Neutron Scattering Data on Reconstituted Complexes of fd Deoxyribonucleic Acid and Gene 5 Protein Show That the Deoxyribonucleic Acid Is near the Center†

Donald M. Gray,* Carla W. Gray, and R. Douglas Carlson

ABSTRACT: We have performed low-angle neutron scattering studies on reconstituted complexes of fd DNA and the gene 5 protein that is produced during infection of *Escherichia coli* by filamentous fd phage. Essentially identical helical complexes have been made with normal protonated DNA or DNA in which at least 87% of the nonexchangeable protons are replaced by deuterium. From neutron scattering profiles of both complexes over a range of D_2O/H_2O solvent mixtures, the DNA deuteration is shown to have a dramatic influence on the measured cross-sectional radius of gyration. Most importantly, data for the complex containing deuterated DNA lead to a more negative slope in a plot of the square of the cross-sectional radius of gyration vs. the inverse of the solute–solvent contrast, compared with the slope of a plot of data for the complex containing protonated DNA. This means that, in a cross-sectional view of the complex, the DNA is near the center of the structure. By our analysis, the DNA has a cross-sectional radius of gyration of 17.6 ± 3 Å, while the protein has a cross-sectional radius of gyration of about 33.5 Å. Therefore, the model for the structure of the helical complex that has been proposed from X-ray diffraction studies on gene 5 protein crystallized with oligodeoxynucleotides [McPherson, A., Jurnak, F., Wang, A., Kolpak, F., Rich, A., Molineux, I., & Fitzgerald, P. (1980) *Biophys. J. 32*, 155–170] is not valid for the complex in solution. From our neutron diffraction data we have also obtained values for the solvent-excluded volume and mass per unit length. The relation of our findings to the solution structure of the complex is discussed.

In the course of infection of *Escherichia coli* cells by the filamentous phage fd, an intracellular filamentous complex is formed between progeny single-stranded fd DNA molecules and the protein product of fd gene 5 (Alberts et al., 1972; Oey & Knippers, 1972). The role of the gene 5 DNA-binding protein in the control of fd phage replication has been reviewed by Ray (1978), and the morphogenesis to form mature phage particles from the precursor complex of fd DNA and gene 5 protein, which involves an exchange of gene 5 protein molecules for coat protein molecules at the cellular membrane, has been reviewed by Webster & Cashman (1978). The structure of the complex between gene 5 protein and fd DNA is particularly intriguing, since dimers of the protein act to coalesce the circular single-stranded DNA into a linear rod (Alberts et al., 1972). The helical nature of the structure has been confirmed by electron microscopy and optical diffraction studies (Gray et al., 1982; C. W. Gray, unpublished results).

The 87 amino acid sequence of the gene 5 protein is known (Nakashima et al., 1974), and the three-dimensional structure at 2.3-Å resolution of the crystallized protein has been determined by X-ray diffraction (McPherson et al., 1979a). Crystallographic studies of gene 5 protein complexed with oligodeoxynucleotides have shown that the protein complexed with oligomers forms crystals with 12 monomers/asymmetric unit in what is probably a hexameric ring of six protein dimers (McPherson et al., 1979b, 1980a,b). On the basis of the hexameric ring structure and a tentative orientation of the DNA-binding sites on the outer surface of the hexameric ring, a model has been proposed by McPherson et al. (1979b, 1980a) for the structure of long complex helices. In this model, both the "up" and "down" antiparallel DNA strands are located on the outer surface of the helix, one turn of which contains six protein dimers.

The solution structure of the fd DNA–gene 5 protein complex has been studied by using the technique of small-angle neutron scattering (Torbet el al., 1981). In that study it was found that the neutron scattering data for the complex in solution were in reasonable agreement with the model proposed by McPherson et al. (1979b, 1980a), with the important exception of the DNA position. Both the contrast dependence of the cross-sectional radius of gyration and the calculated maximum radius of the DNA backbone indicated that the DNA is not on the outer periphery of the helical complex. However, since DNA comprises only about 12% of the total mass in the complex, an unambiguous analysis of the DNA position is difficult from neutron scattering data without a selective enhancement of the scattering from the DNA.

We have now succeeded in reconstituting essentially identical complexes containing either heavily deuterated or normal protonated DNA. The enhanced neutron scattering length of the deuterated DNA significantly affects the radius of gyration measured at various solute–solvent contrasts and has allowed a determination of the relative radii of gyration of the DNA and protein components. In this paper we present our characterization of these reconstituted complexes and the resulting neutron scattering data obtained at five different contrasts for each. Our analysis shows that the DNA is in the central region of the complex with gene 5 protein and that the protein in the complex could not be oriented with its DNA binding sites on the exterior as proposed by McPherson et al. (1979b, 1980a). Thus, the DNA need not undergo a marked change in radial position as gene 5 protein molecules are

† From the Program in Molecular Biology, The University of Texas at Dallas, Richardson, Texas 75080 (D.M.G. and C.W.G.), and the Department of Biology, Brookhaven National Laboratory, Upton, New York 11973 (R.D.C.). *Received July 6, 1981; revised manuscript received November 17, 1981.* This work was supported by National Institutes of Health Grant GM-19060 and Grant AT-503 from the Robert A. Welch Foundation (D.M.G.) and by funds awarded under National Institutes of Health Biomedical Research Support Grant S07 RR07133 (C.W.G.). Research carried out at Brookhaven National Laboratory was under the auspices of the U.S. Department of Energy. Acknowledgment is also gratefully made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for partial support of this work.

0006-2960/82/0421-2702$01.25/0

replaced by coat protein molecules during viral morphogenesis.

Materials and Methods

Isolation of Gene 5 Protein. The gene 5 protein was isolated from about 70 g of infected cells by using essentially the procedure described by Day (1973). Fractions were eluted from the DNA–cellulose column with 0.2, 0.4, 0.6, and 0.8 M NaCl; the first portion of the fraction eluted with 0.6 M NaCl contained the purest gene 5 protein as indicated by its ultraviolet (UV) absorption spectrum. The fraction eluted at 0.4 M NaCl and a portion of the fraction eluted at 0.6 M NaCl were dialyzed to remove NaCl and reapplied to the DNA–cellulose column, and again fractions that eluted at 0.6 M NaCl contained the purest gene 5 protein. The major portions of these fractions were dialyzed at 4 °C for 2 days against 0.01 M Tris-HCl,[1] pH 6.9, and 10^{-4} M β-mercaptoethanol, concentrated 5–11-fold by air-drying, and dialyzed for another 3 days. The dialyzed fractions were clarified by centrifugation and stored at −20 °C.

$NaDodSO_4$–polyacrylamide gel electrophoresis on the pooled fractions used for reconstitution with DNA (see below) showed that the gene 5 protein preparation was essentially homogeneous in molecular weight, close to the expected value of 9690. Overloaded gels showed only one minor band of protein at a molecular weight of about 24 000, representing less than 2% of the total protein. The UV absorption and CD spectra of the protein agreed with those published by Day (1973). Concentrations of gene 5 protein were determined by using an extinction coefficient of 0.73 $mg^{-1}\cdot cm^2$ at 276 nm (Day, 1973).

Isolation and Deuteration of DNA. fd phage was obtained from Dr. D. A. Marvin, European Molecular Biology Laboratory, Heidelberg, West Germany. Phages containing normal protonated DNA (H-DNA) were grown on a host *E. coli* strain Ton 1101 (recommended and donated by Dr. K. Geider, Max-Planck-Institut für Medizinische Forschung, Heidelberg, West Germany) in a medium consisting of 32 g of tryptone, 20 g of yeast extract, 5 g of NaCl, and 1 g of Tris base per L of water (Osborn et al., 1970). Phages containing deuterated DNA (D-DNA) were grown on the same host in a minimal medium made by dissolving 0.5 g of NaCl, 8 g of KCl, 0.1 g of $MgSO_4$, 12.1 g of Tris base, and 0.087 g of K_2HPO_4 in 1 L of 99.8% D_2O, adjusting the pH to 7.4 with HCl, and then adding 37 mL of a mixture containing 0.124 g of $CaCl_2$, 1 g of NH_4Cl, 3 g of sodium pyruvate, 1 μg of $FeCl_3$, 0.50 g of casamino acids (Difco, Vitamin Assay), and 0.010 g of vitamin B_1 in 99.8% D_2O. As described by Moore & Engelman (1976), a high percentage uptake of deuterium from D_2O into nucleic acids in *E. coli* occurs when carbon sources are restricted to pyruvate and amino acids. We find that the uptake of deuterium into the DNA of fd phage is also high (see below) when the phage are grown on an *E. coli* host in such a medium.

For phages containing normal H-DNA, cells were removed by low-speed centrifugation 1.6 h after infection at a multiplicity of about 10 phages/cell at a concentration of 3×10^8 cells/mL. Phages were isolated from the supernatant by overnight precipitation at 4 °C in the presence of 0.5 M NaCl (added first), 3% (w/v) poly(ethylene glycol) (average M_r 6000), and 0.1% (w/v) sarkosyl. After low-speed centrifugation of the precipitate, the pellet was resuspended overnight in approximately 0.1 the original volume of 0.1 M Tris-HCl, pH 8, and the precipitation was repeated without the sarkosyl. For phages containing D-DNA, the infection was allowed to proceed overnight before removal of the cells and precipitation of the phages. Both types of phage were further purified by CsCl density-gradient centrifugation and were finally dialyzed into 0.05 M Tris-HCl, pH 8. DNA was extracted from the phages by gently rocking the phage solutions for 5 min with equal volumes of buffer-saturated, redistilled phenol. The first extraction was performed with the phenol preheated to 65 °C, and three additional extractions were done at room temperature as recommended by Marvin & Schaller (1966). The aqueous phases containing the DNA were separated from the phenol phases by centrifugation when necessary, and the final aqueous phase was dialyzed into 0.05 M Tris-HCl, pH 8.

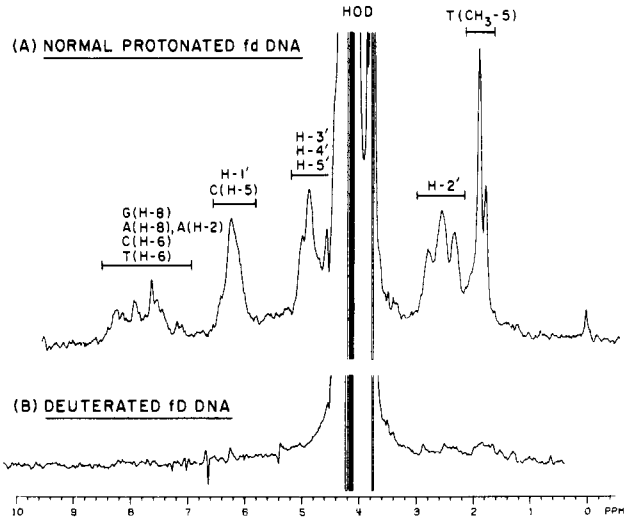

FIGURE 1: 200-MHz NMR spectra of equal concentrations of samples of (A) normal protonated and (B) deuterated fd DNAs at 90 °C. No reference compound was added to the samples; the abscissa was positioned so that the characteristic thymine CH_3 resonances occur at 1.72–1.82 ppm, as reported by Patel & Canuel (1979) for these resonances in an octadeoxynucleotide at 90 °C. Positions of the other base and sugar resonances as labeled are then in agreement with published values (Patel & Canuel, 1979; Patel, 1977; Fang et al., 1971). The spectrum of the deuterated DNA sample was aligned with that of the protonated DNA sample by the large HOD peak. The deuterated DNA shows a major loss of every proton resonance. (The spectra were generously taken by Dr. Ben Shoulders, Department of Chemistry, The University of Texas at Austin, with a Nicolet NT-200 NMR spectrometer.)

We previously estimated the molar extinction coefficient, ϵ, of fd DNA at 259 nm to be 8350 $L\cdot mol^{-1}\cdot cm^{-1}$ in 0.01 M Tris-HCl, pH 8 (Torbet et al., 1981). We have now made a second estimate of ϵ for fd DNA in 0.05 M Tris-HCl, pH 8, again relying on the determination of ϵ made by Berkowitz & Day (1974) for fd DNA in the presence of 0.195 M Na^+. This second estimate of ϵ at 259 nm is 8500 $L\cdot mol^{-1}\cdot cm^{-1}$, which is within 2% of the previous value we found for DNA in a Tris buffer of low concentration. We have used our more recent value for determining DNA concentrations in these experiments.

The extent of deuteration in the D-DNA sample was determined from the loss of resonances in its proton nuclear magnetic resonance (NMR) spectrum. Portions of the D-DNA and H-DNA samples were dialyzed against 99.8% D_2O, lyophilized, and redissolved in 99.96% D_2O. Solutions of the DNAs containing equal concentrations of 1.50 ± 0.02 mg/mL (determined by the absorbance of a weighed dilution of each solution of redissolved DNA) were used to obtain the NMR spectra at 90 °C shown in Figure 1. It is clear from these spectra that there was almost a total loss of the nonexchangeable proton resonances from the D-DNA sample. The extent of deuterium replacement was no less than 87% as determined from the areas under the proton peaks in the two

[1] Abbreviations: Tris, tris(hydroxymethyl)aminomethane; $NaDodSO_4$, sodium dodecyl sulfate; CD, circular dichroism; EDTA, ethylenediaminetetraacetic acid; OD, optical density.

Table I: Samples of Complex Used for Neutron Scattering Measurements

% (v/v) D_2O of samples		concn of complex (mg/mL)	length measurements by electron microscopy: no. of molecules measured	av length (μm)	% mass in lengths ≥ 0.1 μm
complex containing D-DNA:	0.0	17.8	89 (27)[a]	0.13 (0.55)	81 (99.5)
	14.8 ± 0.2	17.1	80	0.12	81
	66.4 ± 0.2	15.4	102	0.15	85
	79.5 ± 0.2	14.6	92	0.13	86
	99.3 ± 0.1	14.3	43	0.23	94
complex containing H-DNA:	0.0	19.3	(28)	(0.34)	(97)
	14.8 ± 0.2	19.1	73	0.13	80
	66.4 ± 0.2	19.4	93	0.11	73
	79.5 ± 0.2	19.5	72	0.10	72
	99.3 ± 0.1	19.5	82	0.14	83
	99.3 ± 0.1	9.6			
	99.3 ± 0.1	5.1			

[a] Data in parentheses are for portions of the original samples not used for neutron scattering experiments.

spectra. The absence of unidentified peaks in these spectra is a confirmation that the samples were essentially free of contaminants.

Reconstitution of Complexes. The reconstitution of complexes was monitored by circular dichroism (CD) measurements. Upon mixing the gene 5 protein and DNA components, we observed a decrease in the magnitude of the 228.5-nm band of the gene 5 protein and a change to more negative values in the CD of the DNA at wavelengths above 250 nm. These changes are qualitatively similar to those reported by others (Day, 1973; Pretorius et al., 1975; Anderson et al., 1975). Figure 2 shows a series of spectra for five mixtures at different protein to nucleotide molar ratios, in 0.05 M Tris-HCl, pH 8, all the spectra representing the same DNA concentration. That is, the spectra are all on an $\epsilon_L - \epsilon_R$ scale where the molar concentration is that of DNA nucleotides. (The actual mixtures all contained the same protein concentration of 2.2×10^{-5} M). We found that the change in the CD of the DNA was accompanied by a shift in the crossover wavelength to 277 nm as the DNA became saturated with protein. The end point of the CD changes occurred at a protein to nucleotide molar ratio of 0.34, with about one protein monomer present per three nucleotides.

Our data are quite close to those previously reported by Anderson et al. (1975) for the binding of gene 5 protein to fd DNA in a buffer of 0.01 M Tris-HCl and 1 mM Na_2-EDTA, pH 8. We note that saturation also appears to occur at a molar ratio of 0.33 in their plot of the decrease in CD at 270 nm of fd DNA at increasing ratios of protein to DNA nucleotides. However, these authors rely more on the ratio observed for binding of gene 5 protein to poly[d(AT)] for their reported finding of a saturation ratio of 0.25. The CD spectrum shown by Anderson et al. (1975) for a complex reconstituted with fd DNA at a protein to nucleotide ratio of 0.33 has a crossover wavelength of 276 nm and is very close to our spectrum shown in Figure 2 for a complex at a ratio of 0.34. If a dissociation constant close to the maximum of 5×10^{-8} M estimated by Anderson et al. (1975) is also valid for our reconstituted complex, our solutions would contain a negligible amount of free protein.

On the other hand, in a detailed study, Day (1973) found by titrations in other buffers that the protein to nucleotide ratio is close to 0.25 at saturation, indicating that the gene 5 protein binding site accommodates four, not three, nucleotides. Results of sedimentation experiments and filter binding assays also give saturation ratios closer to 0.25 than to 0.33 (Alberts et al., 1972; Oey & Knippers, 1972; Cavalieri et al., 1976). Thus,

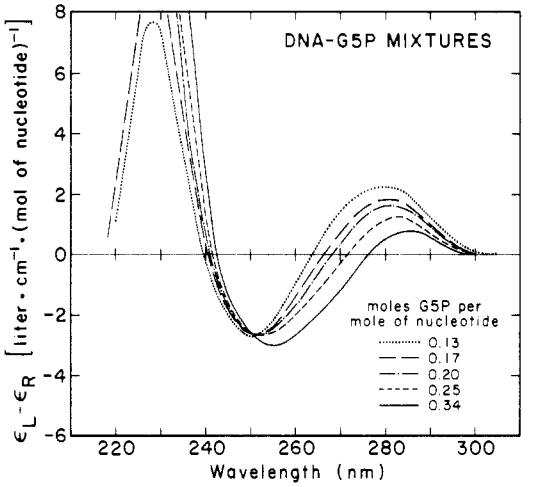

FIGURE 2: CD spectra of a series of mixtures of gene 5 protein (G5P) and fd DNA at different protein to nucleotide molar ratios. Separate mixtures were made at each designated ratio. Concentrations of the separate DNA and gene 5 protein components were determined from absorption spectra and the extinction coefficients given in the text. $\epsilon_L - \epsilon_R$ values are given for the DNA component in units of L·mol^{-1}·cm^{-1}, per mol of nucleotide. The buffer was 0.05 M Tris-HCl, pH 8, and spectra were taken at 20 °C. (CD spectra for Figures 2 and 3 were taken on a Cary Model 61 circular dichrometer, calibrated with *d*-10-camphorsulfonic acid to give a value within 2% of 0.336 degrees ellipticity at 290.5 nm for a 0.1% (w/v) solution at 20 °C.)

it is possible that the binding of gene 5 protein may not be very tight in our buffer of 0.05 M Tris-HCl, pH 8, and that there is a significant amount of free, native protein present at saturation in the experiment shown in Figure 2, which increases the apparent saturation ratio from 0.25 to 0.33. If this is so, we estimate from our data that the free protein concentration present at DNA saturation in this buffer would be $(5.5–6.0) \times 10^{-6}$ M. [We have no evidence for denatured gene 5 protein in our preparations. For example, the magnitude of the 228.5-nm CD band of the protein shown in Figure 3 is similar to that reported by Day (1973), and this band is very sensitive to denaturation, being completely lost in denatured samples.]

This does not mean that there would be a significant concentration of free protein present in the concentrated solutions of complex used for the neutron scattering experiments. By analogy with the binding of T4 gene 32 protein to poly(rA) under nonstoichiometric conditions (Kowalczykowski et al., 1981), we expect that the concentration of free protein present at a given percentage of polynucleotide saturation will be independent of the polynucleotide concentration. The concentrations of complex used for neutron scattering were no less

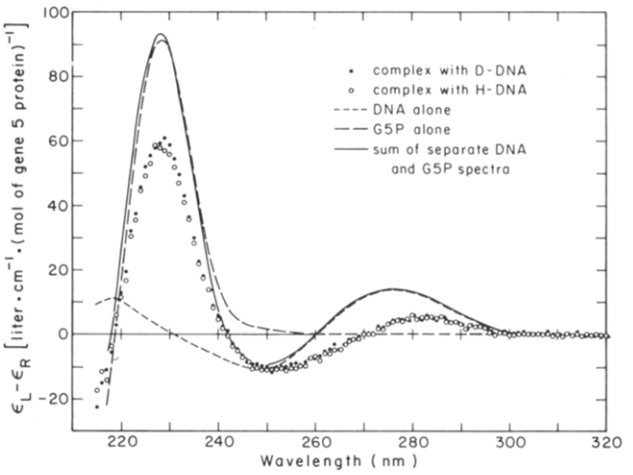

FIGURE 3: CD spectra of reconstituted complexes containing D-DNA (●) or H-DNA (○) mixed with gene 5 protein at a ratio of four nucleotides per protein monomer. Individual data points are plotted to show the error in each spectrum. Also shown are spectra of the separate DNA and protein components and their sum, to show the changes that occurred upon complex formation. $\epsilon_L - \epsilon_R$ values are given in units of L·mol^{-1}·cm^{-1}, per mol of gene 5 protein (or per 4 mol of nucleotide). The buffer was 0.05 M Tris-HCl, pH 8, and spectra were taken at 20 °C.

than 5 mg/mL (see Table I) and contained gene 5 protein at a total concentration of 4.5×10^{-4} M, with sufficient DNA present to bind all of the protein at a protein/nucleotide ratio of 0.25. A free gene 5 protein concentration of $(5.5–6.0) \times 10^{-6}$ M would correspond to a percentage of free protein of about 1.3% in the least concentrated of the solutions used for neutron scattering, and the DNA would be essentially saturated with protein. As discussed by Torbet et al. (1981), 1.3% of free gene 5 protein would make a negligible contribution of about 0.2% to the cross-sectional scattering from solutions of the complex.

Large-scale reconstitution of complexes was performed by adding the gene 5 protein to H-DNA and D-DNA in small beakers while stirring, to give a final ratio of four nucleotides per protein monomer. The concentrations of complex in the mixtures were 1.1–1.5 mg/mL, with final volumes of about 20 mL, in a buffer of 5 mM NaCl, 3–4 mM Tris-HCl, pH 7–8, and 7×10^{-5} M β-mercaptoethanol. From measurements of the absorption spectra of the mixtures and a knowledge of their concentrations (obtained from absorption spectra of the components and the extinction coefficients given above), the extinction coefficient of the complex could be estimated. We obtained a range of $\epsilon(260) = 40\,000 \pm 600$ L·mol^{-1}·cm^{-1}, or an $A(260) = 3.63 \pm 0.06$ cm^2·mg^{-1}, for the two mixtures. The average is 4% higher than the value previously estimated for complex isolated from infected cells (Torbet et al., 1981).

The mixtures were then dialyzed at 4 °C against 0.6 M NaCl and 0.05 M Tris-HCl, pH 8, followed by 0.05 M Tris-HCl, pH 8, to dissociate and more gently reanneal the complex. The solutions were concentrated by a factor of 4–5 by blowing air over the dialysis tubings. After further dialysis in smaller tubing, the solutions were each again concentrated by a factor of 4–5 and finally were simultaneously dialyzed into 0.05 M Tris-HCl, pH 8, at concentrations approaching 20 mg/mL. Stock concentrations were determined by absorbance measurements of weighed dilutions of the stock solutions. A total of 1–2 mL of each solution of concentrated complex was obtained.

The CD spectra of dilutions of the concentrated complexes are shown in Figure 3. The preparations of complex containing H-DNA and D-DNA show no significant spectral differences.

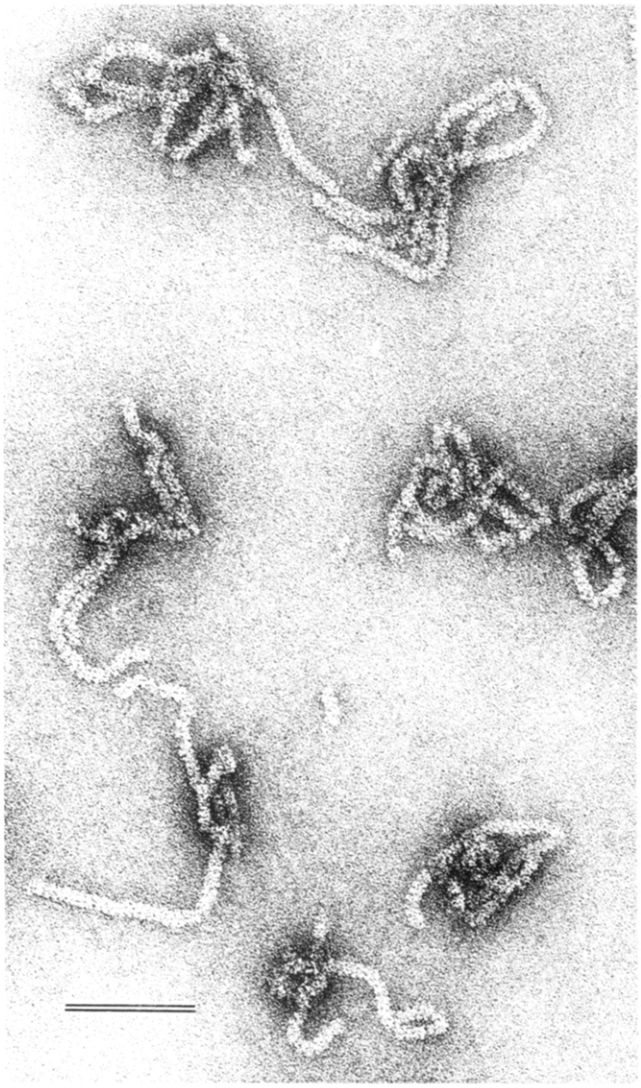

FIGURE 4: Electron micrograph of segments of the complex formed by mixing fd gene 5 protein with fd H-DNA. The complex was diluted into 0.01 M ammonium acetate, pH 7, fixed with glutaraldehyde, and adsorbed to a carbon film that was floated onto the solution (Valentine et al., 1968). The specimens were negatively stained with 2% aqueous uranyl acetate and photographed in a Siemens Elmiskop IA, calibrated with a difffraction grating replica. Bar = 0.1 μm.

To show the changes in CD that have occurred upon formation of the complexes, we included spectra of the individual H-DNA and gene 5 protein components and also their sum in Figure 3.

Electron micrographs of the complexes showed flexible, helical structures similar to those previously observed for both in vitro reconstituted and in vivo complexes (Torbet et al., 1981). Figure 4 shows a micrograph of segments from the preparation of concentrated complex containing H-DNA (negatively stained to show the nucleoprotein structure); the complex containing D-DNA (not shown) was similar in structure. A sample of D-DNA complex (in 14.8% D_2O) was found to be unchanged in structure (except for some fragmentation as noted below) after it had been used for neutron diffraction experiments. When the same preparations of complex were spread with cytochrome *c*, no naked DNA strands were seen.

Preparation of Samples for Neutron Scattering. Half of each solution of concentrated complex was dialyzed against several changes of 0.05 M Tris-HCl, pD 8.0, made up in 99.8% D_2O (Stohler Isotope), until the D_2O concentration in the final dialyzate was calculated to be 99.7%. The actual D_2O con-

centration in the final dialyzate was determined to be 99.3 ± 0.1% by comparing its infrared absorption at 16 600 Å (due to residual H_2O) with that of dilutions of a freshly opened bottle of 99.96% D_2O (Crespi & Katz, 1961). Measurements were made on a Cary Model 14R spectrophotometer. Both samples of complex were simultaneously dialyzed into the same D_2O solvent. The pD of the D_2O solvent was taken to be the measured pH + 0.4 (Glasoe & Long, 1960). All four samples (complex containing H-DNA and complex containing D-DNA, each in D_2O as well as H_2O solvents containing 0.05 M Tris-HCl) were finally clarified by centrifuging for 10 min in an Eppendorf Model 5412 microcentrifuge. The CD spectra of dilutions of the complexes in the D_2O solvent (not shown) were not significantly different from those of the complexes in the H_2O solvent (shown in Figure 3). The samples were stored at 4 °C.

Samples of complex at three intermediate percentages of D_2O were obtained by mixing weighed portions of the samples in the pure H_2O and 99.3% D_2O solvents. For the scattering experiments, sample volumes of about 125 μL were contained in cylindrical microcells of 2-mm path length. Concentrations of complex were calculated from the known stock concentrations and the proportions of the stock solutions in the mixtures. Concentrations were checked by measuring the optical densities at 295 nm in the microcells. The optical densities at 260 nm could not be directly measured since they were in excess of 10 for the concentrated solutions, even in 2 mm path length cells. The calculated and measured concentrations were never in disagreement by more than 3%. The optical density at 320 nm, taken as an indication of scattering from precipitated material, was always within 1.4% of the absorbance at 260 nm, a value comparable with that of freshly centrifuged samples and which indicated a minimum of precipitated material. A check of the optical densities at 295 and 320 nm was made for every sample after exposure to the neutron beam. Optical densities at 295 nm generally were slightly decreased, with concomitant increases in the optical densities at 320 nm, after exposure to the neutron beam. This indicates that some small amount of precipitation occurred during the exposures. The decrease in the optical density at 295 nm was greatest for the samples in 66.4 and 14.8% D_2O, for which the neutron exposure time was the longest. In any case, the decrease in OD(295) was never more than 3%. The average of the optical densities before and after the neutron scattering measurements was used to calculate the concentrations of complex in the beam during data collection. The percentage of D_2O and the concentration of complex in each mixture are listed in Table I. The concentrations are accurate to about ±3%, the largest error being that of the extinction coefficients of the gene 5 protein (±3% or less; Day, 1973) and the DNA (±4%, including the uncertainty in our estimate of the extinction coefficient in 0.01–0.05 M Tris; Day, 1973). The error in the DNA to protein ratio in our samples is thus about ±5%. We assume that there was no selective loss of either component after reconstitution.

In an effort to assess the extent of fragmentation during the experiment due to manipulation of the samples during mixing and cell loading and due to exposure to the neutron beam over a period of up to 3 days at 20 °C, length measurements were made on segments of complex from each of the samples after completion of the neutron scattering experiments. All segments were measured in each of several electron micrographs of cytochrome *c* spreads. Photographs were calibrated with a diffraction grating replica, and lengths of the segments were measured with a Numonics line integrator. The length measurements did not include lengths below 0.01 μm, which were too small to detect. The number of segments measured from each sample and their average lengths are listed in Table I. A complex containing a full-length molecule of fd DNA is about 1 μm long under these conditions. Therefore, most of the segments were not full length. However, most of the mass was in segments with lengths greater than 0.1 μm even at the conclusion of the experiments. The percentage of mass in measured lengths 0.1 μm or longer is given in the last column of Table I. Segments 0.1 μm in length have a length:width ratio of 10:1.

A graph of the log (number of segments) vs. measured lengths of segments from the samples used for neutron diffraction gives a straight line of nonzero slope (not shown), indicating that the fragmentation was due to a random process (Charlesby, 1954) and was not due to shearing by manipulation. The slope of this line gives a number-average length of 0.16 μm and, with the use of an equation for the distribution of molecular weights in a randomly fragmented sample, leads to a percentage of mass of 87% in lengths greater than 0.1 μm. From the measured lengths, the average length at the conclusion of the experiments was 0.13 ± 0.04 μm, and the percentage of mass in lengths 0.1 μm or longer was 81 ± 6%. The agreement between the measured values and those calculated on the basis of random breakage assures us that there was no significant number of segments of lengths less than 0.01 μm, which could not be measured, and that the neutron scattering data we obtained were from samples with 81–87% of the mass remaining in rodlike segments throughout the experiments.

The CD spectrum of a sample of the complex containing H-DNA (in 79.5% D_2O) that had been used to obtain the neutron scattering data showed spectral characteristics unchanged from those shown in Figure 3. Together, the electron microscopy and CD data suggest that, except for some fragmentation, the structure of the complex was unchanged during the neutron diffraction experiments.

Neutron Scattering. Neutron scattering experiments were conducted at the H4-S low-angle neutron spectrometer in the High Flux Beam Reactor at Brookhaven National Laboratory (BNL). The design of the spectrometer and counter and their parameters have been described (Schoenborn et al., 1978). The sample-to-detector distance was 198.3 cm, and the nominal wavelength was 2.33 Å ($\Delta\lambda/\lambda = 0.02$). Collimation was by two circular apertures 6 mm in diameter, 1.4 m apart. Previous calculations at BNL have shown that deconvolutions of the data to account for smearing effects of this collimation geometry and wavelength spread were not necessary. As many as seven samples were held in an automatic sample changer kept at constant temperature by circulating liquid from a water bath at 20 ± 0.5 °C. Sample exposures to the neutron beam were for a preset number of monitor counts rather than for preset times, to eliminate the effect of fluctuations in beam intensity. Moreover, the buffer and the two samples of complex containing D-DNA and H-DNA, all at the same percentage of D_2O, were alternately exposed to the beam to minimize any variations in background, beam fluctuations, or sample deterioration during data collection, which was for a period of from 1 to 3 days.

The incident neutron flux at the sample was 7.72 ± 0.12 × 10^5 $n \cdot cm^{-2} \cdot s^{-1}$. The flux was determined by measuring the attenuation of the beam by five different thicknesses of lucite slabs and extrapolating the data to zero attenuation.

The scattered neutrons were detected with a two-dimensional detector (Schoenborn et al., 1978). The number of

counts in successive annular rings of width 2.4 mm was calculated and divided by the ring area to reduce the two-dimensional array of counts to a single scattering curve, $I(Q)$, of intensity (in counts·mm^{-2}) as a function of Q, Q being the amplitude of the scattering vector. Transmissions were measured with the backstop removed and the beam attenuated. The detector response was checked with an incoherent scatterer and was sufficiently uniform in area so that it was not necessary to correct for nonuniformities.

The scattering curve of the solute was obtained by using

$$I(\text{solute}) = I(\text{sample})/T_s - (1 - \bar{v}c)I(\text{buffer})/T_b - (\bar{v}c)I(\text{empty cell})/T_e \quad (1)$$

In this equation I(sample), I(buffer), and I(empty cell) are the scattering curves of the sample at a given percentage of D_2O, of the buffer at the same percentage of D_2O, and of the empty cell, respectively; T_s, T_b, and T_e are the fractions of the incident beam transmitted by the sample cell containing solute dissolved in buffer, by the sample cell containing buffer, and by the empty sample cell, respectively; c is the concentration of complex in milligrams per milliliter; $\bar{v}$ is the partial specific volume of the complex in cubic centimeters per gram. We calculated $\bar{v}$ from the scattering length density of the solvent at the contrast match point for the complex from the previous work (Torbet el al., 1981), knowing the sum of scattering lengths and molecular weight for a gene 5 protein monomer plus a tetranucleotide. Then, after determining the contrast match point for our present sample of complex containing H-DNA, we recalculated $\bar{v}$ to be 0.735 cm^3·g^{-1}. This latter value was used in a final determination of solute scattering curves, which altered the contrast match point by only 0.3% D_2O, well within the error of the match point determination. The value of $\bar{v}$ was assumed to be the same for both types of complex in applying eq 1.

Radii of gyration were determined by using I(solute) directly from eq 1. For the calculation of mass per unit length, the scattering curves were further corrected to an absolute scale with the equation

$$I'(\text{solute}) = I(\text{solute})r^2/(I_0 a) \quad (2)$$

where I(solute) is in counts (for a given number of monitor counts) per square millimeter of detector surface, I_0 is the incident flux in counts per square centimeter for the same number of monitor counts, r is the sample-to-detector distance in millimeters, and a is the area in square centimeters of sample exposed. As long as the sample area exposed is equal to the full beam cross section used to determine I_0, a cancels and need not be explicitly known. The multiplication by r^2 makes I' the relative scattered intensity per steradian subtended by the detector at a given value of Q.

Results

Guinier Plots and Cross-Sectional Radii of Gyration. The coherently scattered intensity from rodlike particles at a given percentage of D_2O can be approximated over a limited range of Q by the Guinier equation:

$$IQ = C\exp(-R^2Q^2/2) \quad (3)$$

I is as defined in eq 1, Q is the amplitude of the scattering vector [related to the scattering angle 2θ, at wavelength λ, by $Q = (4\pi \sin\theta)/\lambda$], R is the cross-sectional radius of gyration, and C is a constant. Thus, a graph of ln (IQ) vs. Q^2 in the allowed range of Q will have a slope proportional to R^2 [see a review by Kratky & Pilz (1972)]. In these experiments we were limited to the range of Q^2 above 0.8×10^{-3} Å^{-2}. In the previous work (Torbet el al., 1981) it was found that the scattering data for the complex of fd DNA and gene 5 protein satisfies the linear relationship described by eq 3 for Q^2 well into this range. In the present experiments we have used data points up to a Q^2 of 2.4×10^{-3} Å^{-2}, and our values of R for complex containing H-DNA are generally consistent with those obtained earlier.

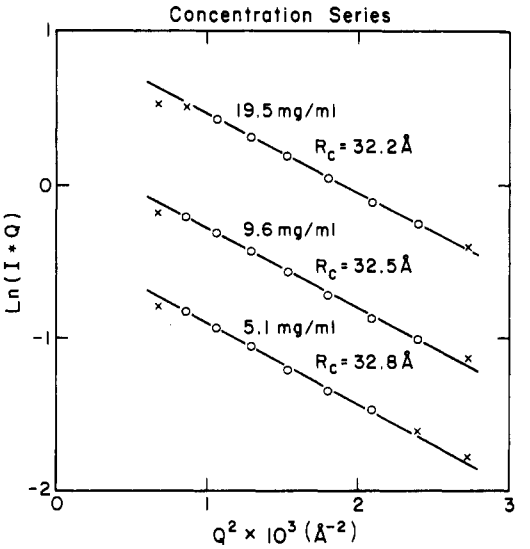

FIGURE 5: Guinier plots for a concentration series of the complex containing H-DNA. Slopes [shown as (—)] were obtained by least-squares linear regression of the points shown as (○) at each concentration. Points [shown as (×)] which departed from linearity at lower and higher values of Q^2 were omitted from the least-squares regression to find the slopes in this figure and Figures 6 and 7.

As shown in Table I, our samples of the complex were at concentrations approaching 20 mg·mL^{-1}. Data were also collected for two dilutions of the sample of complex containing H-DNA (in 99.3% D_2O) to test for interparticle interference. The Guinier plots of ln (IQ) vs. Q^2 for the two dilutions and the concentrated sample are shown in Figure 5. The slopes of the curves were obtained by least-squares linear regressions. For concentrations of 19.5, 9.6, and 5.1 mg·mL^{-1}, the slopes gave R values of 32.2 ± 0.2, 32.5 ± 0.4, and 32.8 ± 0.6 Å, respectively. The errors are ±1 standard deviation as determined from the regression analyses. Within the error of these values, R exhibited no concentration dependence, indicating the absence of significant intermolecular interference.

Figures 6 and 7 show the Guinier plots of the samples of complex containing H-DNA and D-DNA, respectively. In agreement with Torbet et al. (1981), the slopes of the plots for complex containing H-DNA shown in Figure 6 do not vary greatly with the percentage of D_2O. This is shown more clearly in Figure 8, where R^2 is graphed as a function of the inverse of the contrast $\bar{\rho}$. The contrast is the difference between the solvent and mean solute scattering densities, a difference that is proportional to the percentage of D_2O in the solvent. By analogy with the expression derived by Ibel & Stuhrmann (1975), the relation between R^2 and $\bar{\rho}$ for a cylindrically symmetric structure is

$$R^2 = R_c^2 + \alpha/\bar{\rho} \quad (4)$$

where R is the cross-sectional radius of gyration measured at contrast $\bar{\rho}$, and R_c is the cross-sectional radius of gyration at infinite contrast ($1/\bar{\rho} = 0$). The slope, α, provides information about the radial distribution of neutron scattering density. For complex containing H-DNA, R_c(H-DNA complex) = 31.9 ± 0.3 Å and $\alpha = (-0.2 \pm 0.3) \times 10^{-4}$, where the errors are ±1 standard deviation from a least-squares linear regression analysis of the data in Figure 8. These values are in reasonable

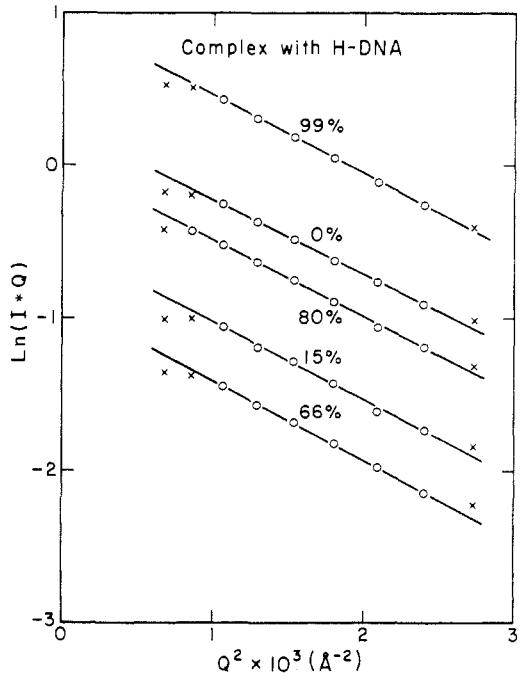

FIGURE 6: Guinier plots for complex containing H-DNA at various percentages of D_2O in D_2O–H_2O mixtures. Slopes were obtained by least-squares linear regression of the points shown as (O) at each percentage of D_2O.

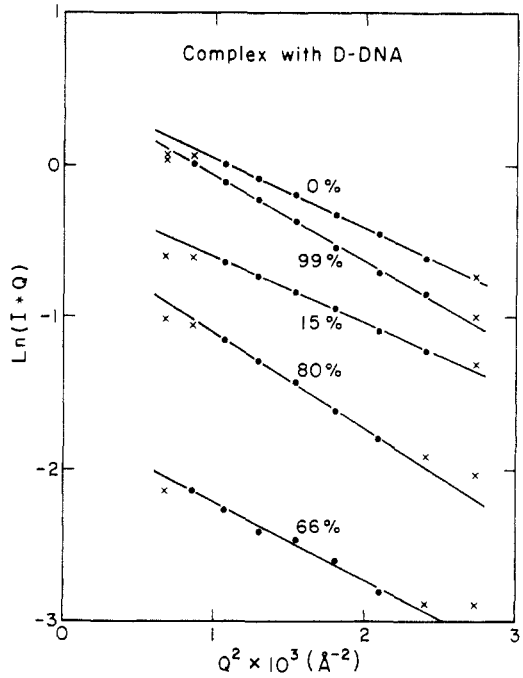

FIGURE 7: Guinier plots for complex containing D-DNA at various percentages of D_2O in D_2O–H_2O mixtures. Slopes were obtained by least-squares linear regression of the points shown as (●) at each percentage of D_2O.

agreement with those previously obtained, $R_c = 34.5 \pm 1$ Å and $\alpha = (-0.8 \pm 0.2) \times 10^{-4}$, considering that our present values of R were obtained from a more limited range of Q. The near-zero value of α shows that the cross-sectional distribution of scattering density is fairly uniform. Even though the DNA and protein components differ significantly in their scattering densities, the DNA constitutes only 12% of the mass of the complex. Thus, it is possible that significant differences in the radial distribution of the two components would not be obvious with changing contrast, for complex containing H-DNA.

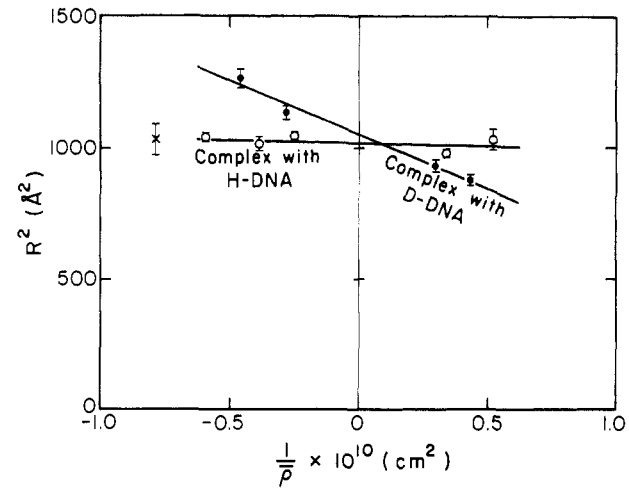

FIGURE 8: Square of the cross-sectional radius of gyration, R, as a function of the reciprocal of the contrast, $\bar{\rho}$. Complex containing H-DNA (O); complex containing D-DNA (●). Slopes were found by least-squares linear regressions, the point shown as × being omitted for the complex containing D-DNA. The intercepts at infinite contrast ($1/\bar{\rho} = 0$) give values for R_c^2. Error bars show ±1 standard deviation from linear regression analyses of the Guinier plots in Figures 6 and 7.

By deuteration of the DNA, the difference in the scattering densities of the DNA and protein is enhanced. Figure 7 shows that the slopes of the Guinier plots for complex containing D-DNA have an obvious dependence on the percentage of D_2O. The DNA component must thus have a different radius of gyration from the protein component of the complex; if the two radii of gyration were the same, the deuteration of one component would not change the slopes of the Guinier plots relative to those for the complex containing protonated DNA. The slope of the Guinier plot for complex containing D-DNA at 66.4% D_2O is the least precisely defined due to the low intensities of the scattered neutrons from this sample, which had the lowest contrast of all the samples we used. The increased scatter in the points for this sample is obvious in Figure 7. At this contrast, the intensity scattered by the solute is small compared with the intensity scattered by the solvent, and small differences in the sample and buffer D_2O concentrations could account for the fact that the radius of gyration for this sample is not consistent with the other four on a graph of R^2 vs. $1/\bar{\rho}$. As shown in Figure 8, with the exception of the point obtained at 66.4% D_2O (shown as X), the data obtained for the complex containing D-DNA lie on a line with R_c (D-DNA complex) = 32.4 ± 0.3 Å and $\alpha = (-4.15 \pm 0.42) \times 10^{-4}$. The radii of gyration at infinite contrast ($1/\bar{\rho} = 0$) for the two types of complex are the same within experimental error, as they should be. The definite negative slope in Figure 8 for the complex containing deuterated DNA means that the DNA must be centrally located in the helical complex structure. If the DNA were wrapped on the exterior of the protein, the slope would have become more positive for the complex containing deuterated DNA relative to the slope for the complex containing normal protonated DNA. Values of R_c as well as other data to be discussed for the complexes are summarized in Table II.

Solvent-Excluded Volume and Mass per Unit Length. The data shown in Figures 6 and 7 may be extrapolated to $Q = 0$ and scaled to give $(I'Q)_0$, where I' is on an absolute intensity scale as defined by eq 2 under Materials and Methods. $(I'Q)_0$ can be related to the solvent-excluded volume, V, and the mass per unit length, μ, by the formula given in Torbet et al. (1981):

$$(I'Q)_0 = (\pi c d N_A \mu / M^2)(\sum b_i - V\rho_s)^2 \qquad (5)$$

Table II: Parameters of the Complex As Determined by Neutron Scattering[a]

parameter	complex containing H-DNA	complex containing D-DNA of 87% deuteration
R_c(complex) (Å)	31.9 ± 0.3	32.4 ± 0.3
R_c(DNA) (Å)		17.6 ± 3.1
R_c(protein) (Å)		33.5 ± 0.4
V, solvent-excluded volume (Å^3)	13460 ± 230	13170 ± 140
$\bar{v}$, partial specific volume[b] (cm^3·g^{-1})	0.742 ± 0.013	0.724 ± 0.008
μ, mass per unit length in H_2O[b] (daltons·Å^{-1})	1860 ± 95	2090 ± 90
L, contour length[b] (Å)	9410 ± 480	8400 ± 350
axial rise per subunit[b] (Å)	5.9 ± 0.3	5.2 ± 0.2

[a] Errors are ±1 standard error resulting from linear regression analysis of plotted data and from estimated errors in concentrations and incident beam flux. The errors we show for R and μ are errors from statistical analyses of the scattering curves over a given range of Q; these would probably be increased to the errors of ±1 Å and ±110 daltons·Å^{-1}, respectively, previously reported by Torbet et al. (1981), if we could include the uncertainty in fitting different ranges of Q, as in the previous work. Even if the error in R_c(complex) were increased to ±1 Å, the error in R_c(DNA) would only be increased to ±3.5 Å since this error is dominated by the error in α, not in R_c(complex). [b] Taking the molecular weight of a subunit (protein monomer plus four nucleotides) of complex containing H-DNA or D-DNA to be 10922 or 10955, respectively.

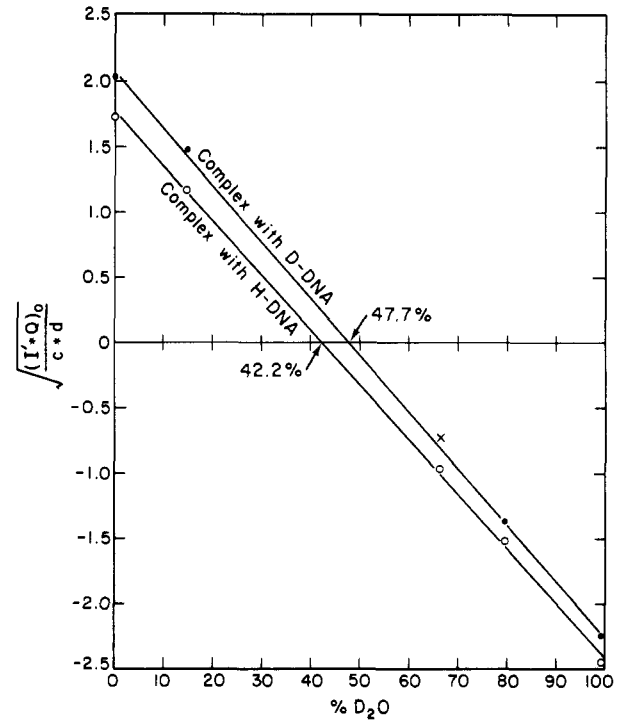

FIGURE 9: Variation of the square root of $I'Q$, extrapolated to Q = 0 and divided by the sample concentration and path length, as a function of the percentage of D_2O. Complex containing H-DNA (○); complex containing D-DNA (●). Lines drawn are from least-squares linear regressions, the point shown as × being omitted for the complex containing D-DNA. The intercepts at $(I'Q)_0 = 0$ are the contrast match points.

In this equation, c is the concentration of complex, d is the cell path length, N_A is Avogadro's number, μ is the mass per unit length, M is the molecular weight of a subunit (gene 5 protein plus a tetranucleotide), $\sum b_i$ is the sum of the scattering lengths in the subunit, V is the solvent-excluded volume, and ρ_s is the solvent scattering density. ρ_s is proportional to the percentage of D_2O, and $\sum b_i$ will be proportional to the percentage of D_2O if the exchangeable protons are at equilibrium. Thus, $[(I'Q)_0/(cd)]^{1/2}$ should vary linearly with the percentage of D_2O. Figure 9 shows that this is true for both types of complex, excluding the data point for D-DNA complex in 66.4% D_2O. The fact that our data fall on straight lines in this graph shows that our extrapolated $(IQ)_0$ values from the Guinier plots are consistent.

The contrast match point for complex with H-DNA occurs at 42.2 ± 0.7% D_2O, and for complex with D-DNA the match point occurs at 47.7 ± 0.5% D_2O. For comparison with these values, we have calculated values using the data given in the review by Jacrot (1976) for the neutron scattering lengths and volumes of the peptide residues and for the scattering lengths of the protonated and deuterated nucleotide residues, using 290 Å^3 for the volume of a nucleotide (Cohen & Eisenberg, 1968). We assumed that the solvent-excluded volumes of the gene 5 protein and DNA are additive and that the exchangeable protons are in equilibrium with the solvent. Combining these data and assumptions with the known amino acid composition of gene 5 protein (Nakashima et al., 1974) and the nucleotide composition of fd DNA (Day & Wiseman, 1978), we calculate that the match point for complex containing protonated DNA should be at 41.8% D_2O and that the match point for complex containing DNA with 87% of the nonexchangeable protons replaced by deuterium should be at 46.2% D_2O. The measured contrast match point for complex containing H-DNA is in excellent agreement with the calculated value. Thus, a large majority of the possible exchangeable protons appear to be in equilibrium with the solvent. If only 80% of the potentially exchangeable protons are assumed to be accessible to the solvent, the contrast match point for complex containing H-DNA is calculated to be 40.1% D_2O, well below the measured value. The fact that the measured contrast match point for complex containing D-DNA is slightly higher than calculated suggests that the replacement of hydrogen by deuterium in the DNA may be greater than the 87% estimated from NMR spectra (see Materials and Methods). The calculated match point for complex containing DNA with 100% of the nonexchangeable protons replaced by deuterium is 46.8% D_2O.

The solvent-excluded volumes V may be calculated for the zero contrast condition when $[(I'Q)_0/(cd)]^{1/2} = 0$. In this case, from eq 5, $V = \sum b_i/\rho_s$, where V is the solvent-excluded volume of a gene 5 protein monomer plus an associated tetranucleotide and $\sum b_i$ and ρ_s are the contrast match point values. For complexes containing H-DNA and D-DNA, we obtain respectively V(H-DNA) = 13460 ± 230 Å^3 and V(D-DNA) = 13170 ± 140 Å^3, with the assumptions that 100% of the exchangeable protons are in equilibrium with the solvent and that in D-DNA the nonexchangeable protons are 87% replaced by deuterium. If the D-DNA were 100% deuterated, the latter value would be increased to 13360 ± 140 Å^3. Thus, the two experiments are consistent and give solvent-excluded volumes that are not significantly different from the additive volume of 13600 Å^3 for the amino acids in gene 5 protein plus four nucleotides. The value of V obtained by Torbet et al. (1981) was about 4% larger than this sum. The values calculated for V are reduced by 4% if it is assumed that only 80% of the exchangeable protons are in equilibrium with the solvent.

The partial specific volume $\bar{v}$ may be calculated from V as $\bar{v} = VN_A/M$ for the complexes. Using the above values of V = 13460 Å^3 and 13170 Å^3 for the two complexes, we obtain an average $\bar{v}$ of 0.733 cm^3·g^{-1}, almost identical with the value of 0.735 cm^3·g^{-1} we used in applying eq 1, as explained under Materials and Methods. Values of V and $\bar{v}$ for the complexes are listed in Table II.

The mass per unit length, μ, of the complexes in H_2O solution may be obtained by using eq 5 and the values of $[(I'Q_0/(cd)]^{1/2}$ extrapolated to 0% D_2O from Figure 9. Taking V to be the same for both complexes and equal to that obtained for complex containing H-DNA, we obtain a mass per unit length of 1860 ± 95 daltons·Å^{-1} for complex containing protonated DNA and 2090 ± 90 daltons·Å^{-1} for complex containing 87% deuterated DNA. (For the calculation of these values we take the molecular weight of a protein monomer plus four nucleotides, without cation, to be 10 922 and 10 955 for complexes containing H-DNA and D-DNA, respectively.) The values for mass per unit length are in reasonable agreement with each other, especially since the latter value would be reduced by about 3% if the D-DNA were closer to 100% deuterated. Moreover, these values correspond to contour lengths that are in agreement with those measured by electron microscopy. Contour lengths, L, may be calculated from the values of mass per unit length with the expression $L = 6408M/(4\mu)$, where 6408 is the number of nucleotides in full-length fd DNA (Beck et al., 1978). The calculated contour lengths are 9410 ± 480 and 8400 ± 350 Å for complexes containing full-length molecules of H-DNA and D-DNA, respectively. These two values bracket the value of 8800 ± 200 Å measured by electron microscopy for naturally occurring (in vivo) fd complex adsorbed to carbon from 0.01 M ammonium acetate, pH 7 (C. W. Gray, unpublished results; Gray et al., 1982; Torbet et al., 1981), conditions under which the structure of the complex is similar to that in 0.05 M Tris-HCl, pH 8.

The value of mass per unit length obtained in this work is for complex reconstituted at a protein:nucleotide ratio of 1:4. It is consistent with the value previously obtained by Torbet et al. (1981) for complex reconstituted at a protein:nucleotide ratio of 1:4.5, if the average number of nucleotides bound by one protein monomer was indeed 4.5 and not an integer of either 4 or 5. Then the mass per unit length obtained in the previous study is computed to be 1500 ± 110 daltons·Å^{-1}. Using our present value for complex containing H-DNA, we would predict a mass per unit length of approximately (4/4.5)(1860 ± 95) ≃ 1650 ± 85 daltons·Å^{-1} for a complex containing an average of 4.5 nucleotides bound/monomer. The agreement seems very reasonable, considering that different neutron spectrometers were used and the incident beam fluxes were determined by different procedures.

From the value of mass per unit length and the molecular weight of a subunit, the axial rise per subunit is easily calculated as M/μ. We obtain 5.9 ± 0.3 and 5.2 ± 0.2 Å as the rise per subunit from the data on complexes containing H-DNA and D-DNA, respectively. The above parameters are all summarized in Table II.

Subsidiary Maximum. As in the previous work (Torbet et al., 1981), we observed a subsidiary maximum at a Q of 0.07–0.08 Å^{-1} in each neutron scattering profile of the complex (not shown). We do not find a significant difference in the Q value of the maximum in comparisons of the scattering profiles from the complex containing H-DNA with those from the complex containing D-DNA at the same percentage of D_2O. In profiles for complex containing H-DNA, a slight shift of the maximum to higher Q values as the D_2O concentration changed from near the DNA match point was previously suggested to be evidence that the DNA was not on the surface of the complex (Torbet et al., 1981). However, we do not find, at D_2O concentrations below 85% where the effect should be maximal, that the profiles shift to still higher values upon deuteration of the DNA. Difference profiles for complexes containing D-DNA and H-DNA do not show a discernible extremum at $Q < 0.13$ Å^{-1}. Therefore, it appears that the subsidiary maximum at $Q = 0.07$–0.08 Å^{-1} is almost entirely due to the protein component of the complex.

Discussion

The above results show that the complexes containing protonated and deuterated DNA have the same cross-sectional radii of gyration at infinite contrast (R_c) and similar if not identical solvent-excluded volumes (V) and masses per unit length (μ). Combined with the CD, NMR, and electron microscopy data on the preparations, these results leave little doubt that the complexes are essentially identical except for the deuteration of the DNA in one of them. Moreover, the internal consistency of our data and the reasonable agreement of (1) our present R_c and μ values with those from previous work (Torbet et al., 1981), (2) the V values with the calculated sum of the volumes of the components, and (3) the contour lengths calculated from the μ values with those observed by electron microscopy (Gray et al., 1982) give us confidence that our data are quantitatively as well as qualitatively correct. The limited range of Q to which we were restricted could account for some of the differences in our present results compared with the previous work. Also, the complexes used in the previous work were reconstituted at a different nucleotide to gene 5 protein ratio (4.5 as compared with 4 in the present work). Therefore, minor differences in the results from the two investigations are not unexpected.

The fragmentation we found in the complex in this work (no assessment of fragmentation was done in the earlier work) would not have led to very different slopes in our Guinier plots. We have estimated the effect of fragments on the cross-sectional radius of gyration by calculating [according to Malmon (1957)] the scattering curves for cylinders having different lengths and a radius of 45.2 Å (chosen to give $R_c = 32$ Å for infinitely long cylinders). As the length-to-diameter ratio is decreased to as little as 3:1, the calculated Guinier plots remain linear over the range of Q^2 we have used in this work, and the cross-sectional radius of gyration *increases* gradually by less than 2 Å. We have experimentally observed a slight decrease in R_c compared with that of the earlier work. It should be emphasized that in spite of some fragmentation the preparations used in the present work were remarkably stable during the lengthy exposures to the neutron beam at 20 °C (see Materials and Methods).

Cross-Sectional Radii of Gyration of the DNA and Protein Components. A major result of our present investigation is that the slope of the plot of R^2 vs. $1/\bar{\rho}$ (Figure 8) is dramatically more negative for complex containing D-DNA than for the complex containing normal protonated DNA. Making use of the slope of this graph for complex containing D-DNA, we can estimate the cross-sectional radius of gyration of the DNA component. We showed above that the match point density of the complex containing H-DNA could be accurately predicted from the scattering lengths of the amino acids and nucleotides; in a similar manner, we have calculated that the match point density of the protein alone would be 0.0220×10^{-4} Å^{-2} (39.7% D_2O). At this density, the protein would be matched, and the measured radius of gyration of the complex would be that of the DNA alone, R(DNA). Thus, $[R(\text{DNA})]^2$ may be calculated from eq 4, taking R_c(complex) = 32.4 ± 0.3 Å, $\alpha = (-4.15 \pm 0.42) \times 10^{-4}$, and $\bar{\rho} = (0.0276 - 0.0220) \times 10^{-4}$ Å^{-2}. This value for $\bar{\rho}$ is the difference between the measured match point density of the complex containing D-DNA and the calculated match point density of the protein alone. In fact, the contrast, $\bar{\rho}$, for the complex at the match

point of the protein would be slightly less than this value due to the decreased density of the complex at the lower percentage of D_2O that is the match point of the protein. Therefore, the value of R(DNA) obtained will be an undisputed maximum. R(DNA) thus calculated is 17.6 ± 3 Å at 39.7% D_2O, clearly indicating an interior placement of the DNA.

The radius of gyration of the gene 5 protein in the complex can now be calculated knowing the R(DNA) and the R_c for the complex as a whole. At infinite contrast the DNA and protein are indistinguishable, and the radii of gyration of the individual components are related, in proportion to their volumes, to the radius of gyration of the whole complex. Taking the volume of the protein and of a tetranucleotide to be 12 440 Å^3 and 1160 Å^3, respectively, from the volumes of the amino acid residues (Jacrot, 1976) and using half the volume of a base pair (Cohen & Eisenberg, 1968), we calculate the volume of a subunit to be 13 600 Å^3. This is close to the match-point volumes obtained for both complexes (see Results). Thus, we have the relation

$$[R_c(\text{complex})]^2 = \frac{1160}{13\,600}[R_c(\text{DNA})]^2 + \frac{12\,400}{13\,600}[R_c(\text{protein})]^2 \quad (6)$$

With R_c(complex) = 32.4 ± 0.3 Å and R_c(DNA) = 17.6 ± 3 Å, eq 6 yields R_c(protein) = 33.5 ± 0.4 Å, which is only slighter larger than the radius of gyration of the complex as a whole. This calculation is based on the assumption that the radius of gyration of the DNA does not significantly change with contrast. This is very reasonable since the DNA with its exchangeable hydrogens lies within a radius that is narrow compared with the radius of the whole complex. In the subsequent discussion it is also assumed that the radius of gyration obtained for D-DNA is the same as the radius of gyration of normal H-DNA in the complex.

We can make a second estimate of R_c(protein), using eq 4, as a check of the above value. Using measured values of R_c and α for the complex containing D-DNA and a calculated value of $\bar{\rho}$ at the match point of the DNA (which is 108.9% D_2O for 87% deuterated DNA), we obtain R_c(protein) = 34.1 ± 0.4 Å. This is quite close to the above value of R_c(protein) = 33.5 ± 0.4 Å at infinite contrast and confirms the reasonableness of the latter value.

The above values for R_c(DNA) and R_c(protein) are consistent with the small negative α we observe for complex containing normal protonated DNA. If we make the approximation that the radii of gyration of the individual components do not change with contrast and use eq 4 with calculated values of $\bar{\rho}$ at the match points of H-DNA and gene 5 protein, we obtain $\alpha = (-0.85 \pm 0.60) \times 10^{-4}$, a range of values not inconsistent with our measured α of $(-0.2 \pm 0.3) \times 10^{-4}$.

Cross-Sectional View of the fd DNA–Gene 5 Protein Complex. We conclude from our values for R_c(DNA) and R_c(protein) that the DNA is near the center of the complex, bound to the inner surface of the gene 5 protein molecules. This is illustrated in Figure 10, which shows a cross section of the complex with the areas probably occupied by the DNA and protein. In constructing this cross section, we have taken the maximum radial thickness of a DNA strand to be 10 Å and the radial thickness of a gene 5 protein molecule to be 30 Å. (Gene 5 protein would have dimensions of about 25 by 30 Å in the cross-sectional plane, if the length of about 45 Å is assumed to be oriented parallel to the complex axis; McPherson et al., 1979a,b.) The inner and outer radii shown for the DNA and protein have been calculated to be consistent

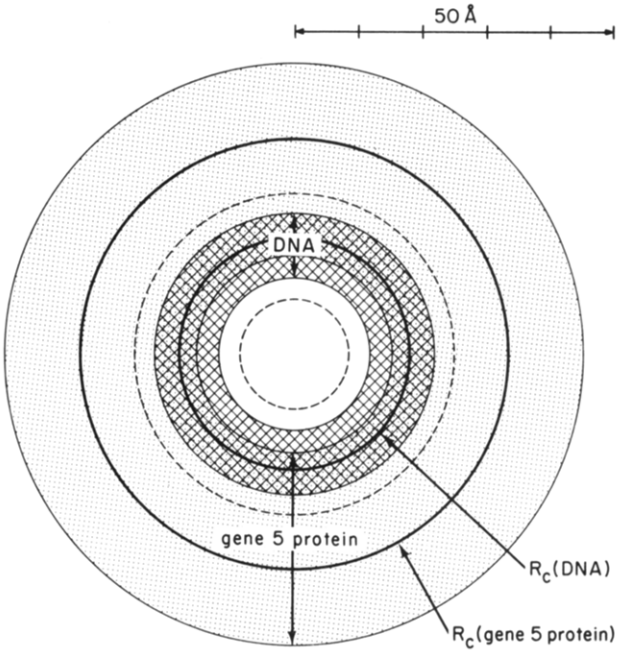

FIGURE 10: Cross section of the complex between DNA and gene 5 protein. R_c(DNA) and R_c(protein) are the radii of gyration at infinite contrast for the DNA and protein components. The cylindrical shell occupied by the DNA (crosshatched) extends from 11.8 to 21.8 Å, and the shell occupied by the protein (stippled) extends from 14.9 to 44.9 Å. The dashed rings show the uncertainty in the DNA shell due to the error of ±3 Å (±1 standard deviation) in our determination of the maximum R_c(DNA) as described in the text.

with the measured R_c values, with the assumption that the components are uniform in density along their radial extents. The extreme outer radius of the protein is near 45 Å, giving a diameter consistent with the 95–105 Å observed by electron microscopy (Gray et al., 1982; Torbet et al., 1981). Figure 10 shows that the DNA overlaps at most the inner 10 Å of the protein and the DNA is clearly not close to the outer periphery of the complex, as in the model proposed by McPherson et al. (1979b, 1980a).

The cross-sectional domains shown in Figure 10 for the DNA and protein are not fully occupied. The values of mass per unit length we obtain correspond to axial subunit repeat distances of 5.9 ± 0.3 and 5.2 ± 0.2 Å for complexes containing H-DNA and D-DNA, respectively (see Table II). The volumes calculated with these repeat distances and an outer radius of 45 Å are 2.75–2.45 times the subunit excluded volume of 13 600 Å^3. Therefore, as noted before by Torbet et al. (1981), the complex has a rather open structure.

The helical complex structure can be formed with circular fd DNA, requiring that two DNA strands be coalesced into one helix as noted by Alberts et al. (1972). The gene 5 protein exists as a dimer in solution (Cavalieri et al., 1976) and in crystals (McPherson et al., 1979a), with the presumed DNA binding sites of the two halves being on the same side of the dimer (McPherson et al., 1979a). Therefore, in constructing Figure 10 we have assumed that both the up and down strands of DNA have the same radius of gyration. If in fact one of the strands were at the axis of the complex, the other could not be further removed in radius than ~25 Å and still give a structure with the measured R_c(DNA). Thus, our data also exclude the possibility that one strand is on the exterior and the other is near the center of the complex.

Relationship to Other Studies. Our present results are supported by previous calculations (Gray et al., 1982; Torbet et al., 1981) of the maximum outer radius of the DNA strands in fd complex, based on geometrical relationships and on the

assumption that both of the DNA strands are at the same radius and have fully extended sugar–phosphate backbones. These calculations use the equation

$$l^2 = (2\pi r)^2 + P^2 \quad (7)$$

where l is the contour distance covered by the DNA in one turn of the helix, r is the radius of the DNA backbone, and P is the helix pitch. If l is a maximum, r will be a maximum for a helix of known pitch. Gray et al. (1982) used parameters derived from electron microscopy to estimate the maximum DNA radius as 34.4 Å. Using the rise per residue from neutron diffraction combined with the pitch observed by electron microscopy, Torbet et al. (1981) estimated the maximum DNA radius to be 33.4 Å. Therefore, even from geometrical considerations, both DNA strands cannot be on the outer periphery of the fd DNA–gene 5 protein complex. Moreover, calculations based on data from electron microscopy clearly demonstrate that the DNA must be located at a smaller radius than the bulk of the protein in an analogous helical nucleoprotein complex recently discovered in *Pseudomonas aeruginosa* bacteria infected by the filamentous virus Pf1 (Gray et al., 1982). These authors show that the Pf1 complex has a radius of about 55 Å but that its DNA has a calculated maximum radius of only 20.5 Å. Our present finding that the DNA is near the center of the fd DNA–gene 5 protein complex is thus in agreement with the geometric calculations for the fd complex and indicates that an interior location of the DNA is a common feature of both known precursor complexes formed during filamentous phage morphogenesis.

The model proposed by McPherson et al. (1979b, 1980a) for the helical complex between DNA and gene 5 protein is based on the hexameric ring of six dimers of protein that these authors infer to exist when the protein is crystallized with oligodeoxynucleotides (McPherson et al., 1979b, 1980a,b) and on the identification of the DNA binding site in the protein monomer and the orientation of the monomers in the hexameric ring. They also assumed that the closed ring becomes a "lockwasher" unit of the helix. Their work has provided a useful, testable model for the solution structure of the complex. Although our data do not confirm the position of the DNA to be as they propose, it may only be that their final assumption (that the hexameric ring is closely related to a helical turn of the complex) is incorrect. If the identification and orientation of the DNA binding site is correct in the hexameric ring of dimers, then it appears that the dimers are in reverse orientation with respect to the ring axis in the hexameric structure, compared with their orientation with respect to the helix axis of the complex in solution. If so, our studies suggest that the structure of the complex in solution may not be easy to infer from X-ray studies of crystals where packing forces are influential.

Our determination that the DNA is near the center of the complex means that the DNA does not have to undergo a great radial displacement during its packaging into mature fd phage particles during viral morphogenesis. The DNA structure is different in the two situations, however. The DNA occupies a core of less than 25 Å in diameter in the phage (Wachtel et al., 1974; Banner et al., 1981). Since the length of the phage particle is 8800–9000 Å (Day & Wiseman, 1978), not much different from the contour length of the complex in solution (Table II), the DNA does become more compacted in volume by about a factor of 2 in the phage. It is also known that the DNA is hypochromic in the phage, whereas it is almost fully hyperchromic in the precursor complex (Day & Wiseman, 1978). Regardless of the detailed structural changes that occur in the DNA, we may now visualize that gene 5 protein molecules are replaced by coat protein molecules on the surface of the DNA genome as it passes through the bacterial membrane.

Acknowledgments

We are especially grateful to Marilyn Vaughan, whose expertise was crucial throughout the preparation of the complexes. We thank Dr. Benno Schoenborn (Brookhaven National Laboratory) for his encouragement to undertake this work and Dr. Ben Shoulders (University of Texas at Austin) for the NMR spectra of the H- and D-DNAs. We also appreciate a critical reading of the manuscript by Dr. James Torbet (Max-Planck-Institut für Festkörperforschung, Grenoble, France).

References

Alberts, B., Frey, L., & Delius, H. (1972) *J. Mol. Biol. 68*, 139–152.

Anderson, R. A., Nakashima, Y., & Coleman, J. E. (1975) *Biochemistry 14*, 907–916.

Banner, D. W., Nave, C., & Marvin, D. A. (1981) *Nature (London) 289*, 814–816.

Beck, E., Sommer, R., Auerswald, E. A., Kurz, Ch., Zink, B., Osterberg, G., Shaller, H., Sugimoto, K., Sugisaki, H., Okamoto, T., & Takanami, M. (1978) *Nucleic Acids Res. 5*, 4495–4503.

Berkowitz, S. A., & Day, L. A. (1974) *Biochemistry 13*, 4825–4831.

Cavalieri, S. J., Neet, K. E., & Goldthwait, D. A. (1976) *J. Mol. Biol. 102*, 697–711.

Charlesby, A. (1954) *Proc. R. Soc. London, Ser. A 224*, 120–128.

Cohen, G., & Eisenberg, H. (1968) *Biopolymers 6*, 1077–1100.

Crespi, H. L., & Katz, J. J. (1961) *Anal. Biochem. 2*, 274–279.

Day, L. A. (1973) *Biochemistry 12*, 5329–5339.

Day, L. A., & Wiseman, R. L. (1978) in *The Single-Stranded DNA Phages* (Denhardt, D. T., Dressler, D., & Ray, D. S., Eds.) Cold Spring Harbor Laboratory Monogr. Ser., pp 605–625, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Fang, K. N., Kondo, N. S., Miller, P. S., & Ts'o, P. O. P. (1971) *J. Am. Chem. Soc. 93*, 6647–6656.

Glasoe, P. K., & Long, F. A. (1960) *J. Phys. Chem. 64*, 188–190.

Gray, C. W., Kneale, G. G., Leonard, K. R., Siegrist, H., & Marvin, D. A. (1982) *Virology 116*, 40–52.

Ibel, K., & Stuhrmann, H. B. (1975) *J. Mol. Biol. 93*, 255–265.

Jacrot, B. (1976) *Rep. Prog. Phys. 39*, 911–953.

Kowalczykowski, S. C., Lonberg, N., Newport, J. W., & von Hippel, P. H. (1981) *J. Mol. Biol. 145*, 75–104.

Kratky, O., & Pilz, I. (1972) *Q. Rev. Biophys. 5*, 481–537.

Malmon, A. G. (1957) *Acta Crystallogr. 10*, 639–642.

Marvin, D. A., & Schaller, H. (1966) *J. Mol. Biol. 15*, 1–7.

McPherson, A., Jurnak, F. A., Wang, A. H. J., Molineux, I., & Rich, A. (1979a) *J. Mol. Biol. 134*, 379–400.

McPherson, A., Jurnak, F. A., Wang, A. H. J., Kolpak, F., Molineux, I., & Rich, A. (1979b) *Cold Spring Harbor Symp. Quant. Biol. 43*, 21–28.

McPherson, A., Jurnak, F., Wang, A., Kolpak, F., Rich, A., Molineux, I., & Fitzgerald, P. (1980a) *Biophys. J. 32*, 155–170.

McPherson, A., Wang, A. H. J., Jurnak, F. A., Molineux, I., Kolpak, F., & Rich, A. (1980b) *J. Biol. Chem. 255*, 3174–3177.

Moore, P. B., & Engelman, D. M. (1976) *Brookhaven Symp. Biol. No. 27*, 12–23.
Nakashima, Y., Dunker, A. K., Marvin, D. A., & Konigsberg, W. (1974) *FEBS Lett. 40*, 290–292; erratum *43*, 125.
Oey, J. L., & Knippers, R. (1972) *J. Mol. Biol. 68*, 125–138.
Osborn, M., Weiner, A. M., & Weber, K. (1970) *Eur. J. Biochem. 17*, 63–67.
Patel, D. J. (1977) *Biopolymers 16*, 1635–1656.
Patel, D. J., & Canuel, L. L. (1979) *Eur. J. Biochem. 96*, 267–276.
Pretorius, H. T., Klein, M., & Day, L. A. (1975) *J. Biol. Chem. 250*, 9262–9269.
Ray, D. S. (1978) in *The Single-Stranded DNA Phages* (Denhardt, D. T., Dressler, D., & Ray, D. S., Eds.) Cold Spring Harbor Laboratory Monogr. Ser., pp 325–339, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Schoenborn, B. P., Alberi, J., Saxena, A. M., & Fischer, J. (1978) *J. Appl. Crystallogr. 11*, 455–460.
Torbet, J., Gray, D. M., Gray, C. W., Marvin, D. A., & Siegrist, H. (1981) *J. Mol. Biol. 146*, 305–320.
Valentine, R. C., Shapiro, B. M., & Stadtman, E. R. (1968) *Biochemistry 7*, 2143–2152.
Wachtel, E. J., Wiseman, R. L., Pigram, W. J., Marvin, D. A., & Manuelidis, L. (1974) *J. Mol. Biol. 88*, 601–618.
Webster, R. E., & Cashman, J. S. (1978) in *The Single-Stranded DNA Phages* (Denhardt, D. T., Dressler, D., & Ray, D. S., Eds.) Cold Spring Harbor Laboratory Monogr. Ser., pp 557–569, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Both Parental Deoxyribonucleic Acid Strands at Each Replication Fork of Replicating Simian Virus 40 Chromosomes Are Cut by a Single-Strand-Specific Endonuclease†

Yoriaki Tsubota,‡ M. Anwar Waqar, Leslie R. Davis, Loretta Spotila, and Joel A. Huberman*

ABSTRACT: We have measured the relative accessibility to a single-strand-specific endonuclease of the single-stranded DNA on the leading and lagging sides of replication forks in replicating simian virus 40 (SV40) chromosomes. To do this we have digested replicating SV40 chromosomes with a single-strand-specific endonuclease (P1 nuclease) and then characterized the intermediate and final products of digestion by sucrose gradient sedimentation and agarose gel electrophoresis. P1 nuclease rapidly and specifically cleaves parental DNA strands at replication forks, yielding intermediate and final cleavage products which are consistent with an approximately equal rate of nuclease cleavage on both sides of the fork. Thus, single-stranded DNA is approximately as accessible to P1 nuclease on the leading side of the fork as on the lagging side; the simplest interpretation of this observation is that the stretch of single-stranded DNA on the leading side is as long as that on the lagging side.

When the two parental DNA strands unwind from each other at a replication fork, double-stranded DNA is converted to single-stranded DNA. On the lagging side of the replication fork (the side on which the direction of nascent strand synthesis is opposite the direction of fork movement), the length of single-stranded stretch depends on the frequency of initiation of Okazaki pieces (see Figure 1). Available evidence suggests that new Okazaki pieces are started at irregular intervals averaging 135 nucleotides, with a maximum interval of 290 nucleotides [reviewed by DePamphilis & Wassarman (1980)]. Thus a single-stranded stretch as large as 290 nucleotides may be found in front of or behind each Okazaki piece (Figure 1). On the leading side of the fork (the side on which nascent strand synthesis takes place in the same direction as fork movement), the length of single-stranded stretch is determined by the distance from the 3′ end of the nascent strand to the point at which parental strands separate (Figure 1). Nothing is known about this aspect of replication fork structure.

Herman et al. (1979) have previously shown that both deproteinized replicating simian virus 40 (SV40)[1] DNA and the DNA in replicating SV40 chromosomes are hydrolyzed by S1 nuclease (a single-strand-specific nuclease from *Aspergillus oryzae*) at least 10 000-fold more rapidly than are the corresponding form I, nonreplicating DNAs (deproteinized or in chromosomes), but the replicating DNA in chromosomes is cut about 200-fold more slowly than the deproteinized replicating DNA. Thus single-stranded stretches can be detected in replicating DNA, and when the replicating DNA is associated with chromosomal proteins, these stretches are partially protected from nuclease action. Herman et al. (1979) did not distinguish between hydrolysis of the leading side of the fork and hydrolysis of the lagging side.

In the studies described in this paper, we have analyzed the intermediate and final products of cleavage of SV40 chromosomes by P1 nuclease, a single-strand-specific nuclease from *Penicillium citrinum* (Fujimoto et al., 1974). This nuclease has a higher pH optimum than S1 nuclease. By using P1 nuclease, we were able to carry out nuclease digestions at our standard pH for SV40 chromosome preparations (pH 6.8).

† From the Department of Cell and Tumor Biology, Roswell Park Memorial Institute, Buffalo, New York 14263. *Received November 18, 1981.* This work was supported by National Science Foundation Grants PCM77-14451 and PCM80-11246.

‡ Present address: Kyushu Chuo Hospital, 882 Shiobaru, Minami-ku, Fukuoka 815, Japan.

[1] Abbreviations: SV40, simian virus 40; RI DNA, replicating intermediate DNA; form I DNA, covalently closed circular, supercoiled DNA; form II DNA, circular DNA containing one or more nicks or small gaps; form III DNA, linear duplex DNA of full genome length; EDTA, ethylenediaminetetraacetic acid; $NaDodSO_4$, sodium dodecyl sulfate; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; PMSF, phenylmethanesulfonyl fluoride; Tris, tris(hydroxymethyl)-aminomethane.

0006-2960/82/0421-2713$01.25/0