molecule. The L2 region shows two small loops in the right center of the molecule. (b) A type IV molecule. L1 is at the top left and L3 is at the bottom left.

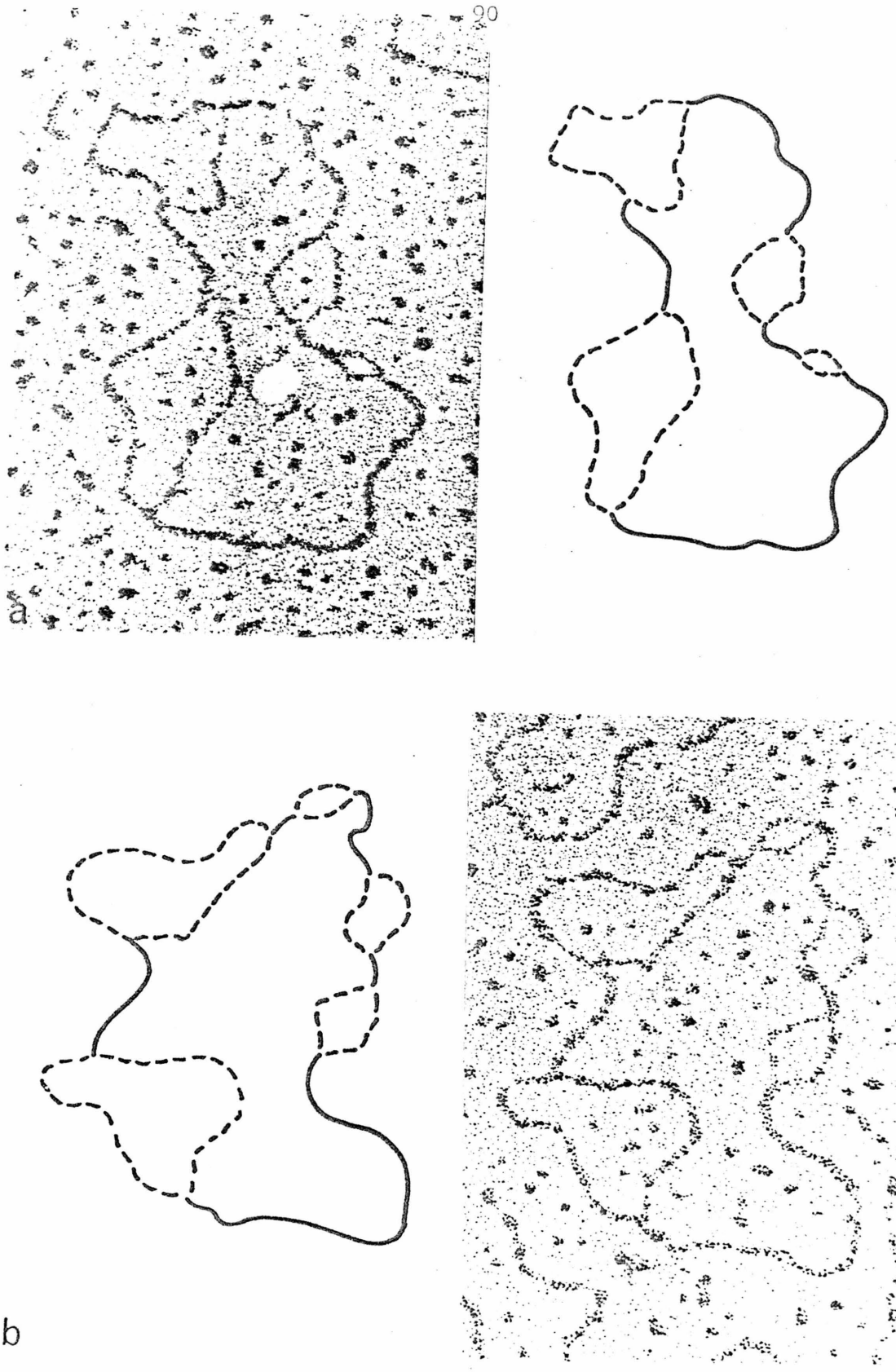
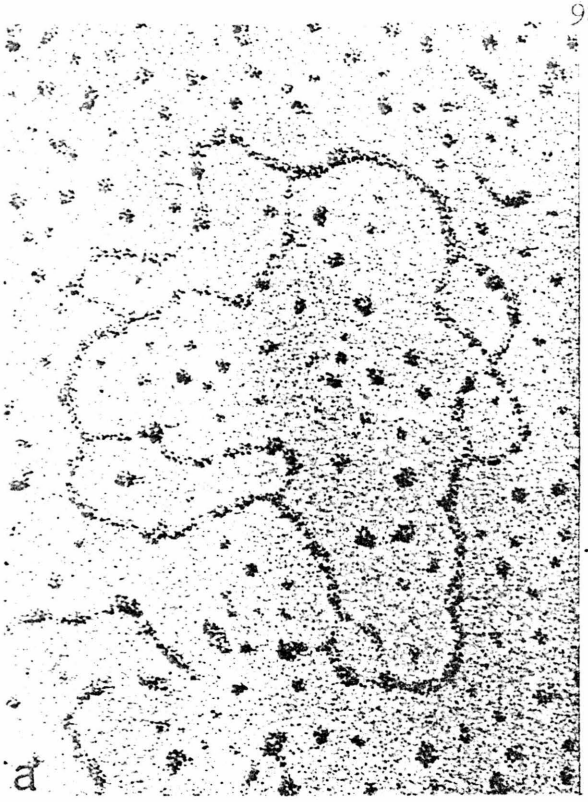
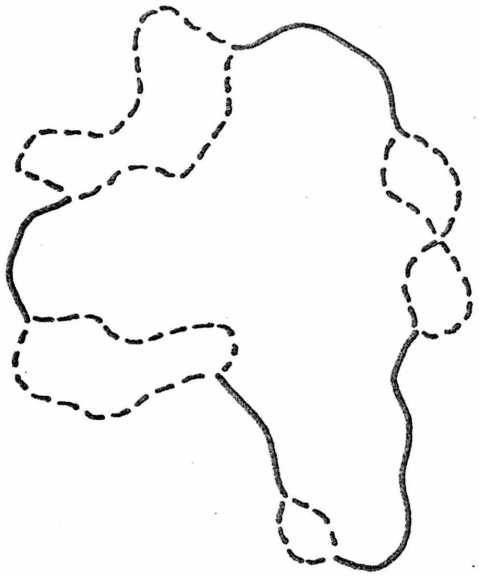
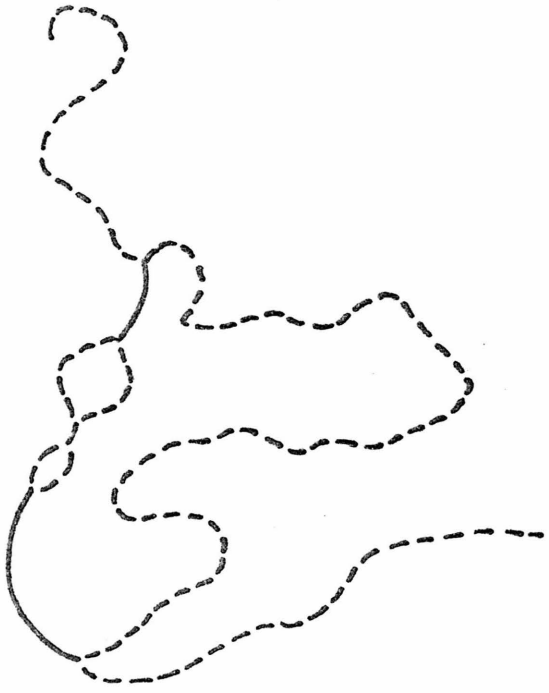

90
a
b

PLATE III: ϕX174/S13 heteroduplex molecules containing loops in the duplex region between L2 and L3. (a) A molecule from standard conditions with a single loop at the bottom of the molecule. (b) A heteroduplex including a linear fragment. Because of the proportions of circles and linears in the two DNA preparations this probably represents an S13 circle and a ϕX174 linear. This molecule was spread from 50% formamide and contains two small loops in an otherwise duplex region which corresponds in size to the region between L2 and L3 under these conditions.



a





b

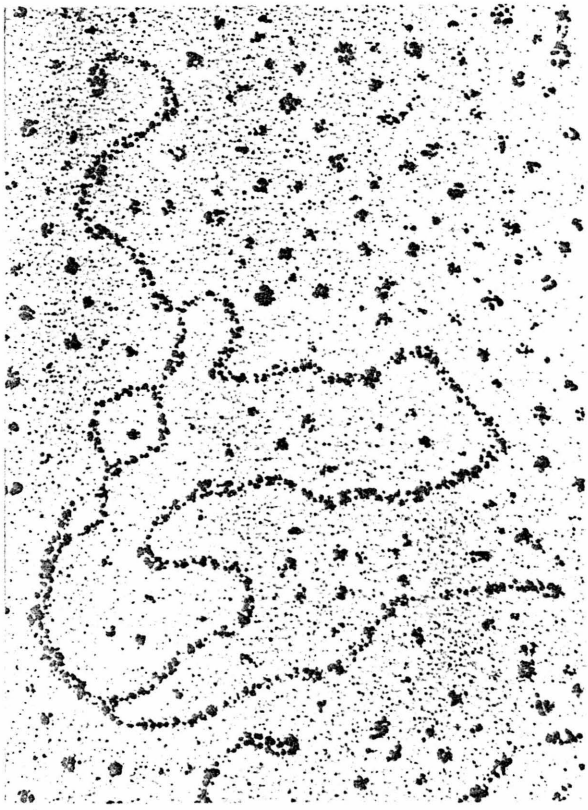
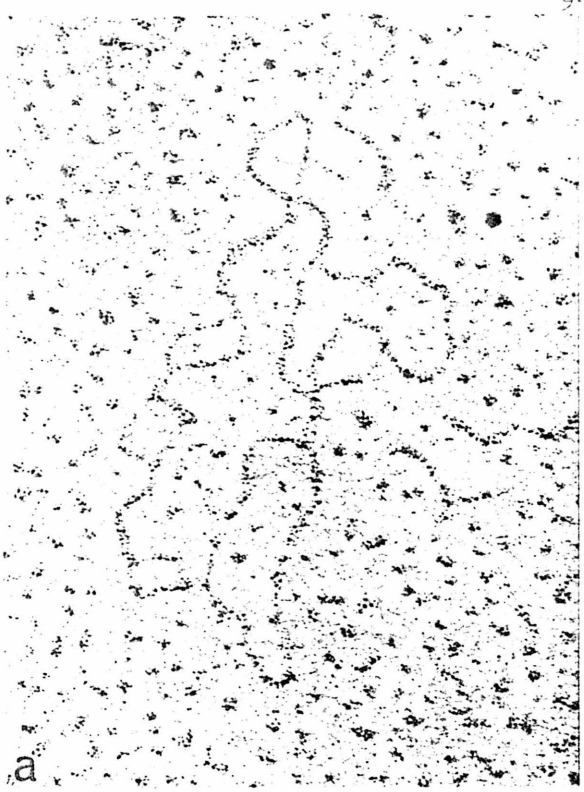
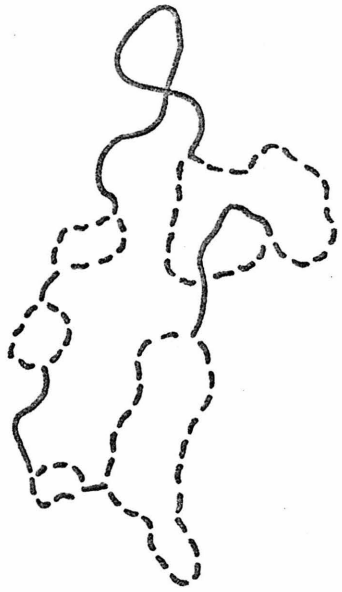
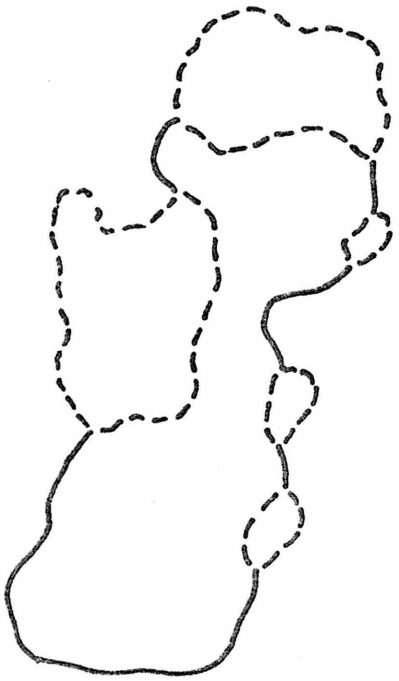


PLATE IV: ØX174/S13 heteroduplexes spread from different formamide concentrations. (a) A type IV molecule from 30% formamide. (b) A type IV molecule from 40% formamide (standard conditions).



a





b

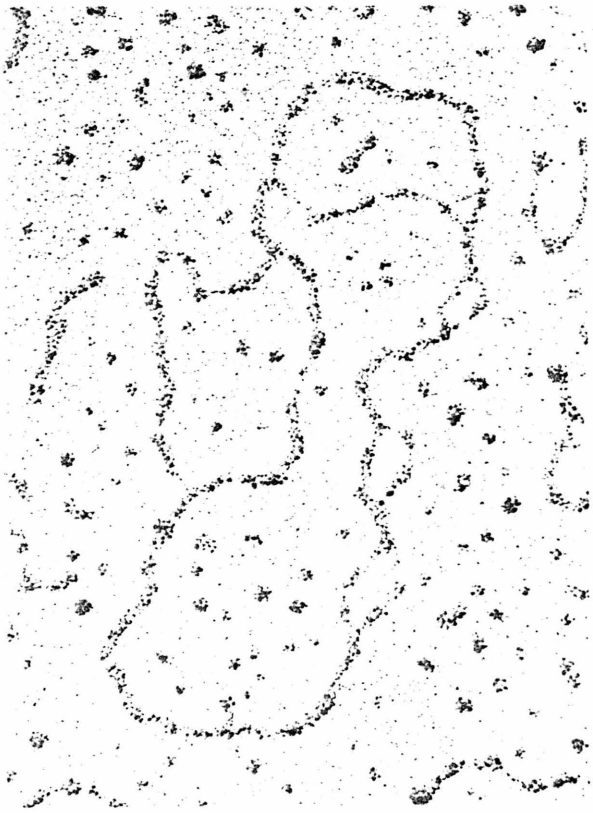
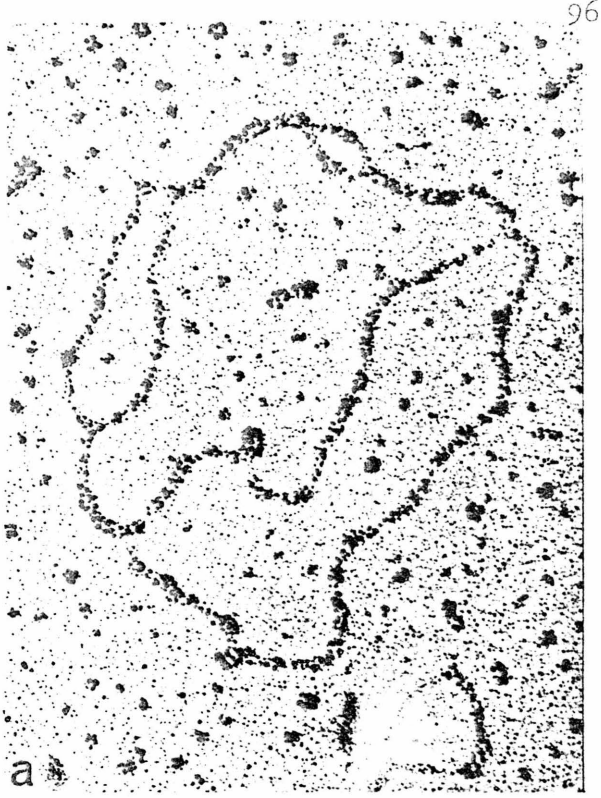
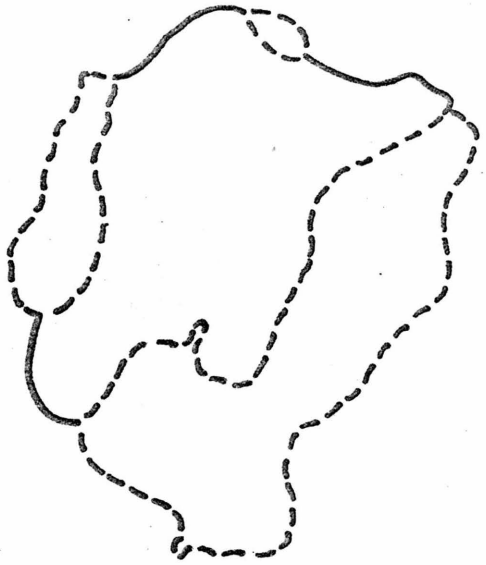
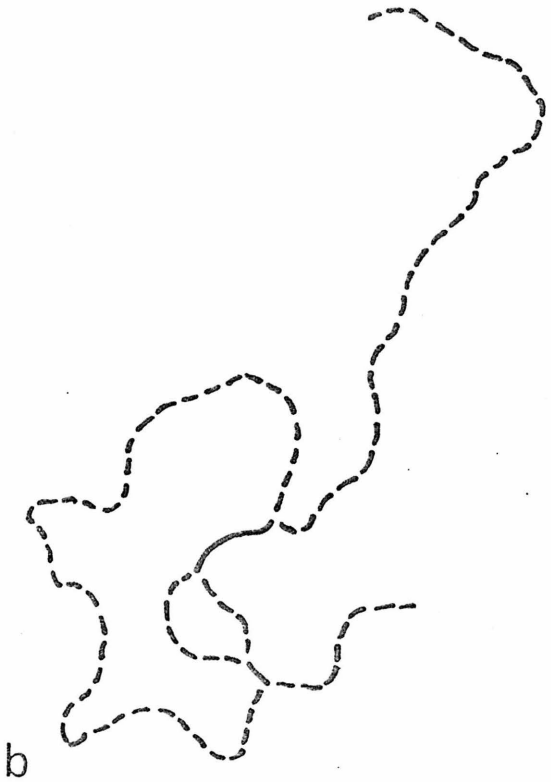


PLATE V: Heteroduplexes spread from different formamide concentrations (continued). (a) A molecule from 50% formamide in which the duplex region between L1 and L2 has melted to form one large loop. The small loop at the top is in the long duplex region between L2 and L3. (b) A molecule from 60% formamide showing the two conserved duplex regions in the right hand side of the circular strand.


a




b

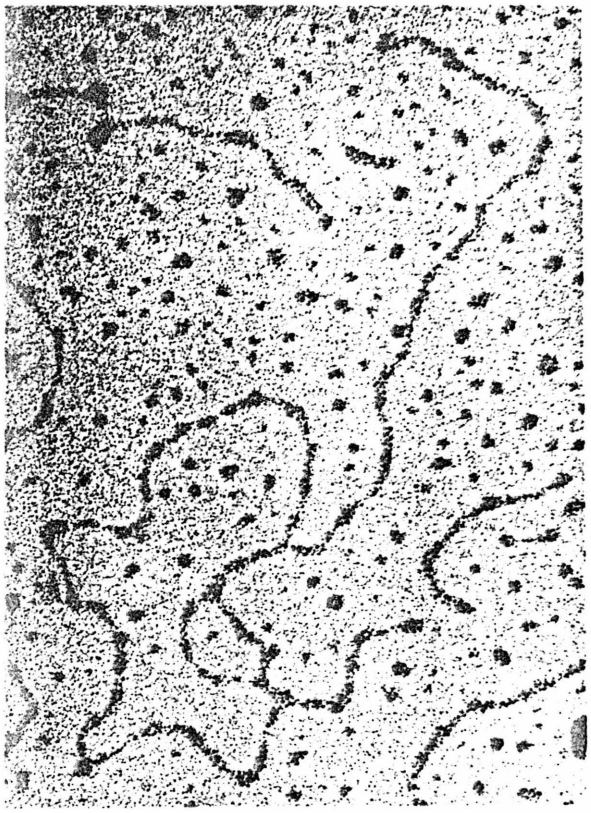
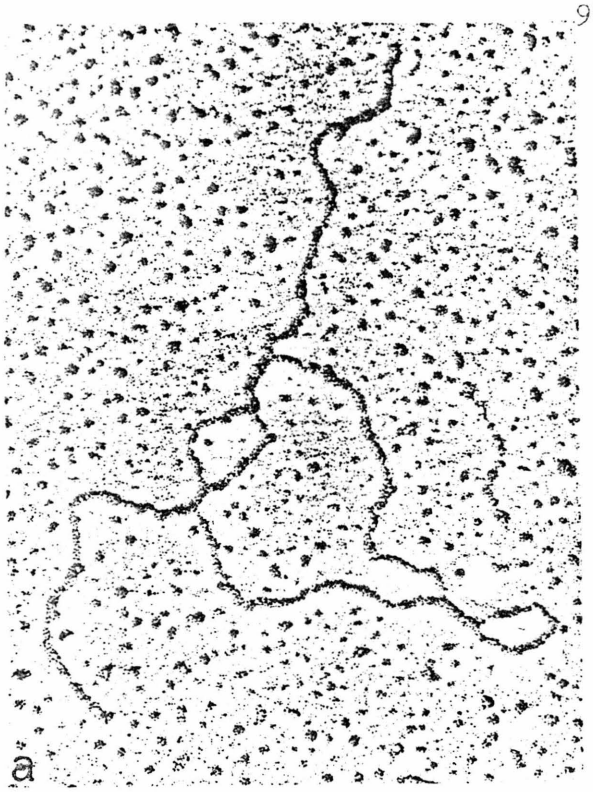
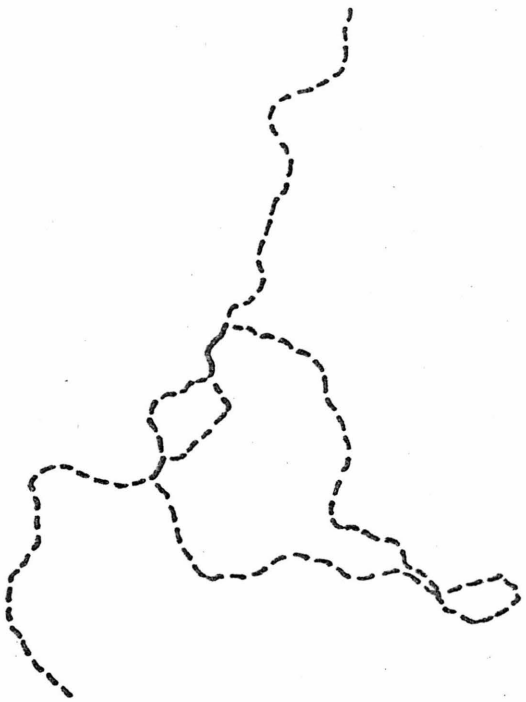
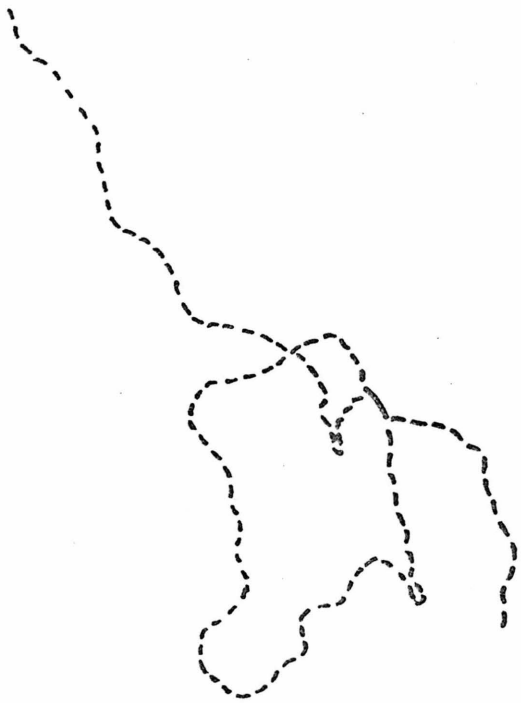


PLATE VI: Heteroduplexes spread from different formamide concentrations (continued). Molecules spread from 80% formamide showing (a) two duplex regions and (b) one duplex region.



a





b

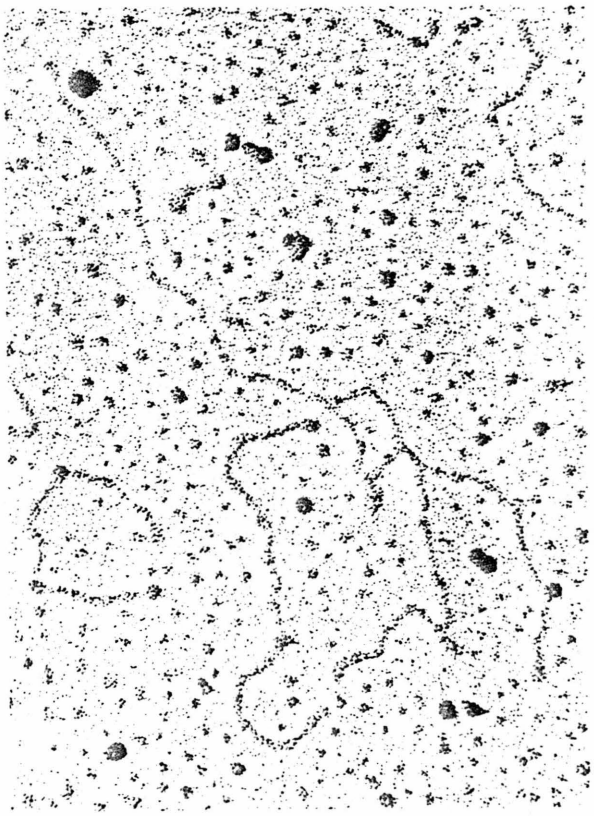
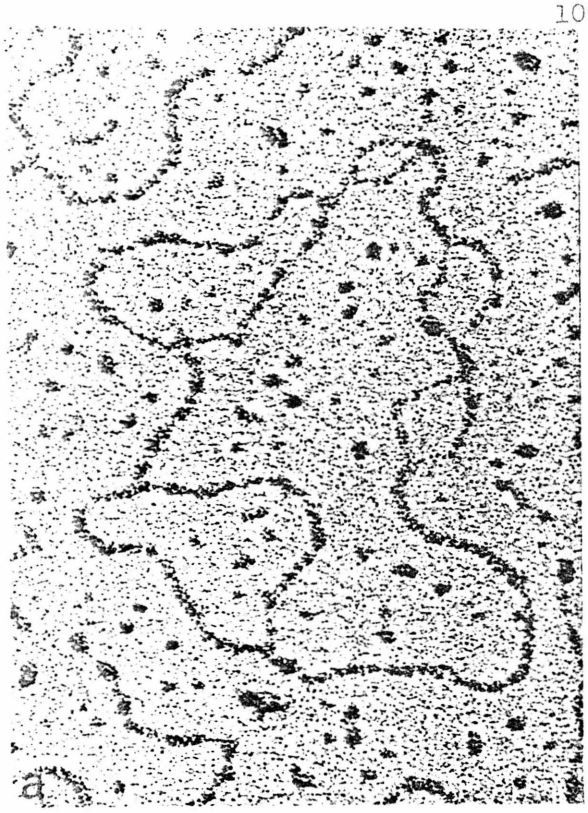
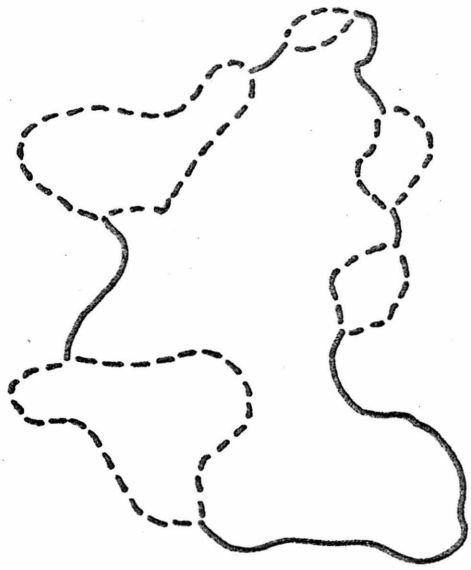
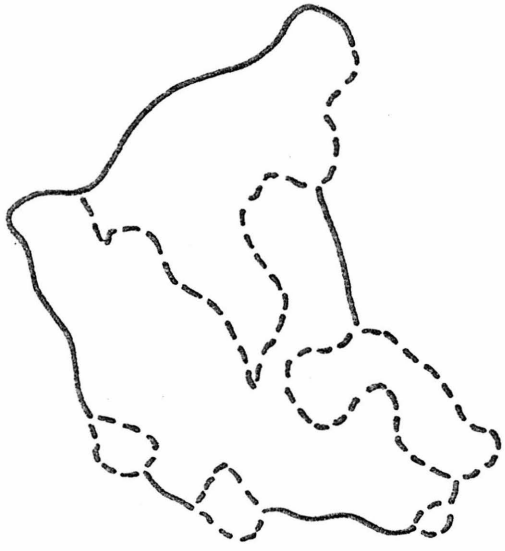


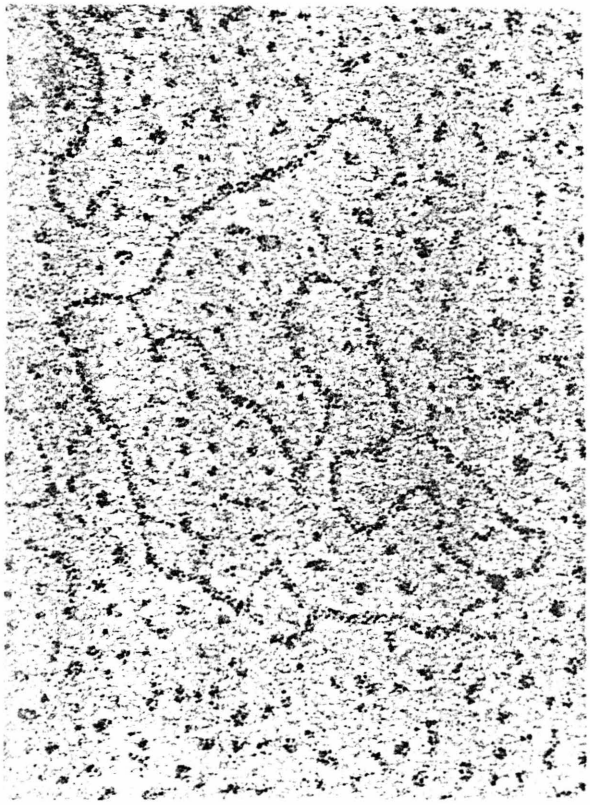
PLATE VII: ØX174/S13 heteroduplexes containing inserted R fragments. (a) The same type IV molecule as shown in Plate IIb, reproduced here for comparison. (b) A heteroduplex containing R1. L1 is in the lower right hand corner and the enlarged L3 is at the top. The fragment can be seen as a double-stranded region in the upper arm of L3 and is indicated by a solid line in the drawing.



a



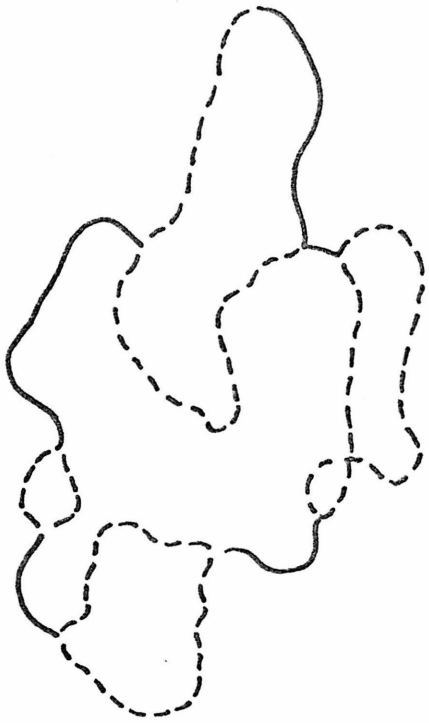


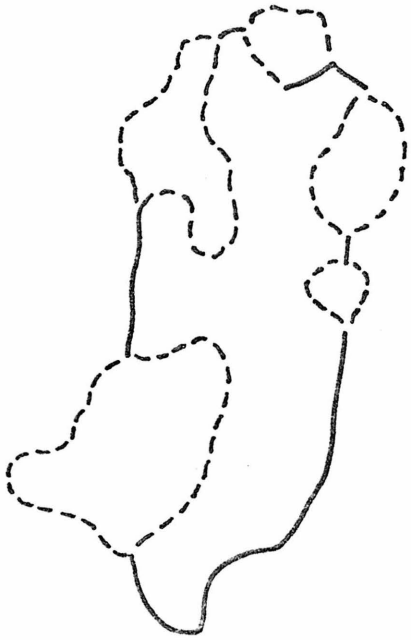


b

PLATE VIII: Heteroduplexes containing R fragments (continued). (a) A heteroduplex containing R2. L3 is at the top of the molecule. The fragment can be seen in the upper arm of L3 and is indicated by a solid line in the drawing. (b) A heteroduplex containing R8. The enlarged loop in the region between L1 and L2 is at the top of the molecule. The drawing indicates the probable position of the fragment, but it is difficult to discriminate between single- and double-stranded DNA for such a small segment.

a





b