Molecular Weight of Single-Stranded fd Bacteriophage DNA. High Speed Equilibrium Sedimentation and Light Scattering Measurements†

Steven A. Berkowitz and Loren A. Day*

ABSTRACT: The molecular weight of circular single-stranded fd NaDNA has been obtained by high speed equilibrium sedimentation, $(1.92 \pm 0.06) \times 10^6$, and by light scattering, $(1.96 \pm 0.12) \times 10^6$. The weighted average of these values with an additional value from sedimentation-diffusion measurements (J. Newman, H. L. Swinney, S. A. Berkowitz, and L. A. Day (1974), *Biochemistry 13,* 4832) is $(1.90 \pm 0.07) \times 10^6$, which corresponds to 5740 ± 210 nucleotides. The radius of gyration, R_G, and second virial coefficient, B, in 0.15 M NaCl–0.015 M sodium citrate (pH 8) (SSC) are 41.6 ± 3.5 nm and $(0.86 \pm 0.44) \times 10^{-4}$ g^{-2} cm^3 mol, respectively. The molecular weights are based on a density increment, $(\partial\rho/\partial c)_\mu$, of 0.483 ± 0.010, determined by resonant oscillator densimetry, and refractive index increments, $(\partial n/\partial c)_\mu$, 0.175 ± 0.019 g^{-1} cm^3 at 436 nm and of 0.170 ± 0.015 g^{-1} cm^3 at 546 nm, determined by differential refractometry, for fd NaDNA in SSC. The three molecular weight determinations are absolute and independent; only the uncertainty in concentration (1.4%) from phosphorus analysis is common to all methods. The uncertainties quoted are 95% confidence limits.

The many molecular weights reported for fd bacteriophage DNA span a large range of values (for a review, see Marvin and Hohn, 1969). We have undertaken an accurate determination of this parameter for its use in studies of filamentous phages and to help establish reliable DNA molecular weight standards for other purposes. We were also interested in investigating the shape and interactions of fd DNA in solution.

Only absolute methods for determining the molecular weight of macromolecules were considered satisfactory for our needs. Empirical equations which relate DNA weight to sedimentation velocity or to intrinsic viscosity, such as those of Studier (1965), Eigner and Doty (1965), and Crothers and Zimm (1965), even if based on reliable molecular weight standards, might not lead to dependable weights in a given application because of the possible effects of base composition on the density increment and topology on the frictional coefficient. Molecular weight determinations by electron microscopy from contour lengths of DNA molecules prepared for microscopy by monomolecular protein film techniques (Kleinschmidt and Zahn, 1959) are subject to uncertainties in the mass per unit length as discussed by Inman (1967), Lang (1970), and Freifelder (1970).

In this study we have measured the molecular weight of fd DNA by two independent absolute thermodynamic methods, high speed equilibrium sedimentation and light scattering. There are no theoretical assumptions in the equations which relate molecular weight to the experimental parameters involved in these techniques, and all of the necessary specific parameters for fd DNA have been determined in this study. In an accompanying hydrodynamic study (Newman *et al.*, 1974), a third independent absolute value was obtained from the Svedberg equation. In addition, the dependence of the thermodynamic and hydrodynamic properties on concentration has been investigated.

Materials and Methods[1,2]

(*a*) *DNA Preparation and Concentration Measurement.* The method of preparing fd bacteriophage was similar to that of Yamamoto *et al.* (1970). Purified virus solutions had plating efficiencies of 60–80% and were homogeneous according to analytical density gradient equilibrium centrifugation and SDS[3] gel electrophoresis. DNA was obtained by three phenol extractions at room temperature. Each virus and nucleic acid preparation was monitored by band sedimentation for the fraction of intact circular, singly nicked, and multiply nicked molecules of fd DNA. To minimize nuclease degradation autoclaved buffers and glassware were used and dialysis tubing was boiled in 10^{-3} M EDTA.

The concentration of fd DNA was calculated from the concentration of phosphorus, determined according to Chen *et al.* (1956), and an average sodium nucleotide weight of 331.0 ± 0.2 (Hoffmann-Berling *et al.*, 1963; Schaller *et al.*, 1969; Wiseman *et al.*, 1972). Optical density measurements were made on calibrated Zeiss PQII and Cary 14 spectrophotometers. The extinction coefficient of fd DNA in SSC (0.15 M NaCl-0.015 M sodium citrate (pH 8)), was found to be $\epsilon(P) = 7370 \pm 110$ cm^{-1} M^{-1} (95% C.L.) at

† From the Department of Biochemistry, The Public Health Research Institute of the City of New York, New York, New York 10016. *Received July 11, 1974.* The work was supported by Grants AI-09049 and 1K04-GM-70363 from the U. S. Public Health Service. One of us, S.A.B., was supported by a U. S. Public Health Service Training Grant GM 01234 in the Department of Biochemistry, New York University School of Medicine.

[1] For further details see Berkowitz (1974).

[2] Unless otherwise indicated, uncertainties are quoted at the 95% confidence limit given by the statistical fluctuation in the data. These have been obtained from the standard error of the mean and tables for the t distribution. For those parameters we have not measured, we have tried to estimate the uncertainty at this level of confidence.

[3] Abbreviations used are: SDS, sodium dodecyl sulfate; SD, standard deviation; 95% C.L., 95% confidence limit; EDTA, ethylenediaminetetraacetate (Na salt).

259 nm, where the largest uncertainty arises from the calibration curve of moles of phosphorus *vs.* OD_{827}.

(*b*) *Specific Refractive Index Increment and Specific Density Increment.* Samples were brought to chemical equilibrium of diffusible components by dialysis against SSC for 72 hr at 5° and 12 hr at room temperature. The densities of inner and outer dialysis solutions were measured (at 25.0° controlled to ±0.01°) with a resonant oscillator densimeter (Kratky *et al.*, 1973) manufactured by Paar Instruments, Graz, Austria. The specific refractive index increment of fd DNA, $(\partial n/\partial c)_\mu$, in SSC at 25.0 ± 0.1° was measured at 436 and 546 nm on a Brice-Phoenix differential refractometer calibrated with KCl solutions. After density and refractive index measurements had been made on a given solution, measurements were made of its phosphorus content and uv absorption. All dilutions were by weight.

(*c*) *Analytical Sedimentation.* A Beckman Model E analytical ultracentrifuge with ultraviolet absorption optics (scanner) and interference optics was used. The distances of the counterbalance reference holes to the rotor axis were determined with a microcomparator.

The intactness of fd DNA was monitored by band sedimentation (Vinograd *et al.*, 1963) in 4 M CsCl–0.02 M Na_3PO_4–10^{-3} M EDTA–0.04 N NaOH (pH 13). In several cases analysis was made with 0.015 M sodium citrate in 99.3% D_2O or with 4 M CsCl–0.02 M sodium phosphate–10^{-3} M EDTA (pH 6.9) after reacting the DNA with 1.5% formaldehyde in SSC at room temperature for 20 hr.

High speed sedimentation equilibrium runs (Yphantis, 1964) were made with an AN-J rotor using 12-mm double sector and 30-mm single sector cells (Schachman and Edelstein, 1966). To minimize nuclease contamination, centerpieces were soaked in 95% ethanol for 1 hr and rinsed with sterile water. The speed was 3400 rpm, which gave apparent reduced molecular weights, σ, between 4 and 5, and total run times were about 72 hr. The temperature was 19 ± 1°; the temperature regulation system was not used in order to minimize convection. Experiments with initial DNA concentrations varying from 5 to 25 μg cm^{-3} were done with the uv scanner. Curves relating pen deflection to actual absorbance were obtained from least-squares quadratic fits of measurements on solutions of adenine and fd DNA at the rotor speeds and wavelengths used for the equilibrium experiments. At the end of some runs, the rotor was accelerated to 15,000 rpm for 12 hr, decelerated to the run speed, and scanned for a base line. In all cases the base line corresponded to that obtained from data in the meniscus region in flatness and amplitude. Experiments with initial concentrations varying from 170 to 750 μg cm^{-3} were performed with interference optics. Base-line patterns, used to determine zero concentration levels, were photographed immediately upon reaching speed and at the end of the run after the rotor was stopped, shaken, and accelerated again to the equilibrium speed. Plate readings were made on the microcomparator by averaging the readings for three fringes, which gave an average standard deviation of ±5 μm. Net fringe displacements, $j(r)$, at distance r were converted to concentration in g cm^{-3} by $c = \lambda j(r)/(\partial n/\partial c)_\mu d$, where λ is the wavelength and d is the path length through the cells.

Point reduced weight-average molecular weights, $\sigma_w(r)$, are

$$\sigma_w(r) \equiv \mathrm{d}\ \ln\ c(r)/\mathrm{d}(r^2/2) \qquad (1)$$

where $c(r)$ is the concentration at distance r from the center of rotation. Values for $\sigma_w(r)$ were obtained from fitting routines spanning anywhere from 5 to 21 points centered at r. Only points having net displacements greater than 100 μm or absorbancies above 0.10 OD were used. The apparent reduced number average weights were calculated as

$$\sigma_n(r) = c(r)/\left[\int_{c(m)}^{c(r)} \mathrm{d}c(r)/\sigma_w(r) + c(\mathrm{m})/\sigma_n(\mathrm{m})\right] \qquad (2)$$

The integral was numerically evaluated over adjacent points; the concentration at the meniscus, c(m), was determined from data in the meniscus region, and the apparent reduced number-average weight at the meniscus, σ_n(m), was set equal to σ_w^0 (Yphantis, 1964; Teller, 1973).

The calculations were done by computer with two programs, one which initially smoothed the data (Roark and Yphantis, 1969) and one which left it unsmoothed (Berkowitz, 1974).

The apparent reduced molecular weight moments are related to their values at zero concentration by

$$\begin{aligned} \sigma_n^0/\sigma_n &= 1 + BM_n c + \ldots \\ \sigma_w^0/\sigma_w &= 1 + 2BM_w c + \ldots \\ \sigma_z^0/\sigma_z &= 1 + 4BM_z c + \ldots \end{aligned} \qquad (3)$$

where σ_n^0, σ_w^0, σ_z^0 are the ideal reduced number, weight, and z average molecular weights, respectively, and B is the second virial coefficient. Values for these quantities were obtained by weighted linear least-squares analyses of $1/\sigma$ *vs.* c. Molecular weights determined from equilibrium sedimentation were calculated from their ideal reduced quantities using

$$M_x = \sigma_x^0 RT/(\partial\rho/\partial c)_\mu \omega^2 \qquad (4)$$

where x = n, w, or z and ω is the angular velocity. Casassa and Eisenberg (1964) showed that the complicating effects of preferential interactions between solution components can be eliminated by substituting the density increment $(\partial\rho/\partial c)_\mu$ for the usual buoyancy term. The values of the molecular weight and second virial coefficient are for that species defined by the concentration units used. In this study, all concentrations were based on absolute molar concentrations of phosphorus. By assigning for each phosphorus 331 daltons, the average nucleotide residue weight of fd DNA plus the atomic weight of sodium, one obtains the mass of a hypothetical DNA molecule with one Na^+ per nucleotide.

(*d*) *Light Scattering.* An instrument of the type designed by Wippler and Scheibling (1954) and manufactured by Sofica (now Fica) was used. Measurements were made at 436 and 546 nm at 25.0 ± 0.1° with unpolarized light. Optical alignment and accuracy of angle settings were tested by recording the angular dependence of fluorescein fluorescence; the angular dissymmetry was 1% or less for $30° \leq \theta \leq 150°$. The intensity calibration was based on measurements of benzene at 90°, our primary standard. Its Rayleigh ratio was taken as 16.0×10^{-6} cm^{-1} at 546 nm and as 46.4×10^{-6} cm^{-1} at 436 nm (Kratohvil *et al.*, 1962; Berkowitz, 1974). A glass rod supplied by the manufacturer was calibrated against benzene and used as a working standard. Stray light was reduced with a 3 × 15 mm slit in the incident beam placed directly in front of the cell. Optical cells were thin wall (1 mm) colorimetric comparator tubes made of flint glass having an outside diameter of 32 mm. The cells were fitted with machined collars for proper sit-

ting in the instrument and with stoppers which had entrance and exit ports for solutions. Samples were filtered into the cell through cellulose triacetate membrane filters, 0.2 or 0.45 μm pore size (Gelman Instrument Company), under hydrostatic heads of 10–15 cm. Values of R_{90} for 10^{-3} M NaCl as low as 1.0×10^{-6} cm^{-1} at 546 nm with a dissymmetry (R_{30}/R_{150}) of 1.07 were obtained. Routinely buffer solutions had R_{90} values of $1.1–1.3 \times 10^{-6}$ cm^{-1} and dissymmetries of 1.1–1.2; R_{90} values for DNA solutions ranged from 4 to 32×10^{-6} for the concentration range 5–40 μg cm^{-3}. Measurements began with the highest concentration; samples were recovered, diluted, and then refiltered for measurements at lower concentrations. This caused no change in the fraction of circular fd DNA.

The calculations of the Rayleigh ratio for fd DNA solutions, at angle θ and concentration c, were made according to

$$R_{\theta,c} = [\sin\theta/(1 + \cos^2\theta)](R_B/i_{90,gs})\alpha(n_0/n_B)^2 I_{\theta,c} \quad (5)$$

where $\sin\theta/(1 + \cos^2\theta)$ accounts for the scattering volume and unpolarized light, R_B is the Rayleigh ratio for benzene, $i_{90,gs}$ is the reading for the glass standard, α is the intensity ratio for the primary and secondary standards, $(n_0/n_B)^2$ corrects for the difference in the refractive indices of the solvent, n_0, and benzene, n_B, and $I_{\theta,c}$ is the difference intensity between the DNA solution and buffer, corrected for a very small contribution from reflections (Tomimatsu *et al.*, 1968). The apparent Rayleigh ratios were analyzed with the double extrapolation procedure of Zimm (1948) based on eq 6 and 7

$$Kc/R_{\theta=0,c} = (1/M_w)(1 + 2BM_w c + \ldots) \quad (6)$$

$$Kc/R_{\theta,c=0} = (1/M_w)[1 + (4\pi n_0/\lambda_0)^2\{R_G\}^2 \sin^2(\theta/2) + \ldots] \quad (7)$$

where λ_0 is the wavelength in vacuum, R_G is the z average radius of gyration, and $K = 2\pi^2(\partial n/\partial c)_\mu^2/N_a\lambda_0^4$, where N_a is Avogadro's number. M_w, R_G, and B were obtained from least-squares analyses of Kc/R_θ as functions of c and θ.

Results

(*a*) *Physical State of fd DNA.* In the fd bacteriophage preparations made, between 3 and 8% of the virions contained linear DNA, as determined by band sedimentation in alkaline CsCl (Figure 1). These linear DNA molecules did not result from our phage purification procedure since the amount of linear DNA found after the first and final purification steps of the bacteriophage was the same. The agitation of the growing culture did not cause nicking of circular DNA since a nonagitated culture showed 8% linear DNA after the first purification step. Analysis of purified DNA in neutral CsCl after formaldehyde treatment showed the same fraction of linear molecules as analysis of the original phase preparation in alkaline CsCl. Hence, neither the DNA extraction procedure nor alkaline pH caused cleavage. The amount of linear DNA could not be reduced by alteration of growth conditions and purification techniques (Berkowitz, 1974).

Storage of fd DNA and fd phage can lead to an increase in the fraction of linear DNA originally present. As a result, the fraction of linear DNA in the preparations used in this and the accompanying paper (Newman *et al.*, 1974) was between 3 and 18%.

(*b*) *Density Increment and Refractive Index Increment.*

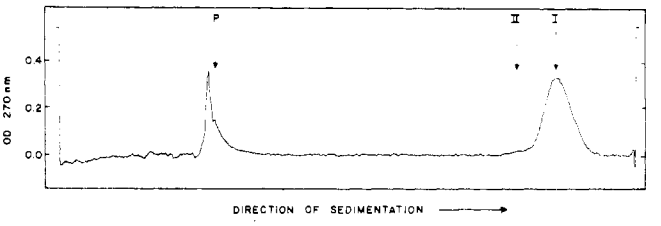

FIGURE 1: Band sedimentation scan of fd in 4 M CsCl–0.02 M Na_3PO_4–10^{-3} M EDTA–0.04 N NaOH (pH 13) at 40,000 rpm. Protein (P) is stripped from the DNA and remains at the meniscus due to the high salt density. DNA is resolved into two bands corresponding to intact circular molecules (I) and singly nicked linear molecules (II). This scan shows the presence of about 5% nicked fd DNA.

Measurements of difference densities in SSC at constant chemical potential of diffusible components, $(\Delta\rho)_\mu = \rho - \rho_0$, where ρ and ρ_0 are the densities of the inner and outer dialysis solutions, respectively, gave linear plots of $(\Delta\rho)_\mu$ *vs.* concentration over the range 1–6.5 mg cm^{-3}. Hence, the slope of this line was taken as the specific density increment at zero concentration, $(\partial\rho/\partial c)_\mu$. Data were also gathered for fd NaDNA in various dilutions of 10 X SSC (1.5 M NaCl–0.15 M sodium citrate (pH 8)). The results are presented in Table I. The value in SSC is about 5% higher than the value reported by Cohen and Eisenberg (1968) for native calf thymus NaDNA in 0.2 M NaCl. This difference is not unreasonable since the buoyant density of single-stranded DNA is greater than that of double-stranded DNA (Meselson and Stahl, 1958). Differences in density increments can also be due to differences in base composition, as demonstrated by measurements on synthetic polynucleotides (Eisenberg and Felsenfeld, 1967; Inners and Felsenfeld, 1970).

A plot of $(\Delta n)_\mu = n - n_0$ *vs.* c was also linear over the concentration range 0.7–3.7 mg cm^{-3}, giving values of $(\partial n/\partial c)_\mu$ for fd NaDNA of $0.175 \pm 0.0.9$ g^{-1} cm^3 (95% C.L.) at 436 nm and 0.170 ± 0.015 g^{-1} cm^3 (95% C.L.) at 546 nm. These values agree with those obtained by Cohen and Eisenberg (1968) for native calf thymus NaDNA and by Krasna *et al.* (1970) and Krasna (1970) for native and denatured calf thymus NaDNA.

(*c*) *Equilibrium Sedimentation.* Representative plots of reduced molecular weights as a function of concentration are given in Figure 2. Consistent with other tests of the Roark and Yphantis program (D. A. Yphantis, private communication), these data show that the smoothing procedures used introduce little, if any, distortion of the data while considerably reducing the level of noise. All parameters calculated with our program and with that of Roark and Yphantis agreed within the computed error limits.

Table II lists values for the various molecular weights and the weight-average virial coefficient obtained from eq 3 and 4. Beyond the random errors indicated in Table II, the only independent variable contributing significantly to the uncertainty in M_x is $(\partial\rho/\partial c)_\mu$, which has a 95% confidence limit of 2.0%. The averaged M_x and their total uncertainties at the 95% confidence limit are shown at the bottom of Table II. Values for the second virial coefficient obtained from reduced number and z averages were $(0.92 \pm 0.16) \times 10^{-4}$ g^{-2} cm^3 mol (SD) and $(0.88 \pm 0.32) \times 10^{-4}$ g^{-2} cm^3 mol (SD), respectively.

(*d*) *Light Scattering.* A typical Zimm plot is given in Figure 3 and the averaged results of five experiments, each done at two different wavelengths, are given in Table III. Only the leading correction terms to eq 6 and 7 were found

TABLE I: Density Increment Values of fd NaDNA in Different Solvents.[a]

	0.01 × SSC	0.1 × SSC	1 × SSC	2.5 × SSC	5 × SSC	7.5 × SSC	10 × SSC
Density (g cm^{-3})	0.9972	0.9982	1.0060	1.0189	1.0309	1.0606	1.0810
Density increment	0.493 ± 0.003	0.484	0.483 ± 0.003[b]	0.435 ± 0.002	0.400	0.363	0.359
No. of experiments	3	1	12	2	1	1	1

[a] Errors indicated are standard deviations. [b] Weighted average of mean obtained by pycnometry, 0.482 ± 0.016, with mean obtained from resonant oscillator densimetry; the uncertainty at the 95% confidence limit is ±0.010 which includes the uncertainty in the extinction coefficient.

TABLE II: Equilibrium Sedimentation Results for fd NaDNA in SSC.

Concentration Ranges Covered at Equilibrium	No. of Experiments	M_n (10^6)	M_w (10^6)	M_z (10^6)	B (10^{-4} g^{-2} cm^3 mol)
4–400 μg cm^{-3} (Scanner)[a]	13	1.85 ± 0.05	1.92 ± 0.03[b]	1.89 ± 0.05[c]	
100–2500 μg cm^{-3} (Interference Optics)[a]	6	1.96 ± 0.06	1.93 ± 0.08	1.94 ± 0.13	0.86 ± 0.17
Weighted averages		1.89 ± 0.13	1.92 ± 0.06	1.90 ± 0.09	0.86 ± 0.44

[a] Errors given in the first two rows are standard deviations from the uncertainty in σ only. Errors given for the weighted averages are at 95% C.L. and include the uncertainty from $(\partial\rho/\partial c)_\mu$. [b] A second virial contribution at low concentration should only be about 0.25–1.0%. The negligible effect of the second virial with scanner data is demonstrated by the fact that the weight-average molecular weight obtained from linear fits of data points above 0.10 O.D. is (1.93 ± 0.01) × 10^6 (SD), which agrees with that from interference data extrapolated to zero concentration. Since there are several ways for calculating the weight-average molecular weight from a given set of data, we quote a weighted average which includes the weight-average value just given, the extrapolated value of $M_w(r)$ to zero concentration, (1.81 ± 0.03) × 10^6 (SD), the extrapolated value of $M_n(r)$ to the base of the cell (1.89 ± 0.24) × 10^6 (SD), and the value from the midpoint moment (1.92 ± 0.03) × 10^6 (SD). Each of these was itself an average from the 13 experiments. [c] We have taken the weighted average of the z average molecular weights obtained by various treatments of the data. This has included the value from extrapolation of $M_z(r)$ to zero concentration, (1.81 ± 0.04) × 10^6 (SD), the value from extrapolation of $M_w(r)$ to the base of cell (2.04 ± 0.05) × 10^6 (SD), the midpoint moment value (1.91 ± 0.03) × 10^6 (SD), and the higher moment value obtained according to Nazarian (1968), (1.91 ± 0.06) × 10^6. The latter value is a complex molecular weight moment higher than the z average moment at the midpoint of the data points.

to be necessary according to statistical F tests of quadratic *vs.* linear fits. This was expected from the low concentrations used and the size of fd DNA in SSC. If dust and other contaminants were contributing to the scattering, calculations in which the highest or lowest angles were removed would give statistically different molecular weights and R_G values; analyses of our data with the three lowest and then the three highest angles deleted gave values which agreed with the values in Table III, and showed no indication of systematic differences. At the low concentrations used, reliable second virial coefficients could not be obtained.

Discussion

(*a*) *State of fd DNA.* A small fraction of virions in a given fd preparation contains linear fd DNA. Similar findings for φX174 bacteriophage have been reported by Schekman and Ray (1971) and Iwaya *et al.* (1973). In addition, Johnson and Sinsheimer (1974) found two types of φX174 viral DNA present in host cells late in infection. One type had a single nick and the second type contained a small gap. Thus, the packaging of fd DNA molecules of different masses could be possible. Since storage of fd DNA leads to molecules with the same sedimentation properties as linear molecules present in phage particles, the naturally occurring nick in fd DNA cannot be a significant gap of missing nucleotides. Our DNA samples are thus monodisperse with respect to mass, as substantiated by the agreement in M_w, M_n, and M_z (Table II). Although intact and singly nicked molecules are essentially of equal mass, their random conformation can differ, leading to differences in frictional and virial coefficients. This difference for single-stranded DNA can be minimized with neutral buffers of high ionic strength (>0.1 M). In SSC, the two forms are not resolved by sedimentation in the analytical ultracentrifuge. Hence, the DNA samples in SSC can also be considered highly monodisperse with respect to size and shape.

(*b*) *Molecular Weight of fd DNA.* In this and the accompanying study, the molecular weight of fd DNA was obtained by three absolute and independent methods: (1.92 ± 0.06) × 10^6 from equilibrium sedimentation, (1.96 ± 0.12) × 10^6 from light scattering, and (1.87 ± 0.06) × 10^6 from sedimentation–diffusion (Newman *et al.*, 1974). The uncertainties cited for each molecular weight include the statistical and systematic uncertainties in each parameter involved. In the case of light scattering, uncertainties in R_B and $(n_0/n_B)^2$, which we did not measure, were estimated at the 95% C.L. A normal propagation of errors from Gaussian statistical fluctuations was assumed.

No assumptions are involved in the relationships used to obtain molecular weights by the thermodynamic techniques (eq 4, 6, and 7). It is assumed in the use of the Svedberg equation that the pressure dependence of $(\partial\rho/\partial c)_\mu$ is negligible and that the change in the chemical potential of solvent components is zero during boundary sedimentation (Eisenberg, 1962). Hence, we consider the different measurements to be of the same quantity, the anhydrous mass

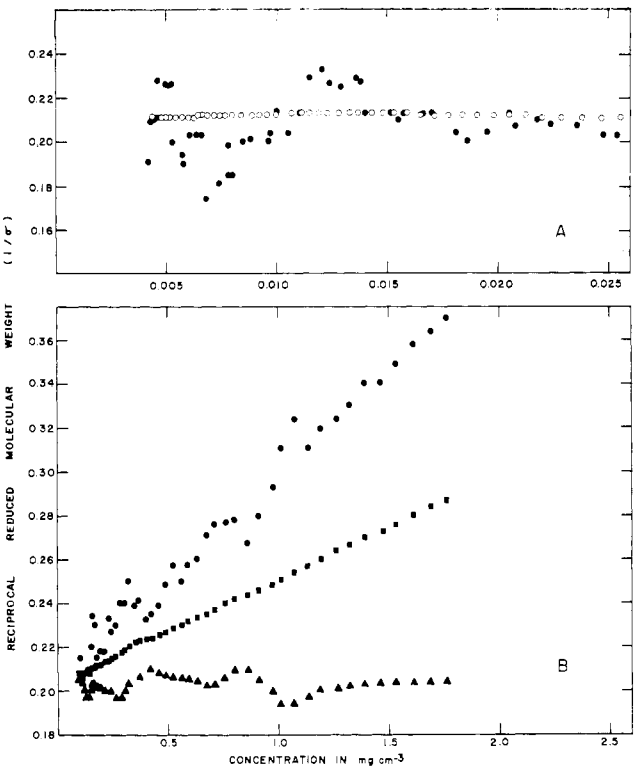

FIGURE 2: (A) Reciprocal reduced weight average molecular weights, σ_w^{-1}, as a function of concentration from an equilibrium sedimentation experiment at an initial concentration of 5 μg cm^{-3}. Open symbols, analysis with smoothing routines; solid symbols, analysis without smoothing routines. (B) Reciprocal reduced molecular weight averages, σ_n^{-1} (■), σ_w^{-1} (●), and $(2\sigma_n^{-1} - \sigma_w^{-1})$ (▲) (Yphantis and Roark, 1972), as a function of concentration in SSC for an interference experiment at an initial concentration of 0.5 mg cm^{-3}.

TABLE III: Light Scattering Results for fd NaDNA in SSC.

Wavelength (nm)[a]	M_w (10^6)	R_G (nm)
436	2.02 ± 0.24	43.3 ± 2.3
546	1.93 ± 0.18	39.8 ± 2.1
Weighted averages	1.96 ± 0.12[b]	41.6 ± 3.5

[a] Errors, quoted at 95% confidence limits, include statistical and systematic uncertainties specific to each wavelength. These include the uncertainties in R_B (Kratohvil *et al.*, 1962; Berkowitz, 1974) which introduce 2.5 and 2.0% uncertainty at 436 and 546 nm, respectively, and in $(\partial n/\partial c)_\mu$, which introduce 10 and 8.3% uncertainty at 436 and 546 nm, respectively, into the molecular weight. The uncertainties in R_G derive only from the statistical uncertainties in the slopes and intercepts of the Zimm plots. [b] In addition to the statistical uncertainty in the weighted mean of M_w, contributions from systematic uncertainties common to both wavelengths have been included. These arise from the uncertainty in the extinction coefficient, which leads to an uncertainty in the molecular weight of 1.4%, and from uncertainty in the exponent in $(n_0/n_B)^2$, which, from the recent work of Wallace *et al.* (1972), contributes at most a 3% error in the molecular weight.

of the hypothetical entity, the fd DNA anion with one Na^+ per nucleotide. We have examined the three independent sets of data to determine whether or not there is evidence of systematic error in one of the techniques. Following Birge

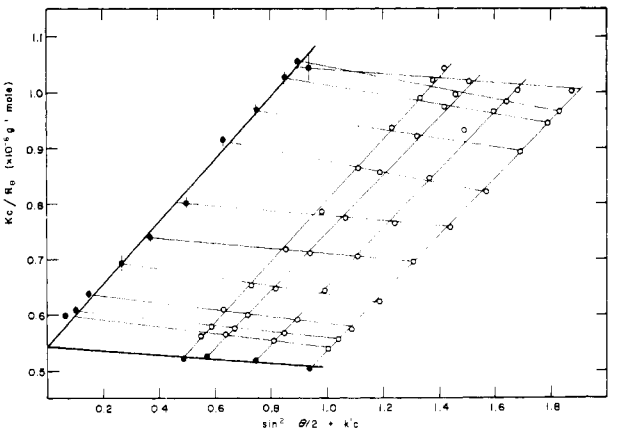

FIGURE 3: A Zimm plot for fd NaDNA in SSC at 436 nm. Solid symbols represent extrapolated points. The angular range was 30° ≤ θ ≤ 150°, and the concentration range was 4.0–18.2 μg cm^{-3}. All lines in this plot are the best fit linear least-squares lines for each set of points.

(1932), we calculated the standard deviation for the molecular weight in two ways: (1) the "external standard deviation," calculated from deviations from the weighted mean molecular weight, and (2) the "internal standard deviation," calculated from the standard deviation of the individual molecular weight values. The external and internal standard deviations agree within the probable relative error in the errors; therefore, the conclusion is that the three molecular weight values show no systematic differences (Birge, 1932). Thus the most reliable average molecular weight from these three values is their weighted average, (1.90 ± 0.07) × 10^6 (95% C.L.), which corresponds to 5740 ± 210 nucleotides.

In their review article of 1969, Marvin and Hohn concluded that fd DNA contained 6600 ± 15% nucleotides on the basis of the data available at that time and pointed out the need for a careful determination. Since then, a few other values have appeared: Bujard (1970) obtained 2.1 × 10^6, Frank and Day (1970) obtained 2.2 × 10^6, and Oka *et al.* (1971) obtained 1.75 × 10^6 from electron microscopic contour lengths; Wiseman *et al.* (1972) obtained 2.5 × 10^6 and 1.8 × 10^6 from sedimentation data and the empirical relations of Eigner and Doty (1965) and Studier (1965), respectively; Halsall and Schumaker (1972) obtained 1.7 × 10^6 for M13 DNA by sedimentation-diffusion with a large residual uncertainty from an assumed partial specific volume; Ikehara *et al.* (1973) reported 1.85 × 10^6 from the Studier (1965) relationship; and Chun *et al.* (1974) reported 1.83 × 10^6 from equilibrium sedimentation analysis, although the poor agreement in their n, w, and z average weights indicate serious problems from a sample heterodispersity. Comparison of our present results with these previous findings shows that none of the earlier individual values should have been considered reliable.

(*c*) *Molecular Weight of φX174 DNA.* A molecular weight for φX174 DNA of (1.58 ± 0.07) × 10^6 (95% C.L.) can be obtained from that of fd DNA and the ratio of lengths of their double stranded forms, 1.20 ± 0.02 (weighted average ratio from Ray *et al.*, 1966; and Frank and Day, 1970). Recent values for the viral strand of φX174 NaDNA are: 1.55 × 10^6 (Daniels, 1969); (1.59 ± 0.08) × 10^6 (average of two values for double-stranded forms, 3.17 × 10^6 and 3.18 × 10^6, and one value for the viral strand, 1.61 × 10^6, with uncertainties of 5% for each (Strider, 1971) and (1.61 ± 0.10) × 10^6 (Campbell and

Jolly, 1973). Daniels' study (1969), which was by equilibrium sedimentation in a density gradient, has been criticized by Schmid and Hearst (1971) and no uncertainty was given. Strider (1971), who also used sedimentation in a density gradient, used the revised value for the term $(1 + \Gamma')/\beta$, but did not include the 3% uncertainty in this term (Schmid and Hearst, 1971) in his uncertainty estimate (W. Strider, personal communication). Although Jolly and Campbell, who used light scattering, delineated their sources of error, their estimate of 1% uncertainty in $\partial n/\partial c$ is probably too low, especially since M is proportional to $(\partial n/\partial c)^{-2}$. The weighted average of our value for ϕX174 NaDNA with those from Strider (1971) and Campbell and Jolly (1973) is $(1.59 \pm 0.05) \times 10^6$ (95% C.L.). This corresponds to 4800 ± 160 nucleotides.

The original values of Sinsheimer (1959a,b) for ϕX174 DNA, cited as $1.6\text{–}1.8 \times 10^6$ for isolated DNA and 1.6×10^6 for the DNA from the bacteriophage weight and chemical composition, are very close to this value. The agreement, though is somewhat fortuitous in the case of the measurements on the DNA since the value of the dn/dc and the extinction coefficient used are both in error by at least 10%. There is also some ambiguity as to whether the weights were cited for the anion or for the Na salt of ϕX174 DNA.

The molecular weight of ϕX174 NaDNA has also been obtained from relative contour length measurements of ϕX174 RFII DNA *vs.* T7 DNA (Davis and Hyman, 1971) and *vs.* λ DNA (Sharp *et al.*, 1972). Davis and Hyman reported that a value of 1.7×10^6 for ϕX174 gave a molecular weight for T7 NaDNA of $(25.0 \pm 0.2) \times 10^6$, in agreement with other absolute values for T7 NaDNA (Schmid and Hearst, 1969, 1971; Leighton and Rubenstein, 1969; Bancroft and Freifelder, 1970; Dubin *et al.*, 1970), while Sharp *et al.* (1972) reported 1.74×10^6 for ϕX174 NaDNA using the recommended molecular weight for λ NaDNA of $(30.8 \pm 1.0) \times 10^6$ (Davidson and Szybalski, 1971). The difference between these values and that which we report from relative contour length measurements of ϕX174 and fd NaDNA (which are similar in size and base composition) indicates the difficulty in obtaining accurate molecular weights from relative contour lengths of DNA molecules differing considerably in size and base composition.

(*d*) *Second Virial Coefficient and Radius of Gyration.* The observed value of the second virial coefficient for fd DNA in SSC at 20°, $(0.86 \pm 0.44) \times 10^{-4}$ g^{-2} cm^3 mol, is larger than that expected for a polymer of 1.9×10^6 daltons based on its mole fraction in a solution which is thermodynamically ideal, 2.5×10^{-12} g^{-2} cm^3 mol. Such a deviation from ideality is not unexpected for polymers because of their excluded volumes. From the relation between the excluded volume of a sphere and the second virial coefficient, $B = 16\pi N R_c^3/3M^2$, we can calculate the radius, R_c, of an equivalent sphere to account for the observed virial which is 30.3 ± 5.8 nm. This radius agrees within experimental error with the radius of the equivalent hydrodynamic sphere, 31.6 ± 0.6 nm, obtained from diffusion measurements (Newman *et al.*, 1974). The ratio $R_c/R_G = 0.73$ agrees with expected values for polymers from statistical thermodynamics (see Tanford, 1961). A comparison of thermodynamic and hydrodynamic virial coefficients is given by Newman *et al.* (1974).

(*e*) *Comments on the Structure of fd Bacteriophage.* Exact knowledge of the number of nucleotides in fd DNA is important in the analysis of structure of filamentous bacteriophages (see Marvin and Hohn, 1969) and the complexes between fd DNA and gene 5 protein, a nucleoprotein complex involved in the replication of fd DNA and in viral morphogenesis (Alberts *et al.*, 1972; Oey and Knippers, 1972; Day, 1973; Pratt *et al.*, 1974). Recent progress in X-ray diffraction studies of filamentous bacteriophages has led to structural models which fit the diffraction data well (Marvin *et al.*, 1974a,b). However, the ultimate success of the model building approach to structural analysis depends on the availability of accurate chemical and physical data other than that from X-ray diffraction. Since the filamentous bacteriophages contain almost 90% protein, their diffraction patterns are dominated by protein reflections. One way of obtaining information on the arrangement of the DNA in fd bacteriophage is to combine the data for the number of nucleotides with contour lengths known from electron microscopy. The length of the fd virion obtained by Frank and Day (1970) is 8800 ± 300 Å (95% C.L.). It is reasonable to assume that the DNA passes from one end of the filament and back again. The projection of the average separation between the 2870 ± 105 bases of one strand along the filament axis is therefore (3.07 ± 0.16) Å (95% C.L.) for the virus under the conditions prevalent during electron microscopy. X-Ray diffraction studies of oriented filaments show that the basic crystallographic asymmetric unit repeats every (15.6 ± 0.1) Å at 0% relative humidity and every (16.3 ± 0.2) Å above 90% relative humidity (Dunker *et al.*, 1974). Although we do not know the corresponding repeat distance for virus which is presumably dehydrated by vacuum for electron microscopy, it seems reasonable that it is less than or equal to 15.6 Å. Since the asymmetric unit contains the DNA strands going in both directions, we conclude that the number of bases in each crystallographic unit is less than or equal to 10.2 ± 0.5. If the length of fd is as long as that claimed for M13 by Pratt *et al.* (1974) (9700 Å, no error given) this number would drop to 9.3. The best model proposed by Marvin *et al.* (1974b) for the fd structure calls for five protein subunits in each of the crystallographic units. From this line of reasoning, it would seem that there are exactly two nucleotides for each coat protein subunit in the virion. However, existing data for the chemical composition (see Marvin and Hohn, 1969) indicates 2.3 ± 0.2 nucleotides per subunit and the structural models which best account for the X-ray diffraction patterns call for a perturbed helical arrangement of protein subunits where the perturbation may come from a nonintegral number of nucleotides per protein subunit (Marvin *et al.*, 1974b). Accurate molecular weight determinations of the virion, as well as redeterminations of its chemical composition, should resolve this question of a nonintegral *vs.* an integral number of nucleotides per subunit.

Similar considerations of filament lengths allow one to conclude, from the electron micrographic data of Delius (Alberts *et al.*, 1972) and Griffith (Pratt *et al.*, 1974), that the projection of the base–base separations in the gene 5 protein–DNA complexes is about 3.8 Å. The actual separation between the individual neighboring bases, not its projection, along a strand may be much greater, since the complex has a diameter of about 150 Å, roughly three times that of the bacteriophage.

Acknowledgments

We thank Margarete Klein for her assistance in preparing materials, Jay Newman and Harry L. Swinney for helpful discussion and comment, William F. Harrington for the

use of his densimeter for some of the measurements, and David A. Yphantis for providing us with the computer program developed in his laboratory.

References

Alberts, B. M., Frey, L., and Delius, H. (1972), *J. Mol. Biol. 68,* 139–152.
Bancroft, F. C., and Freifelder, D. (1970), *J. Mol. Biol. 54,* 537–546.
Berkowitz, S. A. (1974), Ph.D. Thesis, New York University.
Birge, R. T. (1932), *Phys. Rev. 40,* 207–227.
Bujard, H. (1970), *J. Mol. Biol. 49,* 125–137.
Campbell, A. M., and Jolly, D. J. (1973), *Biochem. J. 133,* 209–226.
Casassa, E. F., and Eisenberg, H. (1964), *Advan. Protein Chem. 19,* 287–395.
Chen, P. S. Jr., Toribara, T. Y., and Warner, H. (1956), *Anal. Chem. 28,* 1756–1758.
Chun, R. W., Herschleb, W. P., Downing, D. J., and Krista, M. L. (1974), *Biophys. Chem. 1,* 141–151.
Cohen, G., and Eisenberg, H. (1968), *Biopolymers 6,* 1077–1100.
Crothers, D. M., and Zimm, B. H. (1965), *J. Mol. Biol. 12,* 525–536.
Daniels, E. (1969), *Biopolymers 7,* 359–377.
Davidson, N., and Szybalski, W. (1971), *in* The Bacteriophage Lambda, Hershey, A. D., Ed., New York, N.Y., Cold Spring Harbor Laboratory, Chapter 3.
Davis, R. W., and Hyman, R. W. (1971), *J. Mol. Biol. 62,* 287–301.
Day, L. A. (1973), *Biochemistry 12,* 5329–5339.
Dubin, S. B., Benedek, G. B., Bancroft, F. C., and Freifelder, D. (1970), *J. Mol. Biol. 54,* 547–556.
Dunker, A. K., Klausner, K. D., Marvin, D. H., and Wiseman, R. L. (1974), *J. Mol. Biol. 81,* 115–117.
Eigner, J., and Doty, P. (1965), *J. Mol. Biol. 12,* 549–580.
Eisenberg, H. (1962), *J. Chem. Phys. 30,* 1837–1843.
Eisenberg, H., and Felsenfeld, G. (1967), *J. Mol. Biol. 30,* 17–37.
Frank, H., and Day, L. A. (1970), *Virology 42,* 144–154.
Freifelder, D. (1970), *J. Mol. Biol. 54,* 567–577.
Halsall, H. B., and Schumaker, V. N. (1972), *Biochemistry 11,* 4692–4695.
Hoffmann-Berling, H., Marvin, D. A., and Dürwald, H. (1963), *Z. Naturforsch. B 18,* 875–883.
Ikehara, K., Obata, Y., Utiyama, H., and Kurata, M. (1973), *Bull. Inst. Chem. Res., Kyoto Univ. 5,* 140–152.
Inman, R. B. (1967), *J. Mol. Biol. 25,* 209–216.
Inners, L. D., and Felsenfeld, G. (1970), *J. Mol. Biol. 50,* 373–389.
Iwaya, M., Eisenberg, S., Bartok, K., and Denhardt, D. T. (1973), *J. Virol. 12,* 808–818.
Johnson, P. H., and Sinsheimer, R. L. (1974), *J. Mol. Biol. 83,* 47–61.
Kleinschmidt, A. K., and Zahn, R. K. (1959), *Z. Naturforsch. B 14,* 770–779.
Krasna, A. I. (1970), *Biopolymers 9,* 1029–1038.
Krasna, A. I., Dawson, J. R., and Harpst, J. A. (1970), *Biopolymers 9,* 1017–1028.
Kratky, O., Leopold, H., and Stabinger, H. (1973), *Methods Enzymol. 27D,* 98.
Kratohvil, J. P., Dezelic, Gj., Kerker, M., and Matijevic, E. (1962), *J. Polym. Sci. 57,* 59–77.
Lang, D. (1970), *J. Mol. Biol. 54,* 557–565.
Leighton, S. B., and Rubenstein, I. (1969), *J. Mol. Biol. 46,* 313–328.
Marvin, D. A., and Hohn, B. (1969), *Bacteriol. Rev. 33,* 172–209.
Marvin, D. A., Pigram, W. J., Wiseman, R. L., Wachtel, E. J., and Marvin, F. J. (1974b), *J. Mol. Biol.* (in press).
Marvin, D. A., Wiseman, R. L., and Wachtel, E. J. (1974a), *J. Mol. Biol. 82,* 121–138.
Meselson, M., and Stahl, F. W. (1958), *Proc. Nat. Acad. Sci. U. S. 44,* 671–682.
Nazarian, G. M. (1968), *Anal. Chem. 40,* 1766–1769.
Newman, J., Swinney, H. L., Berkowitz, S. A., and Day, L. A. (1974), *Biochemistry 13,* 4832.
Oey, J. L., and Knippers, R. (1972), *J. Mol. Biol. 68,* 125–138.
Oka, T., Nakamura, T., Watanabe, S., and Takanami, M. (1971), *J. Electronmicrosc. 20,* 67–71.
Pratt, D., Laws, P., and Griffith, J. (1974), *J. Mol. Biol. 82,* 425–439.
Ray, D. S., Preuss, A., and Hofschneider, P. H. (1966), *J. Mol. Biol. 21,* 485–491.
Roark, D. E., and Yphantis, D. A. (1969), *Ann. N. Y. Acad. Sci. 164,* 245–278.
Schachman, H. K., and Edelstein, S. J. (1966), *Biochemistry 5,* 2681–2705.
Schaller, H., Voss, H., and Gucker, S. (1969), *J. Mol. Biol. 44,* 445–458.
Schekman, R. W., and Ray, D. S. (1971), *Nature (London), New Biol. 231,* 170–173.
Schmid, C. W., and Hearst, J. E. (1969), *J. Mol. Biol. 44,* 143–160.
Schmid, C. W., and Hearst, J. E. (1971), *Biopolymers 10,* 1901–1924.
Sharp, P. A., Hsu, M.-T., Ohtsubo, E., and Davidson, N. (1972), *J. Mol. Biol. 71,* 471–497.
Sinsheimer, R. L. (1959a), *J. Mol. Biol. 1,* 37–42.
Sinsheimer, R. L. (1959b), *J. Mol. Biol. 1,* 43–53.
Strider, W. (1971), Ph.D. Thesis, New York University.
Studier, F. W. (1965), *J. Mol. Biol. 11,* 373–390.
Tanford, C. (1961), Physical Chemistry of Macromolecules, New York, N.Y., Wiley, Chapter 4.
Teller, D. C. (1973), *Methods Enzymol. 27D,* 346.
Tomimatsu, Y., Vitello, L., and Fong, K. (1968), *J. Colloid Interface Sci. 27,* 573–580.
Vinograd, J., Bruner, R., Kent, R., and Weigle, J. (1963), *Proc. Nat. Acad. Sci. U. S. 49,* 902–910.
Wallace, T. P., Volosin, M. T., Delumyea, R. G., Gingello, A. D. (1972), *J. Polym. Sci., Part A-2 10,* 193–200.
Wippler, C., and Scheibling, G. (1954), *J. Chim. Phys. Physicochim. Biol. 51,* 201–205.
Wiseman, R. L., Dunker, A. K., and Marvin, D. A. (1972), *Virology 48,* 230–244.
Yamamoto, K. R., Alberts, B. M., Benzinger, R., Lawhorne, L., and Treiber, G. (1970), *Virology 40,* 734–744.
Yphantis, D. A. (1964), *Biochemistry 3,* 297–317.
Yphantis, D. A., and Roark, D. E. (1972), *Biochemistry 11,* 2925–2934.
Zimm, B. H. (1948), *J. Chem. Phys. 16,* 1099–1116.