Molecular Weights of Homogeneous Coliphage DNA's from Density-gradient Sedimentation Equilibrium

CARL W. SCHMID AND JOHN E. HEARST

*Department of Chemistry, University of California
Berkeley, Calif. 94720, U.S.A.*

(Received 17 March 1969)

The problem of obtaining meaningful molecular weights for high molecular weight DNA samples by sedimentation equilibrium in a density gradient has been solved by correcting the theory for thermodynamic non-ideality. Several different methods of virial corrections are suggested and examined to determine the best method of obtaining molecular weights at infinite dilution.

A pronounced virial effect is experimentally demonstrated for high molecular weight DNA's. The concentration distributions at sedimentation equilibrium are studied as a function of DNA concentration for T7, T5 and T4 coliphage DNA's and the results extrapolated to infinite dilution to obtain molecular weights for these DNA's.

The values obtained by this technique are 23.2×10^6 for T7, 65.7×10^6 for T5 and 104×10^6 daltons for T4 DNA. These results are in agreement with recent hydrodynamic predictions but are in disagreement with many literature values. A brief review of the literature is presented to elucidate these discrepancies.

1. Introduction

The molecular weights of *Escherichia coliphage* DNA's have been studied by several methods. Absolute measurements are possible by autoradiographic ^{32}P "star gazing" (Rubenstein, Thomas & Hershey, 1961), terminal labeling (Richardson, 1966); and for molecular weights less than 30×10^6 daltons, light-scattering is possible (Harpst, Krasna & Zimm, 1968*a,b*). Relative values may be obtained through length determinations by either electron microscopy (Thomas, 1966) or ^3H autoradiography (Cairns, 1961). The mass of the DNA is then determined by either assuming a linear mass density or establishing a calibration of length to known molecular weights. Sedimentation velocity, intrinsic viscosity (Crothers & Zimm, 1965) and zone centrifugation (Abelson & Thomas, 1966) may also be used, with known molecular weight samples as standards.

An alternate approach is to obtain an absolute molecular weight of the phage by light-scattering, the sedimentation-diffusion method, or plaque-counting (Rubenstein *et al.*, 1961; Davison & Freifelder, 1962; Cummings & Kozloff, 1960). The molecular weight of the DNA is then calculated from the phage weight with the phosphate analysis of the phage.

Recently molecular weights have been calculated from sedimentation velocity and intrinsic viscosity (Hearst, Schmid & Rinehart, 1968*a*) with the wormlike coil model for which the hydrodynamic properties have been derived in a series of papers (Hearst & Stockmayer, 1962; Harris & Hearst, 1966; Gray, Bloomfield & Hearst, 1967;

Hearst, Harris & Beals, 1966; Hearst, Beals & Harris, 1968*b*). The calculated values are in disagreement with some literature values, and therefore both independent measurements of the molecular weights to test the hydrodynamic theory and a review of the literature to try to determine the most accurate values for the molecular weights are now necessary.

As is well known, molecular weight measurements on high molecular weight DNA's are not readily obtainable by ordinary physico-chemical techniques. Sedimentation equilibrium in a density-gradient has been proposed and investigated for such studies (Meselson, Stahl & Vinograd, 1957; Hearst & Vinograd, 1961*bc*; Hearst, Ifft & Vinograd, 1961). However, molecular weights for DNA obtained by this method have been much lower than the expected values. This anomaly has been attributed to several causes; light-scattering by the DNA molecules distorting the observed distribution (Cummings, 1963), density heterogeneity (Sueoka, 1959); and unknown causes (Thomas & Pinkerton, 1962).

In the present paper, the effects of thermodynamic non-ideality are considered. Thomas & Pinkerton (1962) considered the possibility of such an effect but overlooked it, as will be explained in the Discussion. The corrections for thermodynamic non-ideality are presented in the form of a virial expansion. Several different forms of virial expansions are investigated with data for T7 DNA to choose the best method of correction. The molecular weights obtained by this method for T7, T5 and T4 DNA are compared with values reported in the literature.

2. Theory

At sedimentation equilibrium in a buoyant density gradient; Hearst & Vinograd (1961*b*) have shown that:

$$-M_{s,o} \left(\frac{\partial \rho}{\partial r} \right)_{\text{eff}} \bar{v}_{s,o} r_0 \omega^2 \delta d\delta = \left(\frac{\partial \mu_3}{\partial m_3} \right)_{\mu_1} dm_3 \quad (1)$$

$$M_{s,o} = M_3(1 + \Gamma') = M_3 \left\{ 1 + \left(\frac{\partial w_1}{\partial w_3} \right)_{\mu_1} \right\}, \quad (2)$$

where the symbols have the following meanings:

$\mu,$	chemical potential;
$\left(\frac{\partial \rho}{\partial r} \right)_{\text{eff}},$	effective density gradient;
$r_0,$	position of band center;
$\bar{v}_{s,o},$	partial specific volume of solvated DNA;
$\omega,$	angular velocity;
$\delta,$	distance from band center;
$w,$	grams per thousand grams solvent;
$M_{s,o},$	molecular weight of the solvated DNA;
1 subscript,	low molecular weight solute (water);
3 subscript,	macromolecular component (cesium DNA).

In previous studies, equation (1) has been integrated by assuming thermodynamic ideality for the solvated polymer, which leads to the prediction of a Gaussian distribution:

$$m_3 = m_{3,o} \exp \left[-M_{s,o} \left(\frac{\partial \rho}{\partial r} \right)_{\text{eff}} \bar{v}_{s,o} \frac{\omega^2 r_0 \delta^2}{2RT} \right] = m_{3,o} \exp - \frac{\delta^2}{2\sigma^2}. \quad (3)$$

The right-hand side of equation (1) may be expanded in a power series to correct for thermodynamic non-ideality in a manner analogous to that suggested by Flory (1953):

$$\left(\frac{\partial \mu_3}{\partial m_3}\right)_{\mu_1} = \frac{RT}{m_3} (1 + 2Bm_3 + 3gB^2m_3^2 + \dots). \quad (4)$$

where B is the second virial coefficient at constant chemical potential of the low molecular weight solute (Vinograd & Hearst, 1962), and g is a parameter relating the third virial coefficient to the square of the second virial coefficient. Theoretical values of g are $5/8$ for a sphere (Flory, 1949; Flory & Krigbaum, 1950) and approximately 0.25 for flexible polymers (Stockmayer & Casassa, 1952).

Equation (4) and equation (1) are combined and integrated to give equation (5). The series of equation (4) has been truncated after the third virial term:

$$\frac{-KM_{s,o}\delta^2}{2} = \ln \frac{m_3}{m_{3,o}} + 2B(m_3 - m_{3,o}) + \frac{3}{2}gB^2(m_3^2 - m_{3,o}^2). \quad (5)$$

The parameter K is defined by the equation:

$$K = \left(\frac{\partial \rho}{\partial r}\right)_{\text{eff}} \bar{v}_{s,o} r_0 \omega^2 / RT$$

Equation (5) allows rapid calculation of the molecular weight. This is done by measuring the width, σ_{app} , of the experimental distribution at the concentration $m_{3,o} \exp -\frac{1}{2}$, or 0.606 $m_{3,o}$. For this value, equation (5) may be rewritten as:

$$\frac{1}{M_{\text{app}}} = K\sigma_{\text{app}}^2 = \frac{1}{M_{s,o}} (1 + 1.576 Bm_{3,o} + gB^2 1.898 m_{3,o}^2). \quad (6)$$

Apparent molecular weights obtained by this method are too inaccurate to permit a simultaneous evaluation of $M_{s,o}$, B and g , and therefore reasonable estimates of g will be used so that the data will only require a fit to two unknowns $M_{s,o}$ and B (Flory, 1953). Before discussing this, however, there is an alternative method of combining equation (4) and equation (1) to obtain a useful moments relation assuming rectangular geometry over the band, which is derived below.

If equation (4) is substituted into equation (1) and multiplied by $m_3\delta$, one obtains equation (7):

$$-KM_{s,o}m_3\delta^2d\delta = \delta(1 + 2Bm_3 + 3gB^2m_3^2)dm_3. \quad (7)$$

Equation (7) may now be integrated, and for convenience it will be integrated from $-\infty$ to $+\infty$. These limits of integration and integration by parts which will now follow assume only that the entire band is positioned in the cell with essentially zero macromolecule concentration at the miniscus and cell bottom. Integrating the right-hand side of equation (7) by parts over these limits gives equation (8):

$$\int_{-\infty}^{+\infty} KM_{s,o}m_3\delta^2d\delta = \int_{-\infty}^{+\infty} (m_3 + Bm_3^2 + gB^2m_3^3)d\delta \quad (8)$$

Normalizing equation (8) by dividing by $\int_{-\infty}^{+\infty} m_3d\delta$ and expressing it in a more convenient notation, the moment relation, equation (9), is obtained:

$$KM_{s,o}\langle\delta^2\rangle = 1 + B\langle m_3\rangle + gB^2\langle m_3^2\rangle. \quad (9)$$

An additional simplification is to assume proportionality between $\langle m_3 \rangle^2$ and $\langle m_3^2 \rangle$; for a Gaussian distribution the proportionality is:

$$\langle m_3^2 \rangle = \frac{2}{\sqrt{3}} \langle m_3 \rangle^2.$$

Making this substitution for $\langle m_3^2 \rangle$ and rearranging equation (9), a more useful expression is obtained:

$$K\langle \delta^2 \rangle = \frac{1}{M_{s,o}} \left(1 + B\langle m_3 \rangle + g \frac{2}{\sqrt{3}} B^2 \langle m_3 \rangle^2 \right) \equiv \frac{1}{M_{app}}. \quad (10)$$

The averages appearing in equation (10), $\langle \delta^2 \rangle$ and $\langle m_3 \rangle$, are easily obtained by numerical analysis of the equilibrium distribution. Since the numerical value of K is known (Hearst & Vinograd, 1961*b,c*; Hearst, Ifft & Vinograd, 1961), an apparent molecular weight M_{app} may be measured as a function of concentration $\langle m_3 \rangle$.

Equation (10) as written has the same problem as mentioned for equation (6). The experimental determinations of M_{app} and $\langle m_3 \rangle$ will be too inaccurate to permit a three-parameter fit of the data, and instead it will be better to estimate and use a two-parameter fit.

Both equations (6) and (10) may be written as:

$$\frac{1}{M_{app}} = \frac{1}{M_{s,o}} (1 + B'C + g'(B')^2 C^2) \quad (11)$$

where C is the concentration appropriate for the expression chosen. The parameters B' and g' are the virial constants with the appropriate weighting factors, depending on whether equation (6) or (10) is being used.

The following approximations may now be made to equation (11) by selecting a range of values for g' .

For $g' = 0$, the familiar linear correction for virial effects would apply:

$$\frac{1}{M_{app}} = \frac{1}{M_{s,o}} (1 + B'C). \quad (12)$$

For $g' = 0.25$, the square root of equation (11) may be taken and the Berglund expression is obtained (Flory, 1953):

$$\sqrt{1/M_{app}} = \sqrt{1/M_{s,o}} \left(1 + \frac{B'}{2} C \right). \quad (13)$$

If $g' = 0.5$, equation (11) may be taken as a truncated Maclaurin series for the exponential:

$$\frac{1}{M_{app}} = \frac{1}{M_{s,o}} \exp B'C;$$

or more conveniently:

$$\ln M_{app} = \ln M_{s,o} - B'C \quad (14)$$

This holds for $B'C < 1$. For $B'C \approx 1$ or $M_{s,o}/M_{app} \approx 2.7$, the exponential approximation agrees with equation (11) to within 10%.

For $g' \approx 1$, equation (11) may be taken as a truncated series expansion and rearranged to:

$$M_{app} = \frac{M_{s,o}}{1 + B'C + (B')^2 C^2 + \dots} \approx M_{s,o} (1 - B'C). \quad (15)$$

Equation (15) is also only valid for $B'C < 1$, and will reproduce the value predicted by equation (11) to within 10% for $M_{s,0}/M_{app} \approx 1.7$.

The apparent molecular weight is known at a given concentration by either of the two different methods described in conjunction with equations (6) and (10). A series of linear plots of a function of apparent molecular weight against concentration is suggested by equations (12), (13), (14) and (15) to obtain the desired quantity $M_{s,0}$. These plots are examined in a following section of this paper, to select the most reliable expression.

A number average molecular weight is obtained from the moment relation for $\langle \delta^2 \rangle$, equation (10) (Meselson *et al.*, 1957). Direct measurement of the apparent half-width and use of equation (6) gives a weight average molecular weight in the limit of zero polymer concentration if this concentration distribution is Gaussian.

This may be shown by averaging the expression for the weight average molecular weight

$$M_w(\delta) = \frac{-2}{K} \frac{d \ln C(\delta)}{d \delta^2}$$

(Meselson *et al.*, 1957) over the entire Gaussian concentration distribution:

$$M_w = \frac{\int C(\delta) M_w(\delta) d\delta}{\int C(\delta) d\delta} = \frac{1}{K\sigma^2}$$

where the Gaussian concentration distribution, $C(\delta) = C(0) \exp -\delta^2/2\sigma^2$ defines σ .

3. Experimental Procedure

(a) Materials

Harshaw optical grade CsCl in 0.01 M-Tris (pH 7.2) at densities of about 1.70 g/ml., depending on the DNA being studied, was used as the buoyant medium. The DNA's were characterized by sedimentation velocity in 0.1 M-NaCl in 0.044 M-Na₂HPO₄, 0.056 M-NaH₂PO₄ and 0.001 M-Na₂ EDTA (Gray & Hearst, 1968). The samples of T4^{ort} and T5⁺ were obtained by phenol extractions of the phage followed by exhaustive dialysis of the phenol in the above buffer. For T7 DNA, $S_{20,w} = 32.6$ s, T4^{ort} DNA $S_{20,w} = 61.2$ s and for T5⁺ DNA $S_{20,w} = 50.9$ s (Gray & Hearst, 1968).

(b) Methods

Centrifugation was performed on a Beckman model E ultracentrifuge, equipped with a scanner and modified speed control (Smiriga & Hearst, 1969). All work was done with double-sector optics using double-sector titanium centerpieces (Hearst & Gray, 1968).

The refrigeration unit was adjusted to a high temperature, for which the rotor achieved a steady-state temperature of approximately 20°C without the use of a temperature control unit. This was done to avoid thermal gradients which may arise from the temperature control and which can be predicted to lead to band distortion.

The system was judged to be at equilibrium by following the band height and band width with time until they remained constant. Equilibrium is usually achieved after 72 hr.

Traces of the concentration distribution were taken as a function of slit width. A slit width was selected at which band height and band width were unaffected by further reduction in slit width. This occurs at a slit width of about 0.06 mm for the conditions reported here.

The concentration distribution was studied as a function of wavelength to determine if the band width is affected by light-scattering as Cummings (1963) has suggested.

(c) Numerical analysis

For the moment analysis of the concentration distribution, equally spaced data intervals were selected and integration was performed with the trapezoidal rule. The maximum

and minimum intervals selected were 0.01 mm and 0.002 mm in cell dimensions for the studies on T7 and T4 DNA, respectively. A trial band center was selected by symmetry and a trial apparent molecular weight was then calculated by equation (10). The band center was then re-positioned in an iterative manner until a maximum molecular weight was obtained. In all cases the symmetrical band center was in good agreement with the computed band center. This is evidence for the symmetry of the concentration distribution and precludes the possibility of biasing the apparent molecular weight by improperly selecting the band center.

To determine if the intervals of data selection were sufficiently small to represent properly the concentration distribution, several molecular weights were recalculated with data intervals eight times larger. The calculated molecular weight was unaffected by increasing the width of the intervals.

The effective gradients were calculated assuming the experiments were done at 25°C from the data in Vinograd & Hearst (1962). Schumaker & Wagnild (1965) have estimated that the physical gradient would be 2.5% higher at 20°C, but we have made no attempt to correct for this small effect.

4. Results

(a) *Effects of light-scattering*

The apparent molecular weight was found to be completely independent of the wavelength in the region of 248 to 280 m μ . A trend in band width similar to that previously reported by Cummings (1963) was observed. This trend was due entirely to the change in magnification caused by refocusing the camera lens to compensate for the wavelength change.

(b) *Virial corrections*

In Figure 1, distributions for two concentrations normalized to the same peak height are reproduced. The higher concentration curve has a wider half width, and therefore a lower apparent molecular weight. This is indicative of thermodynamic non-ideality for the solvated polymer.

The apparent molecular weights for T7 DNA have been evaluated by direct measurement of the half width, σ_{app} , as explained in conjunction with equation (6). Infinite dilution molecular weights from the least squares fit of the data have been

FIG. 1. T7 DNA equilibrium concentration distributions; ----, $O.D._{260} = 1.420$, $\langle \delta^2 \rangle^{1/2} = 3.46 \times 10^{-2}$ cm; —, $O.D._{260} = 0.355$, $\langle \delta^2 \rangle^{1/2} = 2.91 \times 10^{-2}$ cm. Optical densities are 0.606 of the maximum peak height O.D.

tabulated in Table 1 using equations (12), (13), (14) and (15) to correct for non-ideality. The data are plotted as $1/M$ (equation (12)) in Figure 2 and $\log M$ (equation (14)) in Figure 3.

FIG. 2. T7 DNA, from the width of the experimental peak at $0.606 \text{ O.D.}_{\text{max}}$, $1/M_{\text{app}}$ versus concentration. Least square lines—curves are explained in text: $\circ$, 25,000 rev./min; $\bullet$, 35,000 rev./min. Concentrations are 0.606 of the maximum peak height $\text{O.D.}_{280}^{\text{1cm}}$.

FIG. 3. T7 DNA, $\log_{10} M_{\text{app}}$ versus concentration with least square lines: $\circ$, 25,000 rev./min; $\bullet$, 35,000 rev./min. Concentrations are 0.606 of maximum peak height $\text{O.D.}_{280}^{\text{1cm}}$.

The apparent molecular weights for T7 DNA from the moments analysis method, equation (10), have similarly been extrapolated to infinite dilution with equations (12), (13), (14) and (15). These molecular weights are also compiled in Table 1. The reciprocal and logarithm of the apparent molecular weights are plotted in Figures 4 and 5.

Definite curvature in plots of M_{app} versus concentration as well as in plots of $1/M_{\text{app}}$ versus concentration is apparent. This curvature reflects an overcompensation

FIG. 4. T7 DNA moment analysis. $1/M_{app}$ versus concentration with least square lines—curves explained in text: $\circ$, 25,000 rev./min; $\bullet$, 35,000 rev./min. Concentrations are the average $O.D._{260}^{1.0m}$ in the band. $\int (o.d.)^2 d\delta / \int (o.d.) d\delta$.

FIG. 5. T7 DNA moment analysis, $\log_{10} M_{app}$ versus concentration with least square lines: $\circ$, 25,000 rev./min; $\bullet$, 35,000 rev./min. Concentrations are the average $O.D._{260}^{1.0m}$ in the band.

for third virial effects in equation (15) and undercompensation for this term in equation (12). Curvature was not noticeable in plots of $\sqrt{1/M_{app}}$ or $\log M_{app}$ versus concentration.

The data were also analyzed by equation (11), using a least squares quadratic. This was done to obtain a crude measure of any curvature and to obtain an alternative measure of the infinite dilution molecular weight. These results are also tabulated in Table 1.

TABLE I
T7 DNA infinite dilution molecular weights by trial virial approximations
 $M_{s,0} \times 10^{-6}$

Plot	M	$\sqrt{1/M}$	$1/M$	$\log M$	Quadratic fit of equation (11)	
					M equation (11)	g' equation (11)
Moment analysis						
25,000 rev./min	37.9	40.5	40.4	38.8	40.2	0.33
35,000 rev./min	37.5	41.0	40.6	40.3	38.5	4.0
Apparent half width						
25,000 rev./min	36.2	39.6	41.9	38.0	37.9	0.31
35,000 rev./min	40.2	43.4	54.0	39.0	34.0	10.0
g' assumed	1	1/4	0	0.5		

In all cases the value of g' obtained by the quadratic fit is positive. Although this measure of curvature is crude, it does permit recognition of consistently positive curvature, or the presence of third-order virial effects, equation (11). This may also be seen in the molecular weights obtained by linear $1/M$ extrapolations, which are consistently higher than the quadratic fit of $1/M$.

The logarithmic plot has been selected for the remainder of this work as it gives the most consistent values and is in the best agreement with both the individual values and average values of the quadratic fit. Less precise values are obtained from plots of $\sqrt{1/M_{app}}$, $1/M_{app}$ and M_{app} against concentration. The visible curvature in the plots of M_{app} and $1/M_{app}$ indicates improper virial corrections with these approximate equations.

Taking the average molecular weight, $M_{s,0}$ found for T7 DNA, namely 39.0×10^6 daltons with the logarithmic correction and the least square slopes of these plots, the exponential form of equation (14) is plotted in Figures 2 and 4. The exponential approximation provides a good fit to the virial behavior of the data. It is interesting to note that the quadratic expansion of these exponential functions gives a poor fit to the data.

(c) Molecular weights of T4, T5, T7 DNA

Apparent molecular weights for T5⁺ DNA and T4^{ort} DNA have been obtained by the moments analysis. The infinite dilution values for these molecular weights were obtained by the logarithmic correction (Figs 6 and 7). Note the magnitude of the virial correction. In Figure 7, the apparent molecular weight for T4 DNA at $\langle \text{o.d.}_{265} \rangle = 0.6$ is 0.4 of the extrapolated value.

FIG. 6. T5 DNA moment analysis, $\log_{10} M_{app}$ versus concentration with least square lines: O, 25,000 rev./min; ●, 35,000 rev./min. Concentrations are the average $\text{o.d.}_{265}^{1.0\text{cm}}$ in the band.

The molecular weights for the hydrated cesium DNA's are $(110 \pm 6) \times 10^6$ and $(173 \pm 10) \times 10^6$ daltons for T5⁺ and T4^{ort} DNA. Since the direct measurement of the apparent half width is subject to greater uncertainties than the statistically more significant moment relationship, we decided to consider only the moment relationship values in reporting molecular weights. The direct measure of half widths gave molecular weights of $(128 \pm 8) \times 10^6$ and $(163 \pm 6) \times 10^6$ daltons for hydrated cesium T5⁺ and T4^{ort} DNA's. With this additional criterion, T7 is taken to have a molecular weight $M_{s,0}$ of $(39.5 \pm 0.6) \times 10^6$ daltons.

These molecular weight values, $M_{s,0}$, are corrected to dry cesium DNA by equation (2), taking $I'' = 0.28$ (Hearst & Vinograd, 1961c). To convert to the sodium salt,

FIG. 7. T4 DNA moment analysis $\log_{10} M_{app}$ versus concentration with least square lines: $\circ$, 25,000 rev./min; $\triangle$, 30,000 rev./min (high concentration point omitted); $\bullet$, 35,000 rev./min. Concentrations are the average $O.D._{265}^{1cm}$ in the band.

M_3 is multiplied by the ratio of the mean residue weights of the sodium and cesium nucleotides. This ratio is 0.75 for T5⁺ and T7 and 0.775 for T4, assuming for T4 that all the 5-HMC are glucosylated with a base composition of 34% G-5-HMC (Erikson & Szybalski, 1964). These number average molecular weights for dry sodium DNA are compiled in Table 2 with values obtained by other techniques.

5. Discussion

(a) Virial effects

This work substantiates that pronounced virial effects are present, including contributions from at least the third virial coefficient, in density-gradient sedimentation equilibrium. This is in contradiction to earlier results by Thomas & Pinkerton (1962). In their work, these authors used "baseline corrections" to obtain linear $\log m$ versus δ^2 plots. Linearity of such a plot forces the data to fit equation (3), which only applies in the absence of thermodynamic non-ideality. These authors probably eliminated any obvious trend of apparent molecular weight with concentration by their baseline manipulations. This problem was avoided in the present work by use of double-sector optics and the photoelectric scanner, which provide a clear baseline and far better data. An example of the plot of $\log m$ versus δ^2 for our data at high and low concentrations is presented in Figure 8. Note the large curvature for the high concentration plot, which is a result of thermodynamic non-ideality. At lower concentrations the slope of the high concentration plot approaches the slope of the low concentration plot. The coincidence of the limiting straight line for infinitely dilute DNA with the low concentration data in Figure 8 reflects the experimental error in selection of the maximum concentration at band center. The experimental distribution for this concentration (Fig. 1) is about 10% broader than the infinite dilution distribution. Integrating numerically the experimental distribution to find $\langle \delta^2 \rangle$ and $\langle m \rangle$ is not sensitive to the value of the maximum concentration at band center.

The alternatives to using high-order virial corrections, working at lower concentrations and selection of a theta solvent, do not appear promising at present. The lower limit on concentrations which may be reasonably studied with existing equipment is

FIG. 8. Analysis of data for two concentrations of T7 DNA assuming thermodynamic ideality. $\log O.D._{260}^{1.20}$ versus δ^2 : $\circ$, $O.D._{260, max} = 1.420$; $\bullet$, $O.D._{260, max} = 0.355$. Limiting straight line for infinitely dilute DNA with a molecular weight of 39.0×10^6 daltons and $\sigma = 2.62 \times 10^{-2}$. The data were obtained from the width of the experimental distribution at equally spaced concentration intervals and correspond to the distributions traced in Fig. 1.

about 0.1 O.D.^{1cm} . Although cells of long path length could be used to reduce this minimum concentration, an additional optical error would then be caused by the increased bending of the light due to the longer path length (Hearst & Vinograd, 1961a). An increase in path length also magnifies any thermal gradients which might exist, leading to band distortion. Solvent requirements are already stringent, and it is not likely that a theta solvent could be found which would be a satisfactory buoyant-density medium.

Sedimentation equilibrium in a density gradient has also been used as a measure of heterogeneity of composition (Sueoka, 1959, 1961) of DNA samples. This is achieved by comparing the theoretical half-width, calculated from a knowledge of the molecular weight to the broader experimental half-width. The difference between the half-widths is taken as a contribution from density heterogeneity, which is a measure of the heterogeneity of composition. Such studies have previously ignored any band spreading due to thermodynamic non-ideality. In light of this work, values of heterogeneity of composition from band spreading are meaningful only when calculated from band-width at infinite dilution.

(b) Summary of coliphage DNA molecular weights

In Table 2 we have compiled values for the molecular weights of coliphage DNA's. Our purpose in doing this is to determine how these values originate and to discuss

the limitations of the techniques employed in their determination. Hopefully by so doing we may then conclude what the most reasonable values for these quantities are. In this discussion it will be noted that many molecular weights arise either from radioautography or methods calibrated directly or indirectly to the results of radioautography for T2 DNA. It is our belief, as has been previously noted (Thomas, 1966), that a molecular weight of 130×10^6 daltons for T2 DNA is inconsistent with the molecular weight values obtained by other methods. This inconsistency is somewhat surprising, as T2 DNA was used in the calibration of such methods as contour length from microscopy and sedimentation in sucrose gradients with an assigned molecular weight of 130×10^6 daltons.

(c) *Density gradient sedimentation equilibrium*

Determinations of molecular weights with this method are absolute in nature and require no calibrations. A major criticism of this method has been the consideration of the effect of density heterogeneity resulting from heterogeneity of composition on the observed molecular weight. In the cases of T7 DNA (Ritchie, Thomas, MacHattie & Wensink, 1967) and T5 DNA (Thomas & Rubenstein, 1964), there is no reason to suspect any heterogeneity of composition. T-even DNA is thought to be permuted (Thomas & Rubenstein, 1964) and terminally repetitious (MacHattie, Ritchie, Thomas & Richardson, 1967). These authors estimate 1 to 3% terminal redundancy in the permuted T2 DNA. For a 2% piece, 4000 base pairs, the standard deviation of a random distribution of sequences will be approximately 65 base pairs. Heterogeneity of composition of 65 base pairs from the average in the 2% piece would decrease the observed molecular weight only 0.3%. An error associated with the terminal redundancy of T4 DNA is therefore not likely, although possible. The heterogeneity associated with small numbers of single-strand breaks can also be predicted to be negligible.

(d) *Theoretical hydrodynamics*

The hydrodynamic values are calculated for the wormlike coil model from sedimentation coefficients and intrinsic viscosity with the Flory-Mandelkern equation (Hearst *et al.*, 1968a). The range of molecular weight results from the uncertainty in the value of the parameter β in the Flory-Mandelkern equation and is not a measure of the experimental uncertainties in the intrinsic viscosity and sedimentation coefficients. It should be noted that, in addition to uncertainties in β , there are also inconsistencies in comparing values for $S_{20,w}$ and $[\eta]$ obtained from different laboratories. For these reasons T4 and T2 DNA will be considered identical in molecular weight in this study. The differences reported in their molecular weight range and in the range of molecular weights for the T5 mutants by the hydrodynamic theory cannot be considered significant.

The hydrodynamic method of molecular weight determination assumes the applicability of the wormlike coil model. Part of the purpose of the present work is to test the wormlike coil model for DNA hydrodynamics. For this reason we will not comment on the validity of these molecular weights at this point, but merely note that they constitute an absolute molecular weight determination and that they are in good agreement with the density-gradient results of this work (Table 2).

(e) *Autoradiography*

The autoradiography is performed by counting the tracks from disintegration of ^{32}P -labeled DNA embedded in a photographic emulsion in the form of free DNA or

TABLE 2
Molecular weights E. coliphage DNAs
 $M \pm 10^{-6}$

Sample	Sedimentation equilibrium	Hydrodynamics ^a	Autoradiography	Relative sedimentation	Length	% Phage	Miscellaneous
T7	23.1 ± 0.4	21.6—24.0	24.3 ^b			21.3 ± 3 ^e	26.0 ^f 22.8—27.0 ^f
T5 ^{sto}		56.0—62.3	78.5 ^b	81.2 ^{b, d}	77.6 ^e		
T5 +	64 ± 6	59.9—66.6	85.0 ^b	87 ^{b, d}			
T4	105 ± 6	89.6—99.7					
T2		94.4—105	130 ^c		109 ± 8 ^f 118 ^e	102 ^h 160 ^e	

^a Hearst *et al.*, 1968*ab*.

^b Thomas & MacHattie, 1967.

^c Rubenstein *et al.*, 1961.

^d Abelson & Thomas, 1966.

^e Thomas, 1966.

^f Cairns, 1961

^g Davison & Freifelder, 1962

^h Cummings & Kozloff, 1960.

ⁱ Richardson, 1966.

^j Harpst *et al.*, 1968*b*.

intact phage. With the ^{32}P activity the phosphorus content of the center of decay may be calculated from the number of disintegrations. The mass of the center of decay may then be calculated from the mean residue weight. The only errors which could lead to a high value of molecular weight are in the determination of phosphorus activity, which may be determined with accuracy, and in a possible biasing of the distribution of track numbers in the star. It is conceivable that the "star gazer" may overlook single- or double-track stars and bias the average obtained for the molecular weight.

The peaking of the mean for the T2 distribution (Rubenstein *et al.*, 1961) above the Poisson distribution is evidence for a biased distribution (Kahn, 1964). Kahn attributes this systematic error to the inability to count accurately the rays of large stars, and the background criteria of counting only stars with five or more rays (Levinthal & Thomas, 1957) which is evidently the procedure used by Rubenstein *et al.* (1961).

(f) Length determination

Contour length measurements may be made with sufficient accuracy to provide good relative weights, providing the standard DNA is on the same electron microscope grid or film. This has rarely been the case. A calibration of length to known molecular weights or an assumed linear mass density permits the estimation of molecular weight for the measured length.

Cairn's (1961) T2 DNA molecular weight, $109 \pm 8 \times 10^6$ daltons, from ^3H radioautographic length determinations is in agreement with the values from sedimentation equilibrium and hydrodynamics. Unfortunately, this molecular weight is based on an assumed mass per unit length, corresponding to the Watson-Crick B form, which may be an incorrect assumption for the DNA spread on the film.

Thomas (1966) has calibrated length from microscopy to molecular weight (Fig. 9) and has found this calibration to be consistent with a Watson-Crick B form linear mass density, 196 daltons/Å. From this calibration a molecular weight of 77.6×10^6 daltons was obtained from T5^{sto} DNA and 118×10^6 daltons for T2 DNA.

The standards used were the radioautographic molecular weights for T7 and T2 DNA, and molecular weights from zone centrifugation in sucrose gradients using 130×10^6 daltons for T2 DNA as the standard in the zone centrifugation. This means the T2 standard molecular weight is inconsistent with the length calibration plot by 10%. This inconsistency is not the result of glucosylation of the T2 DNA, as non-glucosylated T*2 DNA had an identical length. Since most of the points used in this length calibration use T2 DNA as the standard molecular weight, the molecular weights assigned to the lengths of T5^{sto} and T2 DNA must be viewed with suspicion.

The molecular weights obtained from hydrodynamics and sedimentation equilibrium have been plotted against the lengths reported by Thomas (1966) in Figure 9. The T-even molecular weight has been converted to a non-glucosylated T*-even weight to preserve the homology of the plot. The length of T5⁺ DNA was obtained by zone centrifugation (Abelson & Thomas, 1966; Thomas, 1966) and the plot of $S_{20,w}$ versus length from microscopy. The linearity of the plot is good and the slope corresponds to a linear mass density of 160 daltons/Å. The data are too crude to use this plot for an accurate calibration, but the molecular weight of P1 appears about 10% too high. The molecular weights of λc^{++} and λcb2b5 DNA's, 31×10^6 daltons and 25×10^6 daltons respectively, agree with this curve to within experimental uncertainty.

FIG. 9. Length from electron microscopy *versus*: ○, molecular weight from autoradiography directly or indirectly (Thomas, 1966); |, molecular weight range from hydrodynamics (Hearst *et al.*, 1968a); ●, molecular weight from sedimentation equilibrium T-even corrected for glucosylation T*-even; □, ϕ X RF 174 molecular weight from light-scattering T5⁺ DNA length from a calibration to the results of microscopy (Thomas, 1966; Abelson & Thomas, 1966). Uncertainty in length determinations indicated by horizontal bars on abscissa (Thomas, 1966).

(g) Zone centrifugation

Relative molecular weights (lengths) may be determined by alkaline sucrose gradient centrifugation. This is done by establishing an empirical relationship for the distance traveled in the gradient to molecular weight (length). Abelson & Thomas (1966) undertook such a calibration using the length of T2 and T7 from electron microscopy as the standard. Although under the conditions of the experiment the DNA is single stranded, it may be empirically related to the double-stranded lengths. The choice of double-stranded T7 and T2 lengths as the standards is convenient, because it allows the calculation of molecular weights from the calibration curve of Figure 9. Having used this calibration, Abelson and Thomas arrive at molecular weights of 81.2×10^6 and 87×10^6 daltons for T5st and T5⁺ DNA (Thomas & MacHattie, 1967). The validity of the calibration in Figure 9 has already been discussed.

Studier (1965) has established a molecular weight scale for both native and denatured DNA, assuming the sedimentation coefficient of both forms is related to molecular weight by the equation:

$$S = KM^a \quad (16)$$

where a was evaluated by the ratio of sedimentation coefficients of λ dg DNA whole and half molecules. The light-scattering molecular weight for ϕ X DNA of Sinsheimer (1959) was assumed to be exact and the sedimentation coefficients for linear double-stranded ϕ X DNA used to determine K .

We believe that the sedimentation coefficient of native DNA cannot be related to molecular weight by equation (16) (Hearst *et al.*, 1968a) and that the extrapolation and (dn/dc) used to determine the ϕ X light-scattering molecular weight leaves that number in question.

(h) *Molecular weight of intact phage*

Since the phosphorus content of the phage lies only in the DNA, the molecular weight of the DNA may be obtained from the molecular weight of the phage by a phosphorus analysis. The relatively compact phage is subject to light-scattering and sedimentation-diffusion studies, permitting an absolute molecular weight determination. Plaque-counting affords an additional technique to determine the phage molecular weight, assuming the phage is 100% viable.

Cummings & Kozloff (1960) have obtained a molecular weight of 215×10^6 daltons for T2 phage from sedimentation-diffusion studies. Based on their phosphorus analysis, 3.9% phosphorus, they obtained a molecular weight of 102×10^6 daltons for T2 DNA. Sedimentation-diffusion studies are in principle exact, and there is no reason to doubt the accuracy of this molecular weight.

Plaque-counting has led to a molecular weight of 160×10^6 daltons for T2 DNA (Rubenstein *et al.*, 1961). These authors believed this high value reflected the incomplete viability of their phage sample.

Davison & Freifelder (1962) studied the molecular weight of T7 phage with a variety of methods: light-scattering, sedimentation-diffusion and plaque counting. The range of their molecular weight, $21.3 \pm 3 \times 10^6$ represents the uncertainty of the phosphate analysis ($\pm 1 \times 10^6$ daltons) and the values obtained by the different techniques (19 to 23×10^6 daltons). This also constitutes an absolute molecular weight determination in accordance with the values we suggest.

(i) *Miscellaneous techniques*

Recent improvements in light-scattering techniques (Harpst *et al.*, 1968*ab*) have permitted the determination of a range for the molecular weight of T7 DNA (22.8 to 27) $\times 10^6$ daltons. In addition to the uncertainty of the reported range, there is a 10% uncertainty in instrument calibration (Harpst *et al.*, 1968*a*). Using a more reliable value of dn/dc for the conditions of the experiment (Cohen & Eisenberg, 1968) instead of the value assumed by Harpst *et al.* would raise the reported molecular weights by 24%. Theoretical calculations by Sharp & Bloomfield (1968) on the angular dependence of the scattering indicate the extrapolation to zero angle will result in molecular weights 5 to 10% too high. The results of this work indicate extrapolation to zero concentration ignoring third virial effects will also lead to molecular weights 5 to 10% too high. Because of these arguments for both raising and lowering the reported molecular weight range, the light-scattering results are not definitive enough to assist us in determining an accurate molecular weight for T7 DNA.

Terminal labeling on the 5' hydroxyl end of T7 DNA pretreated with phosphatase has resulted in a molecular weight of 26×10^6 daltons (Richardson, 1966). The reaction on the pretreated material was believed to go to completion, as the calculated molecular weight agreed with some existing literature. We can see no obvious logical error in this experiment beyond the possibility that either the phosphatase reaction or the terminal labeling did not go to completion. The 11% discrepancy between the density-gradient molecular weight and the terminal labeling molecular weight must therefore be resolved in the future.

(j) *Conclusion*

The molecular weights found by sedimentation equilibrium and hydrodynamics for viral DNA's are in good agreement (Table 2). These molecular weights are also inter-

nally consistent when studied as a function of an experimental parameter such as the sedimentation coefficient or intrinsic viscosity (Hearst *et al.*, 1968a) or length (Fig. 9). The molecular weights we suggest are not without precedent in the literature. They are in good agreement with the molecular weights obtained from the phosphorus content and absolute molecular weight of the phage (Cummings & Kozloff, 1960; Davison & Freifelder, 1962).

We have investigated the conditions required to obtain meaningful molecular weights for DNA by sedimentation equilibrium in a density gradient. It is clearly necessary to perform the experiment at several concentrations and extrapolate to zero DNA concentration. We recommend plotting $\log M_{app}$ versus the average optical density, $\langle o.d. \rangle$, in the band where $M_{app} = [K\langle\delta^2\rangle]^{-1}$ and $\langle\delta^2\rangle$ and $\langle o.d. \rangle$ are obtained by graphical or numerical integration of the data. We believe that double-beam optics and a photoelectric scanner are essential for determining the true base line. In one attempt to provide purposely poor temperature control with the RTIC unit, we observed no harmful effects, so the importance of very careful temperature control is questionable. We have only used titanium centerpieces, but have no good reason to believe they are required for the measurement. It is necessary to have the ultra-violet optical system well aligned in order that small slits can be used (0.06 mm). We do not believe that an electronic speed control is essential to the measurement although we used one.

This research was partially supported by USPH grant GM 11180 and NASA grants NsG 243 and NsG 479. We thank Frank Rinehart for determining the sedimentation coefficients, Dr Marcos Maestre for samples of T4^{ort} and T5⁺ phages and Kathy Barrett for a sample of T7 DNA.

REFERENCES

- Abelson, D. & Thomas, C. A. Jr. (1966). *J. Mol. Biol.* **18**, 262.
 Cairns, J. (1961). *J. Mol. Biol.* **3**, 756.
 Cohen, G. & Eisenberg, H. (1968). *Biopolymers*, **6**, 1077.
 Crothers, D. & Zimm, B. (1965). *J. Mol. Biol.* **12**, 525.
 Cummings, D. J. (1963). *Biochim. biophys. Acta*, **72**, 475.
 Cummings, D. J. & Kozloff, L. M. (1960). *Biochim. biophys. Acta*, **44**, 445.
 Davison, P. & Freifelder, D. (1962). *J. Mol. Biol.* **5**, 635.
 Erikson, R. L. & Szybalski, W. (1964). *Virology*, **22**, 111.
 Flory, P. (1949). *J. Chem. Phys.* **17**, 1347.
 Flory, P. (1953). In *Principles of Polymer Chemistry*, chapters 7, 12. Ithaca, New York: Cornell University Press.
 Flory, P. & Krigbaum, W. (1950). *J. Chem. Phys.* **18**, 1086.
 Gray, H. B., Bloomfield, V. A. & Hearst, J. E. (1967). *J. Chem. Phys.* **46**, 1493.
 Gray, H. B. & Hearst, J. E. (1968). *J. Mol. Biol.* **25**, 111.
 Harpst, J., Krasna, A. & Zimm, B. (1968a). *Biopolymers*, **6**, 585.
 Harpst, J., Krasna, A. & Zimm, B. (1968b). *Biopolymers*, **6**, 595.
 Harris, R. A. & Hearst, J. E. (1966). *J. Chem. Phys.* **44**, 2595.
 Hearst, J. E., Beals, E. & Harris, R. A. (1968b). *J. Chem. Phys.* **48**, 5371.
 Hearst, J. E. & Gray, H. B. (1968). *Analyt. Biochem.* **24**, 70.
 Hearst, J. E., Harris, R. A. & Beals, E. (1966). *J. Chem. Phys.* **45**, 3106.
 Hearst, J. E., Ifft, J. & Vinograd, J. (1961). *Proc. Nat. Acad. Sci., Wash.* **47**, 1015.
 Hearst, J. E., Schmid, C. & Rinehart, F. (1968a). *Macromolecules*, **1**, 491.
 Hearst, J. E. & Stockmayer, W. (1962). *J. Chem. Phys.* **37**, 1425.
 Hearst, J. E. & Vinograd, J. (1961a). *J. Phys. Chem.* **65**, 1069.
 Hearst, J. E. & Vinograd, J. (1961b). *Proc. Nat. Acad. Sci., Wash.* **47**, 999.

- Hearst, J. E. & Vinograd, J. (1961c). *Proc. Nat. Acad. Sci., Wash.* **47**, 1005.
- Kahn, P. L. (1964). *J. Mol. Biol.* **8**, 392.
- Levinthal, C. & Thomas, C. A. (1957). *Biochim. biophys. Acta*, **23**, 453.
- MacHattie, L. A., Ritchie, D. A., Thomas, C. A. & Richardson, C. C. (1967). *J. Mol. Biol.* **23**, 355.
- Meselson, M., Stahl, F. & Vinograd, J. (1957). *Proc. Nat. Acad. Sci., Wash.* **43**, 581.
- Richardson, C. C. (1966). *J. Mol. Biol.* **15**, 49.
- Ritchie, D. A., Thomas, C. A., MacHattie, L. A. & Wensink, P. C. (1967). *J. Mol. Biol.* **23**, 365.
- Rubenstein, I., Thomas, C. A. & Hershey, A. D. (1961). *Proc. Nat. Acad. Sci., Wash.* **47**, 1113.
- Schumaker, V. W. & Wagnild, J. (1965). *Biophys. J.* **5**, 947.
- Sharp, P. & Bloomfield, V. (1968). *Biopolymers*, **6**, 1201.
- Sinsheimer, R. L. (1959). *J. Mol. Biol.* **1**, 43.
- Smiriga, S. & Hearst, J. E. (1969). *Rev. Sci. Instr.* **40**, 233.
- Stockmayer, W. & Casassa, E. F. (1952). *J. Chem. Phys.* **20**, 1560.
- Studier, F. W. (1965). *J. Mol. Biol.* **11**, 373.
- Sueoka, N. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1480.
- Sueoka, N. (1961). *J. Mol. Biol.* **3**, 31.
- Thomas, C. A. (1966). *Macro. Met. J. Gen. Physiol.* **49**, No. 6, part II, 143.
- Thomas, C. A. & MacHattie, L. (1967). *Ann. Rev. Biol.* **36**, part II, 485.
- Thomas, C. A. & Pinkerton, T. (1962). *J. Mol. Biol.* **5**, 356.
- Thomas, C. A. & Rubenstein, I. (1964). *Biophys. J.* **4**, II, 93.
- Vinograd, J. & Hearst, J. E. (1962). *Prog. Org. Nat. Prod.* **20**, 372.