Proc. Natl. Acad. Sci. USA
Vol. 73, No. 8, pp. 2649–2653, August 1976
Biochemistry

Psoralen-crosslinked secondary structure map of single-stranded virus DNA

(single-stranded bacteriophage fd DNA/photochemical crosslinking/electron microscopy/bacteriophage fd DNA hairpin map/promoters)

CHE-KUN JAMES SHEN AND JOHN E. HEARST

Department of Chemistry, University of California, Berkeley, Calif. 94720

Communicated by Melvin Calvin, May 21, 1976

ABSTRACT **The photochemical crosslinking of DNA by 4,5′,8-trimethylpsoralen (trioxsalen) has been used to freeze the secondary structures of single-stranded DNA molecules of bacteriophage fd at different ionic strengths. These secondary structures (hairpins or looped hairpins) have been visualized in the electron microscope. Most of the single-stranded circular fd DNA molecules show only one hairpin after irradiation at 15° in 20 mM NaCl in the presence of trioxsalen. As the ionic strength is increased, more hairpins appear on the DNA molecules. To map these secondary structures, double-stranded supercoiled fd DNA (RFI) was cleaved with the restriction enzyme *Hin*dII, which makes only one cut on each RFI molecule. After denaturation and crosslinking of the linear single-stranded fd DNA (a mixture of viral and complementary strands), all the hairpins have been mapped on the DNA molecule with respect to this *Hin*dII site. The results show that these hairpins occur at specific sites. The most stable hairpin has been assigned to the position where the initiation site for the conversion of single-stranded fd DNA to the double-stranded covalently closed form has been mapped. The remaining hairpins map in or near regions corresponding to *in vitro* promoter sites on the fd DNA.**

The filamentous bacteriophage fd contains a single-stranded circular DNA (1) with a molecular weight of about 2×10^6 (2, 3). Upon infection of *Escherichia coli* bacteria by fd, the single-stranded (SS) viral DNA initiates the synthesis of the complementary strand and double-stranded circular DNA (RF) is formed. The RF DNA then undergoes several rounds of semiconservative replications, producing more RF. In the final stage of DNA synthesis the SS fd DNA is formed and packaged, producing mature virus (for a review, see ref. 4).

We have developed a method for visualizing secondary structure of single-stranded DNA in the electron microscope by first stabilizing these regions of secondary structure using the photocrosslinking reaction between DNA and a Psoralen derivative. The Psoralen secondary structure maps of fd DNA are of particular interest because Schaller *et al.* (5) have observed hairpin structures in fd which are resistant to digestion by single-strand-specific nuclease. In particular Schaller *et al.* (6) have isolated one very stable RNA-polymerase-binding hairpin fragment (ori-DNA) from fd DNA, which they have mapped on the viral genome at the site of initiation of synthesis of the complementary DNA strand in the step where the SS fd DNA is converted to RF DNA (SS fd → RF fd). We have observed and assigned a hairpin in SS fd DNA by the Psoralen crosslinking technique to the same site. We find this particular hairpin to be the most stable of the several hairpin structures which SS fd DNA forms. Because fd DNA is initially in infected cells in a SS form, we expect the stability and position of sites of secondary structures in the DNA to have importance as sites of binding of double-strand-specific enzymes such as RNA polymerase. The comparison of the physical map of hairpins and the fd genetic map suggests that secondary structure in the single stranded DNA has an important biological function.

The most commonly used method for examining single-stranded DNA by electron microscopy is the formamide spreading technique (7). However, it has not been possible to control the cytochrome *c* spreading conditions for good DNA extension and good contrast and simultaneously stabilize the secondary structure of single-stranded DNA. In order to crosslink these hairpins, the SS fd DNA has been irradiated with long wavelength UV light in the presence of 4,5′,8-trimethylpsoralen (trioxsalen) at a particular temperature and ionic strength of interest. Trioxsalen and other Psoralen derivatives are known to covalently crosslink pyrimidines on opposite DNA strands (8–10). The DNA can then be spread for electron microscopy at standard isodenaturing conditions for DNA and the crosslinked hairpins can be observed. We have observed as many as eight hairpins in fd DNA by this method.

In order to map these hairpins on the fd genome, the supercoiled replicative form fd DNA (RFI) has been cleaved with the restriction enzyme *Hin*dII. This enzyme recognizes the sequence

5′-G-T-Y-R-A-C-3′
3′-C-A-R-Y-T-G-5′

where Y is pyrimidine and R is purine. All the observed hairpins are located at specific sites on the fd DNA molecule and their positions show some correlation with previously mapped *in vitro* promoter sites on the RFI DNA.

MATERIALS AND METHODS

Bacteriophage fd DNA. The SS fd DNA was a gift from Leroy F. Liu and has been isolated according to standard procedures (1). The supercoiled fd DNA (RFI DNA) was purified by Dr. J. Jacobsen as described by Sugiura *et al.* (13). Both DNAs were kept at 4° in 1 mM Tris·HCl, 0.1 mM EDTA, pH 7.0 (TE buffer) before use.

Digestion of RFI by the Restriction Enzyme *Hin*dII. fd RFI DNA was incubated with the restriction enzyme *Hin*dII at 37° under the following conditions: 10 mM $MgCl_2$, 5 mM NaCl, 6 mM Tris, 6 mM EDTA (pH 8.0). After 4 hr of digestion, the reaction mixture was made 50 mM in EDTA (pH 7.0) and extracted with buffer (100 mM Tris, 10 mM EDTA, pH 7.0) saturated phenol. The residual phenol was extracted with ether, and the aqueous phase containing the restricted RFI (*Hin*dII-RFI) was dialyzed exhaustively against TE buffer.

Gel Electrophoresis of DNA. DNA samples including the supercoiled RFI and *Hin*dII-RFI were analyzed in a 1% agarose

Abbreviations: SS fd DNA, single-stranded (+) strand of fd DNA; SS (±) fd DNA, mixture of single-stranded (+) and (−) strands of fd DNA; RFI DNA, covalently closed double-stranded circular DNA; SS (±) *Hin*dII-RFI DNA, the single-stranded linear DNA obtained by denaturation of RFI circles that have been cleaved once at the site of the restriction endonuclease *Hin*dII.

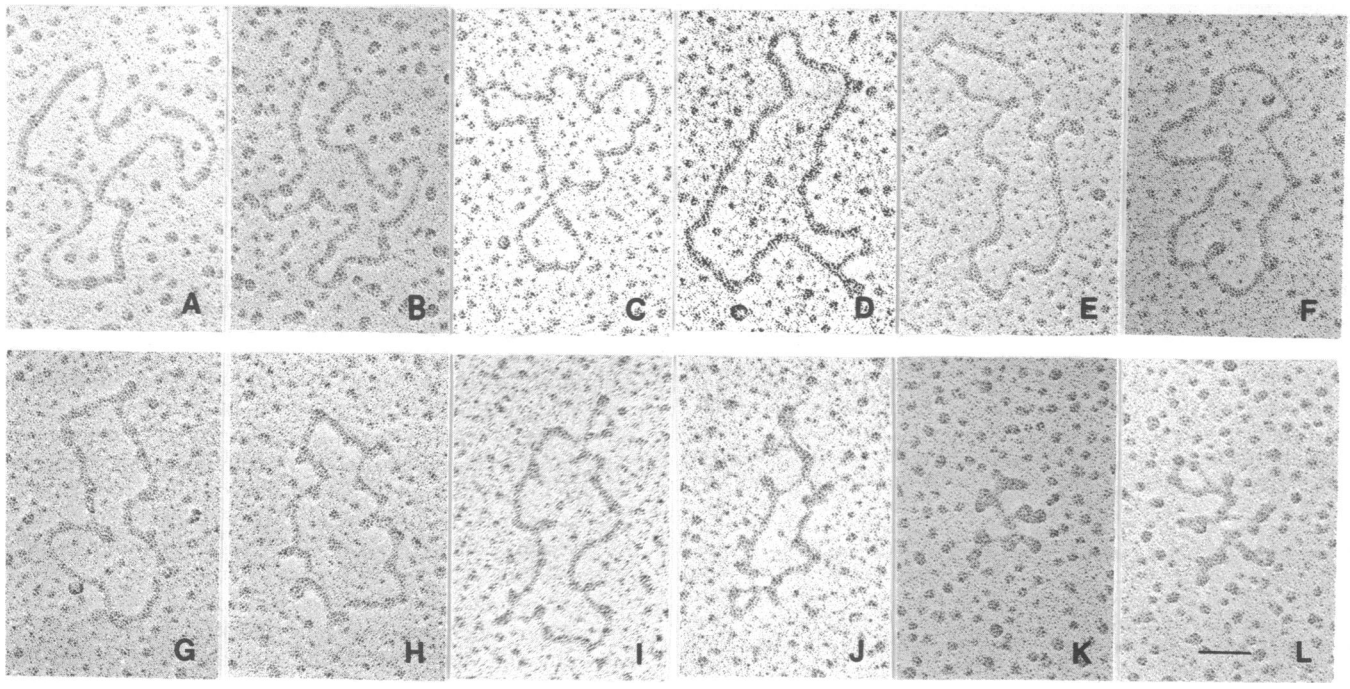

FIG. 1. Secondary structures of single-stranded circular fd DNA. SS circular fd DNA was spread for electron microscopy after photocrosslinking with trioxsalen at 15° and different NaCl concentrations. (A) 0 mM, (B) 10 mM, (C, D) 20 mM, (E, F) 25 mM, (G, H) 30 mM, (I, J) 40 mM, (K, L) 100 mM. The bar in (L) stands for 0.1 μm. Magnification ×72,300.

slab gel as described by Sharp *et al.* (14) and by Shen *et al.* (15).

Preparation of Samples for the Crosslinking Reaction. Concentrated NaCl solutions (0.1 M–1 M) were added directly to the SS circular fd DNA in TE buffer to make the desired final salt concentration. The double-stranded *Hin*dII-RFI DNA was made 0.1 M in NaOH and alkali denatured at room temperature for 10 min. Concentrated HCl solution (5 M) was then added to neutralize it and the sample was dialyzed at 4° against TE buffer. Before the irradiation, the mixture of (+) (viral) and (−) single-stranded DNA, termed SS (±), *Hin*dII-RFI DNA was adjusted to a certain ionic strength by adding concentrated NaCl solution.

Light Source for Irradiation. Long wavelength UV light from two General Electric F15T8 BLB tubes was used in all the photochemical reactions between the DNA and the drug. The incident intensity was 1.0 mW/ml of sample volume as determined by ferrioxalate actinometry (16). The peak irradiation wavelength was determined to be about 365 nm. Light with wavelength below 350 nm was filtered out by a Plexiglas filter.

Drug–DNA Photochemical Reaction. Trioxsalen (Paul B. Elder Co.) was dissolved in 100% ethanol to a final concentration of 1 mg/ml. One microliter of this stock solution was added to 50 μl of 5 μg/ml of DNA solution which had been adjusted to an appropriate ionic strength as described above. Each sample was put into a small glass pipette, sealed, and irradiated at 15° with constant rocking of the pipette. After 2 hr of irradiation, 1 μl of trioxsalen stock solution was added to each sample again and irradiation was continued for another 2 hr.

Electron Microscopy. The Kleinschmidt spreading technique (17) modified by Davis *et al.* (7) was used to visualize the fd DNA molecules in an electron microscope. The hyperphase contained 45% (vol/vol) formamide (M. C. & B. Manufacturing Chemists), 0.1 M Tris, 0.01 M EDTA (pH 8.2), 0.25 μg/ml of DNA, and 40 μg/ml of cytochrome *c*. The hypophase contained 17% formamide, 0.01 M Tris, 0.001 M EDTA (pH 8.5). Samples are picked up with parlodion-coated copper grids, stained with 50 μM uranyl acetate in 90% ethanol, and shadowed with 80% Pt–20% Pd at an angle of 8°. The grids were examined in a Philips 201 electron microscope. The DNA molecules were photographed, the negatives were enlarged by a 500 RR-2 projector (Standard Projector & Equipment Co.), and the contour lengths of the DNA were measured with a planimeter (Numonics Co.).

RESULTS

The Secondary Structures of the SS Circular fd DNA. When SS circular fd DNA is crosslinked and spread, the electronmicrographs show different conformations of DNA depending on the salt conditions used during the irradiation. DNA molecules crosslinked at 15° and salt concentrations lower than 15 mM NaCl appear as smooth circles (Fig. 1A and B). This conformation of fd DNA is maintained for most of the DNA up to a salt concentration of 15 mM NaCl, although a few molecules begin to show evidence of hairpins at this salt condition. After crosslinking fd DNA at 15° in 20 mM NaCl, more than 90% of the molecules have one single hairpin that is approximately 150 base pairs long (Fig. 1C). Some of the DNA molecules at 20 mM NaCl have two or three hairpins (Fig. 1D). As the salt concentration is increased to 25 mM and 30 mM NaCl (Fig. 1E–H), more hairpins show up on each circular fd DNA. Most of the DNA molecules that have been crosslinked at 15° and 30 mM NaCl show four to five hairpins. At 40 mM NaCl concentration, the hairpin pattern becomes more complex (Fig. 1I) and sometimes doublets of hairpins can be found (Fig. 1J). Finally, all of the DNA molecules crosslinked at 100 mM NaCl or higher salt show collapsed, tree-like structures (Fig. 1K and L).

The Secondary Structures on SS (±) *Hin*dII-RFI DNA. The hairpins on the crosslinked SS circular fd DNA seem to have some specificity with respect to their relative positions. In order to confirm this point, supercoiled fd RFI DNA was cut by the restriction enzyme *Hin*dII. The restriction product appears as

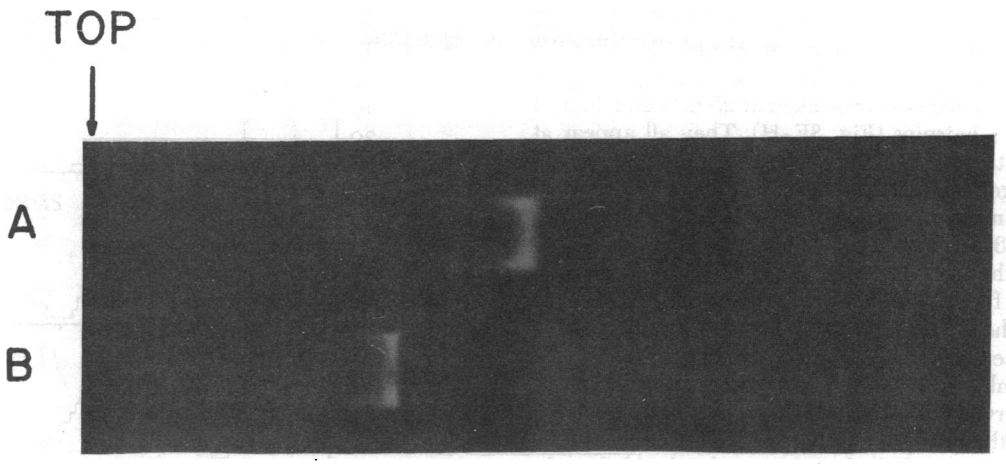

FIG. 2. Gel electrophoresis of double-stranded fd DNA. DNA (0.1 μg) was loaded into each well and electrophoresed in 1% agarose gel at 40 V for 3 hr. (A) fd RFI DNA, (B) *Hin*dII-RFI DNA.

a single band in a 1% agarose gel with molecular weight (3.8 ± 0.2) × 10^6 (Fig. 2, column B). The enzyme *Hin*dII cleaves fd RF DNA once and has been used to establish the gene map and the *in vitro* promoter map of fd DNA (18).

The linear *Hin*dII-RFI DNA was then denatured, neutralized, and crosslinked with trioxsalen at 15°. Typical examples of DNA molecules that have been crosslinked at 0, 10, 20, 25, 30, 40, and 100 mM NaCl concentration are shown in Fig. 3. As can be seen, the general patterns of the hairpins on these linear DNA molecules are very similar to those on the SS circular fd DNA that was crosslinked at corresponding salt conditions. The linear SS (±) *Hin*dII-RFI DNA shows essentially

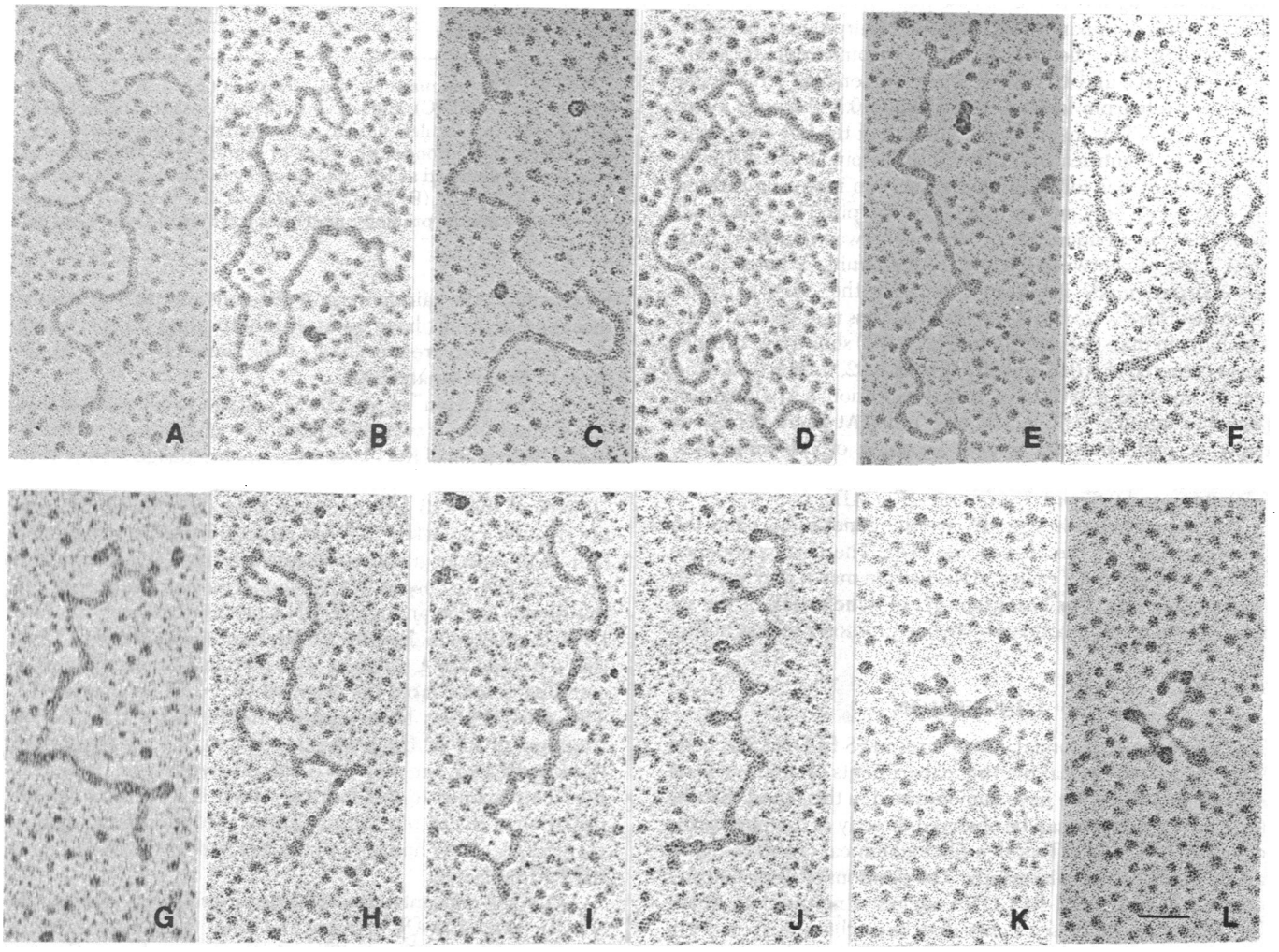

FIG. 3. Secondary structures of single-stranded (±) *Hin*dII-RFI DNA. fd RFI DNA was cut by the restriction enzyme *Hin*dII, denatured, neutralized, and crosslinked at 15° in different NaCl concentrations. (A) 0 mM, (B) 10 mM, (C, D) 20 mM, (E, F) 25 mM, (G, H) 30 mM, (I, J) 40 mM, (K, L) 100 mM. Magnification ×72,300.

no observable hairpins on molecules crosslinked at low salt (below 15 mM NaCl) (Fig. 3A and B). At 20 mM NaCl (Fig. 3C and D), one hairpin has been found on almost all the molecules examined. DNA molecules crosslinked at 25 mM and 30 mM NaCl show more hairpins (Fig. 3E–H). They all appear at specific sites with respect to the "first hairpin" and can be mapped on the genome (see below). It is interesting to note that over half of the molecules crosslinked at 30 mM NaCl and 40 mM NaCl (Fig. 3G–J) have a small hairpin or loop (150–300 Å in length) near the end where the "first hairpin" is located. As with SS circular fd DNA, doublets also appear on the linear SS (±) fd DNA when they are crosslinked at 40 mM NaCl or higher salt concentrations. At 100 mM NaCl, the DNA molecules are crosslinked to a much greater extent but the hairpins can still be observed (Fig. 3K and L).

Mapping of the Hairpins on the fd Genome. An attempt has been made to establish the hairpin map for the fd. Samples of 17 to 23 molecules of the crosslinked SS (±) *Hin*dII-RFI DNA at each of the salt concentrations have been photographed and the lengths of all the hairpins along with the center-to-center distances of adjacent hairpins have been measured. The fd genome was given a unit length of one with the left end (Fig. 4) chosen as the origin of the map. The "first hairpin" that appears at 20 mM NaCl and 15° has been assigned to the left part of the molecule (Fig. 4A). It has been found that the center of the "first hairpin" is located at the position of 0.14 ± 0.03 map units and its length is 167 ± 70 base pairs. At 25 mM NaCl, this hairpin is still dominant so that the other hairpins can be easily mapped (Fig. 4B). The hairpin pattern becomes more complex at 30 mM NaCl and 15°. Sometimes both ends of a molecule contained a hairpin located in the region 0.1–0.2 units away from the end. This made the assignment of the left end of the molecule by identification of the "first hairpin" more difficult. This problem was solved by noting that in many cases those molecules with only one hairpin ("first hairpin") had the small looped hairpin mentioned above located between 0 and 0.1 map units, a hairpin between 0.3 and 0.4 map units, or both. These two hairpins made it easier to identify the "first hairpin" whenever there were two symmetric hairpins near the two ends. The hairpin map thus established (Fig. 4C) shows five distinct peaks centered at 0.046 ± 0.009, 0.16 ± 0.02, 0.36 ± 0.02, 0.54 ± 0.04, and 0.81 ± 0.05 map units. In addition, a small hairpin peak was found at 0.94 ± 0.03 map units. At 40 mM NaCl, as many as eight hairpins appear on some of the molecules. Doublets of hairpins are often found in three regions, 0.25–0.40, 0.50–0.60, and 0.65–0.90 map units (Fig. 4D).

The maps shown in Fig. 4E are the genetic map and the promoter map of the fd genome (18). The correspondence between the promoter map and the hairpin map is striking. This is especially clear for histogram 4C in which each of the six hairpin peaks corresponds well to one promoter region in Fig. 4E.

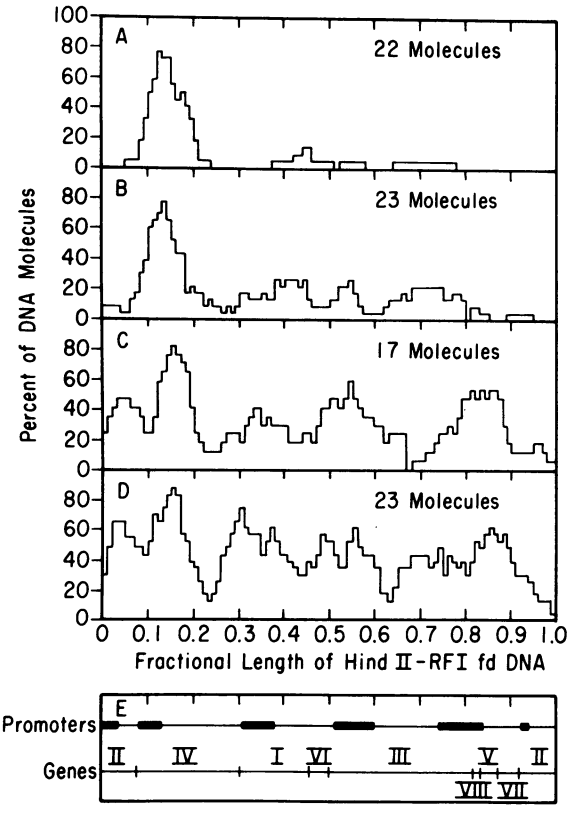

FIG. 4. Hairpin maps of single-stranded (±) fd DNA. The histograms in (A), (B), (C), and (D) show the percentage of SS (±) *Hin*dII-RFI DNA molecules that have a hairpin in a given interval corresponding to one hundredth of the fd genome size. DNA samples have been crosslinked at 15° and one of the following NaCl concentrations: (A) 20 mM, (B) 25 mM, (C) 30 mM, (D) 40 mM. The maps shown in (E) are the promoter map and gene map of the fd virus genome (18).

DISCUSSION

The existence of secondary structures in SS fd DNA has been suggested from optical melting experiments (4) and nuclease digestion data (5). We have demonstrated in this report that the secondary structures can be observed directly as hairpins by the combined use of the Psoralen photochemical crosslinking reaction and electron microscopy. The positions of these hairpins on the SS (±) *Hin*dII-RFI DNA suggest a nonrandom distribution of the DNA segments that have the ability to form hairpins. Mismatched base pairs probably exist in these hairpins, since the melting of SS fd DNA is very broad (4) and since we have observed looped hairpins on the fd DNA in many cases. Furthermore, the hairpin maps at different ionic strengths show that some hairpins have greater stability than others, where these differences are most likely to be caused by different degrees of base mismatching.

The "first hairpin" or most stable hairpin, which maps at 0.14 ± 0.03 map units, is very interesting. The single strand to double strand replication of the filamentous bacteriophage M13 (a virus closely related to fd) has been found to be initiated in a region 10–20% of the M13 genome away from the *Hin*dII cleavage site (19) and a model for the M13 replication has been suggested (20). In this model, a unique hairpin is formed at a specific site on the SS DNA genome. RNA polymerase binds selectively to this duplex region and makes RNA primer that is essential for DNA chain growth. Recently, it has been shown for SS fd DNA that a unique DNA segment with double-stranded character could be isolated from the origin of SS → RF replication (6). This ori-DNA has been mapped in the *Hpa*-H restriction fragment of RF fd DNA which is 0.08–0.14 units away from the *Hin*dII cleavage point. It is very likely that our "first hairpin" corresponds to this hairpin which is important for the initiation of the fd replication. This is the basis for our assignment of the "first hairpin" to the left half of the SS (±) *Hin*dII-RFI molecule.

As the salt concentration increases, more hairpins appear at specific sites of the fd genome and all of them are located in promoter regions (Fig. 4). These hairpins may be a reflection of the general character of the *in vitro* promoter sequences. An examination of sequences of the promoters isolated from dif-

ferent sources (21–27) indicates that most of them have the ability to form hairpins. Although the average lengths of the hairpins we have observed are longer than the promoter sequences that have been isolated (100–250 base pairs versus 29–85 base pairs), the possibility cannot be excluded that the 2-fold symmetry found in the promoter sequences actually extends beyond the RNA-polymerase-protected regions. Alternatively, the promoter sites may be the only regions that have appreciable 2-fold symmetry and the smooth SS DNA seen after crosslinking at low salt (Figs. 1B and 3B) may contain nondetectable short hairpins (<50 base pairs). As the NaCl concentration increases, the bases adjacent to these short hairpins may become more easily crosslinked, extending the length of the duplex region beyond the length of the thermodynamically stable helix.

Although our experiments were performed at 15° and in 0 to 100 mM NaCl, it is conceivable that these hairpins exist under physiological conditions, since 37° is the midpoint of the melting transition of SS fd DNA in 150 mM NaCl (4). Very little is known about the number and size of the *in vivo* products specified by the phage promoters, although genetic experiments indicate a polygenic mRNA for genes III-VI-I (4). Analysis of *in vitro* fd RNA led to the detection of at least four species with unique initiation sequences and length, ATP-initiated chains being of nearly full genome length and GTP-initiated chains approximately one-third or less (28, 29). Most if not all of these hairpin regions could therefore be structural features of the viral messenger RNAs as well. They may function in the processing of the large mRNAs into smaller molecules. More importantly, the hairpins could play a role in the regulation of gene expression at the translational level, as they do in RNA bacteriophage (30).

We have observed hairpins at specific sites in SS SV40 DNA in addition to the SS fd DNA results reported here (Shen and Hearst, in preparation). We believe that, in the near future, the Psoralen photocrosslinking reaction and electron microscopy will be extremely powerful tools for freezing and directly visualizing special DNA structures associated with many genetic activities.

We thank Leroy F. Liu and Dr. John H. Jacobsen for the bacteriophage fd DNA. Dr. M. R. Botchan has kindly provided us the restriction enzyme *Hin*dII. The actinometric determination of lamp intensity was done by Dr. Carl V. Hanson and John Sasaki. C.K.J.S. is grateful for Dr. Thomas R. Cech's patience in teaching him the Kleinschmidt spreading method and other biochemical techniques 2 years ago. This study was supported by National Institutes of Health Grant GM-11180 and American Cancer Society Grant NP-185. C.K.J.S. has been supported by the Earle C. Anthony University Fellowship.

1. Marvin, D. A. & Schaller, H. (1966) *J. Mol. Biol.* **15**, 1–7.
2. Marvin, D. A. & Hoffmann-Berling, H. (1963) *Nature* **197**, 517–518.
3. Berkowitz, S. A. & Day, L. A. (1974) *Biochemistry* **13**, 4825–4831.
4. Marvin, D. A. & Hohn, B. (1969) *Bacteriol. Rev.* **33**, 172–209.
5. Schaller, H., Voss, H. & Gucker, S. (1969) *J. Mol. Biol.* **44**, 445–458.
6. Schaller, H., Uhlmann, A. & Geider, K. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 49–53.
7. Davis, R. W., Simon, M. & Davison, N. (1971) in *Methods in Enzymology*, eds. Grossman, L. & Moldave, K. (Academic Press, New York), Vol. 21, pp. 413–428.
8. Cole, R. S. (1971) *Biochim. Biophys. Acta* **254**, 30–39.
9. Dall'acqua, F., Marciani, S., Vedaldi, D. & Rodighiero, G. (1974) *Biochim. Biophys. Acta* **353**, 267–273.
10. Hanson, C. V., Shen, C. K. J. & Hearst, J. E. (1976) *Science*, in press.
11. Kelly, T. J. & Smith, H. O. (1970) *J. Mol. Biol.* **51**, 393–409.
12. Takanami, M. (1973) *FEBS Lett.* **34**, 318–322.
13. Sugiura, M., Okamoto, T. & Takanami, M. (1969) *J. Mol. Biol.* **43**, 299–315.
14. Sharp, P. A., Sugden, B. & Sambrook, J. (1973) *Biochemistry* **12**, 3055–3063.
15. Shen, C. K. J., Wiesehahn, G. & Hearst, J. E. (1976) *Nucleic Acids Res.* **3**, 931–952.
16. Hatchard, C. G. & Parker, C. A. (1956) *Proc. R. Soc. London* **235**, 518–536.
17. Kleinschmidt, A. K. (1968) in *Methods in Enzymology*, eds. Grossman, L. & Moldave, K. (Academic Press, New York), Vol. 12B, pp. 361–377.
18. Seeburg, P. H. & Schaller, H. (1975) *J. Mol. Biol.* **92**, 261–277.
19. Tabak, H., Griffith, J., Geider, K., Schaller, H. & Kornberg, A. (1974) *J. Biol. Chem.* **249**, 3049–3054.
20. Geider, K. & Kornberg, A. (1974) *J. Biol. Chem.* **249**, 3999–4005.
21. Zain, B. S., Weissman, S. M., Dhar, R. & Pan, J. (1974) *Nucleic Acids Res.* **1**, 577–594.
22. Dhar, R., Weissman, S. M., Zain, B. S., Pan, J. & Lewis, A. M. Jr. (1974) *Nucleic Acids Res.* **1**, 595–614.
23. Maniatis, T., Ptashne, M., Barrell, B. G. & Donelson, J. E. (1974) *Nature* **250**, 394–397.
24. Dickson, R. C., Abelson, J., Barnes, W. M. & Reznikoff, W. S. (1975) *Science* **187**, 27–35.
25. Schaller, H., Gray, C. & Herrmann, K. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 737–741.
26. Sekiya, T., Ormondt, H. V. & Khorana, H. G. (1975) *J. Biol. Chem.* **250**, 1087–1098.
27. Pribnow, D. (1975) *J. Mol. Biol.* **99**, 419–443.
28. Okamoto, T., Suguira, M. & Takanami, M. (1969) *J. Mol. Biol.* **45**, 101–111.
29. Takanami, M., Okamoto, T. & Suguira, M. (1970) *Cold Spring Harbor Symp. Quant. Biol.* **35**, 179–187.
30. Kozak, M. & Nathans, D. (1972) *Bacteriol. Rev.* **36**, 109–134.