The Three-Dimensional Structure of Supercoiled Deoxyribonucleic Acid in Solution

EVIDENCE OBTAINED FROM THE ANGULAR DISTRIBUTION OF SCATTERED LIGHT

By DOUGLAS J. JOLLY and AILSA M. CAMPBELL

Institute of Biochemistry, University of Glasgow, Glasgow G12 8QQ, U.K.

(Received 26 January 1972)

1. The size and shape of superhelical double-stranded circular DNA from bacteriophage ϕ X174 were investigated by light-scattering. The molecular weight of the DNA is 3.17×10^6 and the root-mean-square radius is 103.5 nm. 2. The light-scattering envelopes of various theoretical three-dimensional models for such DNA molecules were calculated by repetitive computational techniques, and the results were compared with the experimental findings. 3. It is concluded that the structure of supercoiled DNA containing -12 superhelical turns in buffer of $I_0.2$ corresponds best to one of the more compact models for superhelix structure such as the branched model, and the commonly employed straight interwound superhelix model is incompatible with the experimental results, at the superhelix density found.

The structure of covalently closed circular duplex DNA has undergone much investigation in the analytical ultracentrifuge (Bauer & Vinograd, 1968, 1970*a,b*) and recently also by viscometry (Revet *et al.*, 1971; Smit & Borst, 1971). These two techniques have shown clearly that the structure is much more compact than that of relaxed circular DNA of the same molecular weight, but the exact nature of the tertiary structure remains obscure. For convenience, the supercoiled DNA is usually represented as a straight interwound superhelix (Fig. 1*a*), although other structures such as the toroid (Fig. 1*b*) or a variety of branched structures, which can be observed in the electron microscope (Fig. 1*c*), have been considered possible.

Light-scattering provides a very valuable tool for distinguishing between some of these possible structures, since it is one of the few techniques that allows the unambiguous measurement of the root-mean-square radius of the molecule. The angular variation of the reciprocal particle-scattering factor $P(\theta)^{-1}$ can then be plotted and compared with the predicted curves for the various models. However, since large amounts of material are required for such experiments, ultracentrifugal analysis remains the more common technique.

In this paper we report a light-scattering analysis of the structure of a small superhelical circular DNA species, ϕ X174 RF I*. In the molecular-weight

range of this DNA, light-scattering experiments do not have to be carried to the very low angles required for the large DNA molecules (Harpest *et al.*, 1968) and it is possible to make accurate determinations of the molecular weight and root-mean-square radius. From these results and from the angular variation of scattering, we have been able to make an analysis of the compact structure of the superhelix.

A preliminary report of this work has been given (Campbell & Jolly, 1972).

Materials and Methods

Preparation and analysis of DNA

The replicative double-stranded form of bacteriophage ϕ X174 DNA was prepared by the method of Rush *et al.* (1967) from infected cells of *Escherichia coli* C. The final separation of the forms I and II was achieved by the method of Hudson *et al.* (1969). The complete removal of dye from DNA solutions was checked by a fluorescence assay, capable of detecting down to 10 ng of dye/ml, before light-scattering. The superhelix density was determined by ethidium bromide titration (Bauer & Vinograd, 1968) as described by Waring (1970) in a Spinco model E analytical ultracentrifuge. This determines not the number of 'physical superhelical turns' but the number of 'superhelical turns' (Wang, 1969), which may not be exactly the same. The difference, however, is thought to be small in supercoiled DNA that is not very tightly twisted, particularly in the buffer used (Maestre & Wang, 1971; Campbell & Lochhead, 1971), and the two numbers have been considered

* Abbreviations: ϕ X174 RF I, intact (un-nicked) circular duplex intracellular replicative form of bacteriophage ϕ X174 DNA; ϕ X174 RF II, nicked circular duplex intracellular replicative form of bacteriophage ϕ X174 DNA.

Fig. 1. Hypothetical structures for various forms of supercoiled DNA of 12 superhelical turns

(a) Rigid interwound superhelix; (b) toroid; (c) two forms of Y-shaped structure, one with all the arms the same length and the other with arms of different lengths.

interchangeable. All experiments were performed under nuclease-free conditions, with sterilized glassware, and the DNA was analysed in the ultracentrifuge both immediately before and after the light-scattering measurements were taken to ensure that only form I was present. This procedure is estimated to detect down to 3% contamination by form II.

Light-scattering measurements

All light-scattering experiments were performed in a Fica 50 photogoniometer (1970 model; Sofica, Paris, France) at $25^{\circ} \pm 0.2^{\circ}\text{C}$ by using a truncated conical cell manufactured to our specifications by Barr and Stroud Ltd., Glasgow, U.K. and calibrated with Ludox (Maron & Lou, 1954; Goring *et al.*, 1957). The Ludox was a gift from Dupont Industrial Chemical Division, Wilmington, Delaware, U.S.A. This calibration with aqueous suspensions was deemed necessary to avoid the large, rather uncertain, refractive-index correction that would be necessary if the makers' calibration with benzene was used. The precautions recommended by Kratochvil

et al. (1962) were taken. With this cell, experiments could be performed with only 3ml of scattering solution. The wavelength of light used was 546nm. All transfers of material were carried out in a dust-free box in a dust-free room. All solutions were clarified by passage through Millipore filters ($0.45\mu\text{m}$ pore size) or Sartorius Membranfilters (25mm diam.), and the dust-free solutions were allowed to run slowly down the sides of the cells. The two types of filter were found to be interchangeable. These procedures gave low and stable background readings for solvent and stable and reproducible readings for DNA. Since the small circular DNA used is not shear-sensitive, filtration is a more suitable method of clarification than is centrifugation. All experiments were performed at constant chemical potential (Cassassa & Eisenberg, 1964).

The concentration of the DNA was measured in a Cary 15 spectrophotometer by using cells of 0.2cm and 1cm path-length and analysing the results by the method of Hirschman & Felsenfeld (1966). This gave an $\epsilon_{(p)}$ (molar phosphorus extinction coefficient) of 6415 with s.d. 2%. The average monomer molecular weight of DNA (sodium salt) was taken as 331. The value for the refractive-index increment of DNA (sodium salt) used was 0.166ml/g (Krasna, 1970).

Buffer. The buffer employed in this experiment was BPES (6mM- Na_2HPO_4 , 2mM- NaH_2PO_4 , 1mM-disodium EDTA and 0.179M- NaCl , pH6.8).

Analysis of data. The scattering curves of the various DNA models were all computed in a PDP/8L digital computer.

Results

Molecular weight

Fig. 2 shows a typical Zimm (1948) plot of covalently closed circular duplex DNA, from bacteriophage ϕX174 , in BPES buffer. The molecular weight from the double extrapolation (mean of three experiments) is $3.17 \times 10^6 \pm 0.19 \times 10^6$. Previous light-scattering investigations on the single-stranded molecule (Sinsheimer, 1959) have suggested that the molecular weight of this form is 1.7×10^6 , and hence the value for the double-stranded form should be 3.4×10^6 . Sinsheimer (1959), however, used a value of dn/dc (refractive-index increment) of 0.201ml/g at 436nm. More recent work has suggested that the value is very much lower (Krasna, 1970; Cohen & Eisenberg, 1968) and Sinsheimer's (1959) results should have given a molecular weight of about 4.5×10^6 for the double-helical form by using the corrected value of 0.175ml/g for dn/dc at 436nm. Our value of 3.17×10^6 is considerably lower than this, a fact that can probably be attributed to stringent modern dust-clearing precautions rather than to DNA breakdown, since this is very easily detected in the ultracentrifuge.

Fig. 2. A typical Zimm (1948) plot of DNA from the replicative form, RF1, of bacteriophage ØX174 in BPES buffer

The formula (Geiduschek & Holtzer, 1958):

$$Kc/R(\theta)_{c=0} = \frac{1}{M} \left[1 + \frac{16\pi^2 R^2}{3(\lambda')^2} \sin^2(\theta/2) + \dots \right]$$

allows the estimation of R , the root-mean-square radius, from the initial slope of the curve for $Kc/R(\theta)$ at zero concentration, and M , the molecular weight, from the intercept. The formula $Kc/R(\theta)_{\theta=0} = 1/M + 2Bc + \dots$ also allows estimation of the second virial coefficient, B , and the molecular weight. $K = 2\pi^2 n^2 (dn/dc)^2 / N\lambda^4$ is a constant for any one type of molecule in a particular solvent, where n is the refractive index of the solvent, N is Avogadro's number, λ is the wavelength of light *in vacuo*, c is the concentration (values for c on the figure are in $\mu\text{g/ml}$); $R(\theta) = r^2 i_\theta / [I_0(1 + \cos^2\theta)]$, where i_θ is the intensity of scattered light at any angle θ , I_0 is the intensity of incident light, and r is the distance of the receiver from the scattering centre. All units are in the c.g.s. system. ●, Experimental points; ○, extrapolated points.

Root-mean-square radius and virial coefficient

From Fig. 2 the root-mean-square radius can be computed to be $103.5 \pm 12 \text{ nm}$. Although theoretical predictions have been made about this value for relaxed circular molecules, relative to the equivalent linear molecules, and these values have been tested by viscosity (Douthart & Bloomfield, 1968) and light-scattering (Dawson, 1970), they bear little relation to the values that can be expected for the closed circular molecule, which is forced into the

supercoiled conformation. No theoretical treatment of the light-scattering properties of such a molecule has yet been attempted.

The second virial coefficient, B , is a measure of the deviation of the solution from ideality and theta conditions (Flory, 1953), when $B=0$. It has been shown to have a small negative value of $-2.6 \times 10^{-4} \text{ mol}\cdot\text{cm}^3\cdot\text{g}^{-2}$. Using much larger DNA molecules in linear form, Harpst *et al.* (1968), in contrast to earlier work, have also found a non-zero but positive second virial coefficient in this solvent. Some degree of interaction among large linear molecules is not an unreasonable hypothesis, but the fact that the small supercoiled DNA molecules show interactions in the opposite sense suggests that this may be an inherent property of this type of DNA molecule.

Number of superhelical turns

The number of superhelical turns was found to be -11.8 ± 1.6 , which corresponds to a superhelix density (Bauer & Vinograd, 1968) of -0.025 ± 0.003 . The critical binding ratio, where forms I and II appear equivalent, was $0.037 \pm 0.005 \text{ mol}$ of ethidium bromide/nucleotide. The necessity of performing the titration by boundary sedimentation increased the error somewhat. Since our models and computer programmes require the number of superhelical turns to be an integer, the value used in the following calculations was -12 , unless otherwise stated.

It has been suggested (Paoletti & Le Pecq, 1971) that ethidium bromide winds rather than unwinds the DNA helix by about the same amount, and hence the superhelical turns are positive.

Without commenting on the relative merits of the evidence for the sign of the superturns, it is sufficient to note that the light-scattering experiments and computer models described here are not affected by this, leaving our results and conclusions unchanged.

Calculation of the particle-scattering factor

The particle-scattering factor, $P(\theta)$, was calculated in all cases from the equation:

$$P(\theta) = \frac{1}{N^2} \sum_n^N \sum_m^N (\sin hr_{nm} / hr_{nm})$$

where N is the total number of scattering points in the molecule, h is $4\pi \sin(\theta/2)/\lambda'$ (θ being the angle at which the scattered intensity is observed), r_{nm} is the absolute value of the vector between scattering points n and m (Geiduschek & Holtzer, 1958), and λ' is the wavelength of light in solution. Three basic types of model were considered: the straight interwound superhelix, the toroidal model and the simple branched structure, two types of which are shown in Fig. 1. Our approach was to calculate the dimensions

of these models from the molecular weight and root-mean-square radius. $P(\theta)^{-1}$ was then calculated for each model and compared with the experimental values.

The problem is to formulate r_{nm} by some general expression for all n, m values. Each r_{nm} is then evaluated numerically, substituted into the above expression and a running total kept. This can be easily done on a modern digital computer, and the calculation repeated at various designated angles.

The contour length of the molecule was taken as 1625.6 nm. This was calculated from the molecular weight by assuming a linear mass density of 1950 daltons/nm, i.e. the DNA duplex is assumed to be in the B form in the supercoil (Maestre & Wang, 1971; Campbell & Lochhead, 1971).

Straight interwound model

The $P(\theta)^{-1}$ curve was calculated for the model in Fig. 1(a). The model molecule contained 12 cross-overs (i.e. physical superhelical turns), and was treated as a rod of length $L = (12R_g^2)^{1/2}$. The equation

$$d/2 = [(P/2K)^2 - 12(R_g/K)^2]^{1/2}/\pi$$

where d is the diameter of the cylinder, K is the number of superhelical turns, R_g is the root-mean-square radius and P is the contour length, was used to calculate d .

The molecule was then imagined to be sliced into $4K$ segments with centres of mass in a straight line up the middle; each segment was of small enough dimension to be considered a point scatterer.

There are then $2(4K-1)$ interpoint distances that are $L/4K$, $2(4K-2)$ that are $2L/4K$, $2(4K-3)$ that are $3L/4K$ etc. The factor 2 in the expression $2(4K-1)$ is to allow for the fact that both the distances r_{ab} and

r_{ba} must be counted. Thus the molecule was assumed to be essentially rigid, a reasonable assumption for such a small DNA molecule (Cohen & Eisenberg, 1966), and end effects were neglected. From the above formulae $d/2 = 19.35$ nm and $L = 358.5$ nm. The reciprocal particle-scattering function, $P(\theta)^{-1}$, calculated from this is shown in Fig. 4.

The toroidal model

The toroidal model is shown in Fig. 1(b). The dimensions were related to the root-mean-square radius as shown in the two following equations (Tanford, 1961):

$$R_g^2 = \sum_i m_i b_i^2 / \sum_i m_i$$

If all m_i are equal,

$$R_g^2 = \sum_i b_i^2 / N$$

where b_i is the distance of the i th mass element of mass m_i from the centre of mass, N is the number of mass elements and R_g is the root-mean-square radius. The centre of mass of the toroid is obviously the centre of the large ring of the toroid. Fig. 3(a) shows a cross-section of the toroid, so that:

$$\begin{aligned} b_i^2 &= R^2 + r^2 + 2Rr \cos(n_i) \\ &= c + m \cos(n_i) \end{aligned}$$

where $c = R^2 + r^2$ and $m = 2Rr$.

$$\therefore R_g^2 = \frac{1}{N} \sum_{i=0}^N \left[c + m \cos\left(\frac{2\pi Ki}{N}\right) \right]$$

Fig. 3. Geometry of toroid model (Fig. 1b) used to calculate the general root-mean-square radius, and $P(\theta)$ function

The meaning of the symbols is defined in the Results section on the toroid model. (a) Cross-section of toroid on z and y axes. (b) Cross-section of toroid on x and y axes. (c) Variation of the angle between adjacent elements (m_s) with distance along the contour length (r_i).

Fig. 4. Comparison of computer-drawn curves for toroid and straight interwound models with experimental data

The reciprocal particle-scattering factors of the various DNA models are compared with the experimental points. Vertical bars represent the spread of experimental values obtained at a particular angle. Curve (1): toroidal model from experimental data; R_g 103.5 nm, 12 superhelical turns. Curve (2): rod model from experimental data; R_g 103.5 nm, 12 superhelical turns. Curves (3) and (4): toroidal model maximum error limits; R_g 116.0 nm, 10 superhelical turns; R_g 91.5 nm, 13 superhelical turns. Curves (5) and (6): rod model maximum error limits; R_g 116.0 nm, 10 superhelical turns; R_g 91.5 nm, 13 superhelical turns.

where K is the number of superhelical turns. This summation can conventionally be changed to an integration, where N is large.

$$\begin{aligned} \therefore R_g^2 &= \frac{1}{N} \int_0^N \left[c + m \cos\left(\frac{2\pi K_i}{N}\right) \right] di \\ &= c + \frac{m}{2K} \sin(2\pi K) \end{aligned}$$

But K is an integer

$$\therefore R_g^2 = c = R^2 + r^2$$

R and r can also be related by imagining the toroidal structure to be cut at one point and straightened out into a single superhelix of axial length $2R$.

Then:

$$2\pi R = [P^2 - (2\pi Kr)^2]^{\frac{1}{2}}$$

where P is the contour length of the molecule. Combining the above equations gives:

$$K = [P^2 - (2\pi KR_g)^2 / 4\pi^2(1-K^2)]^{\frac{1}{2}}$$

Hence both R and r can be obtained from R_g if P is known.

The derivation of the interpoint distances which follows is largely based on the treatment by Gray (1967). A more exhaustive treatment is available (Fuller, 1971), but in our case was thought to be not worthwhile, because first, of the complexity of the mathematical manipulations involved and, secondly, the resolution of our method did not seem to merit it.

The problem with this model is that for each successive identical distance along the contour of the molecule, the angle (m in Fig. 3b) swept out by the large radius varies smoothly about a mean value. This can be represented graphically (Fig. 3c).

Here:

$$(m_s)_{av.} = 2 \arcsin(\pi/N)$$

and generally:

$$m_s = 2 \arcsin[2\pi R/2N(R+r \cos n)]$$

$$\text{Hence } (m_s)_{\max.} = 2 \arcsin[2\pi R/2N(R-r)]$$

$$(m_s)_{\min.} = 2 \arcsin[2\pi R/2N(R+r)]$$

The variation has a sinusoidal appearance, and this formulation was used to describe the variation. Thus $[m_i - i(m_s)_{av.}] = t \cos[(r_i 2\pi K)/P]$ where m_i is the angle m at the i th element, r_i is the distance along the contour length to the i th element and t is a constant. When $r_i = P/K$, $i = N/K$ and then:

$$\begin{aligned} [m_i - i(m_s)_{av.}] &= (m_s)_{\max.} - (m_s)_{av.} \\ &= 2 \arcsin[2\pi R/2N(R-r)] \\ &\quad - 2 \arcsin(\pi/N) \\ &= t \cos(2\pi) \\ &= t \end{aligned}$$

Therefore generally:

$$\begin{aligned} m_i &= 2 [\arcsin(\pi R/N(R-r)) - \arcsin(\pi/N)] \cos \\ &\quad [(r_i 2\pi K)/P] + i[2 \arcsin(\pi/N)] \\ &= 2 [\arcsin(\pi R/N(R-r)) - \arcsin(\pi/N)] \cos \\ &\quad [(r_i 2\pi K)/P] + i(2\pi/N) \text{ if } N \text{ is large} \end{aligned}$$

The second term is the expression expected for m_i if each step $(r_{i+1} - r_i)$ corresponded to an equal increment in m_i . The first term represents the variation in the change of r_i . The angle $n_i = 2\pi Ki/N$ for all i . By using the axes in Figs. 3(a) and 3(b), the parametric equations describing the toroid are:

$$\begin{aligned} x &= [R + r \cos(n)] \sin m \\ y &= [R + r \cos(n)] \cos m \\ z &= r \sin n \end{aligned}$$

Then any interpoint distance:

$$r_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{\frac{1}{2}}$$

which leads to the expression:

$$\begin{aligned} r_{ij}^2 &= [1 - \cos(m_i - m_j)] \{2R^2 + 2Rr [\cos(n_i) + \\ &\quad \cos(n_j)]\} + \{2r^2 [1 - \cos(m_i - m_j) \cos(n_i) \\ &\quad \cos(n_j) - \sin(n_i) \sin(n_j)]\} \end{aligned}$$

The model was divided into 15 K scattering elements of equal mass. These were treated as having their centres of mass at the midpoint of the contour length they occupied.

From the above formulae, the large radius was 101.6 nm and the small radius 19.8 nm when the model was fitted to our experimental root-mean-square radius and molecular weight. The reciprocal particle-scattering function, $P(\theta)^{-1}$, calculated from this is shown in Fig. 4.

Y-shaped model

This model is shown in Fig. 1(c) and is the supposed three-dimensional solution structure of the Y-shapes frequently noted in electron micrographs of superhelical DNA. The angle between the arms and their respective lengths could be varied in the computer model and a number of different sets of values for these parameters were used to compare with the experimental results. The model was then treated as three linked cylinders of similar form to the rod model. For the sake of generality of applicability of the computer programme to all Y-shapes the defining parameter used was d , the diameter of the cylinders; d was chosen and empirically varied until R_g was obtained at the same as the experimental value. R_g was calculated by computer from the formula (Tanford, 1961; Flory, 1969):

$$R_g^2 = \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N r_{ij}^2$$

For Y-shapes where R_g was easily arithmetically calculable, the computer-calculated R_g values were

Table 1. Geometry of Y-shapes as dictated by the experimental root-mean-square radius

Dimensions of three possible Y-shapes (see Fig. 1c) that would have the experimental root-mean-square radius of 103.5 nm. The symbols are explained in the Results section on Y-shaped models.

Features chosen						Resulting dimensions (nm)				Curve no. in Fig. 5
N_1	N_2	N_3	m_{12}	m_{23}	m_{31}	d	L_1	L_2	L_3	
4	4	4	120°	120°	120°	31.1	185.6	185.6	185.6	1
2	4	6	120°	120°	120°	32.0	84.2	168.4	252.6	4
4	4	4	90°	90°	90°	30.0	194.8	194.8	194.8	5

Fig. 5. Comparison of computer-drawn curves for Y-shaped models with experimental data

The reciprocal particle-scattering factors for the Y-shaped model are compared with the experimental points. Vertical bars represent the spread of experimental values obtained at a particular angle. Curve (1): simple Y from experimental data; R_g 103.5 nm, four superhelical turns in each arm, 120° angles between each pair of arms. Curve (2): maximum error for Y; R_g 116.0 nm and four, three and three superhelical turns in the three arms, 120° between each pair of arms. Curve (3): maximum error for Y; R_g 91.5 nm and four, four and five superhelical turns in the three arms, 120° between each pair of arms. Curve (4): unsymmetrical Y from experimental data; R_g 103.5 nm, two, four and six superhelical turns in the three arms, 120° angles between each pair of arms. Curve (5): non-planar Y from experimental data; R_g 103.5 nm, four superhelical turns in each arm, 90° between each pair of arms.

the same. With d defined the length of each arm was calculated from the formula:

$$L_c = N_c [(L/2N)^2 - (d/2)^2]^{\frac{1}{2}}$$

where L_c and N_c are the length and number of supercoils respectively in the arm.

If the arms are numbered 1, 2, 3 with N_1 , N_2 and N_3 the numbers of supercoils respectively in each arm, and m_{12} , m_{23} , m_{31} are the angles between arms 1 and 2, 2 and 3, 3 and 1 respectively, then the set of dimensions shown in Table 1 is dictated for the various variations on the Y-shape chosen. For calculation of $P(\theta)$ (and of R_g), the model was, as in the rod model, imagined to be sliced into four segments per super-

turn. The intersegment distances r_{ij} were calculated as for the rod model for the segments i , j on the same arm. If they were on different arms r_{ij} was calculated from:

$$r_{ij} = (r_{i0}^2 + r_{j0}^2 - 2r_{i0} \cdot r_{j0} \cos m)^{\frac{1}{2}}$$

where m is the angle between the arms on which i and j are, r_{i0} and r_{j0} are the distances from the i th and j th segment respectively to the intersection of the three arms.

Empirically one would imagine that the minimum thermodynamic energy state for the Y-shaped model would be that with all three arms planar, and most of the models chosen assume this. However,

conceivably some torsional strain in the structure might cause the lowest energy conformation to be non-planar, and we have therefore included an example of this in our calculated curves.

The molecule has again been assumed to be essentially rigid and end effects have been neglected, and the curves are shown in Fig. 5.

Errors

Molecular weight. In all cases the molecular weight was between 3.1×10^6 and 3.2×10^6 . The error in the molecular weight can be estimated as $\pm 6\%$. This is broken down as follows.

(a) Ludox calibration constant, $\pm 2\%$. This compared favourably with calibration from benzene, as the absolute value of the Rayleigh ratio of benzene is hardly known to this accuracy, the best estimates ranging from 15.6×10^{-6} to 16.3×10^{-6} (Kratochvil *et al.*, 1962).

(b) Concentration measurements, $\pm 2\%$.

(c) Instrument readings, $\pm 1\%$. The instrument has a digital output and this is integrated over a period of 10s. This and the low dust levels that we consider we have achieved lead to a low instrument reading error.

(d) Value of dn/dc used, $\pm 1\%$. The uncertainty in the intercept of the Zimm plot is the cumulation of these errors. The extrapolations were by eye; $(Kc/R\theta)_{c=0}$ and $(Kc/R\theta)_{\theta=0}$ meet as expected at the intercept. The concentration range of the experimental points was rather smaller than is ultimately desirable owing to experimental limitations. However, correct extrapolation to zero concentration is thought to have been achieved since measurements at a large number of different concentrations were made, and the extrapolations were repeatable. Extrapolation to zero angle is also thought correct, since $P(\theta)^{-1}$ at our lowest angle, 30° , was approx. 1.1, which is considered as quite adequately near 1.0, the value of $P(\theta)^{-1}$ at zero angle.

Root-mean-square radius. The error in the radius of gyration is somewhat larger, owing to the difficulty of estimating the initial slope of curve; it is estimated as $\pm 12\%$.

Number of superhelical turns. The other main source of error that could affect the results in Fig. 3 is in the number of superhelical turns. As mentioned above the number of superhelical turns has been taken as equal to the number of physical superhelical turns. This should be nearly true at the superhelix density (-0.025) determined. There is also an error in the titration experiment itself ($\pm 14\%$).

In an effort to determine the effect of these errors on the resolution of our method we have plotted the $P(\theta)^{-1}$ curve calculable for all the models by using the maximum error data, in Figs. 4 and 5 ($R_g \pm 12\%$, the number of superhelical turns $\pm 14\%$). This shows

that our results establish some real differentiation between possible superhelical structures of DNA in solution.

Comparison of the models

Figs. 4 and 5 show the relationship between the calculated reciprocal particle-scattering factors and the experimental ones. In each case the maximum error limits for each model are also shown. It can be seen that, in BPES buffer, the rigid interwound superhelix is totally incompatible with the experimental results, and that one of the more compact structures is in better agreement. The toroidal model is also essentially incompatible with the experimental data, although the lower limit of the maximum error passes quite close to it. Of the three chosen models the simple branched Y-shape with arms of equal length and angles of 120° between each arm provides the best correlation with the experimental data. Even this model, however, deviates quite substantially from the data at high angles and probably is oversimplified. Variation of the length of the arms and the angle between them does not improve the fit of the model to the experimental points and two examples of this variation are also shown (Fig. 5).

Discussion

The rigid double-helical structure of the supercoil has frequently been invoked as a convenient descriptive model of superhelix structure. Nevertheless, branched structures such as the Y-shape are frequently observed in electron micrographs (see for example, Follet & Crawford, 1967), and the toroidal model has been suggested to be compatible with the sedimentation behaviour at low ionic strength when any crossover between two areas of the highly charged double helix would be most unfavourable (Gray, 1967). The fact that such a structure is not visible under the electron microscope is clearly not evidence that it cannot exist in solution.

The protein monolayer technique (Kleinschmidt & Zahn, 1959), which is normally employed, has the defect of resulting in the visualization of a DNA-cytochrome *c* complex that is less highly charged than is DNA in solution, and is viewed in two dimensions. Upholt *et al.* (1971) have attempted to correlate the tertiary structure (and its transitions) of virus SV40 DNA with sedimentation data. Because the radius of gyration cannot be measured in the ultracentrifuge and an interference curve cannot be obtained by this technique, these attempts have, of necessity, relied heavily on electron microscopy. Thus sedimentation studies are severely limited in this application.

The data presented in this paper show that the rigid double-superhelical structure is totally in-

compatible with the experimental data from light-scattering, and that the toroidal model is also unsuitable. The branched structures that we have envisaged as simple Y-shapes show the greatest agreement with the experimental points, although the high angle-deviation would suggest that this also is too simple a model. As variation of the length of the arms and angles between them does not improve the fit of the data, some more complex alterations in the model must clearly be necessary. Possibly further branching of the structure could occur. It is noteworthy that branched structures have recently been invoked to explain the sedimentation behaviour of bacteriophage λ DNA at high ionic strength (Hinton & Bode, 1971).

The slightly negative second virial coefficient that we have found for this DNA is noteworthy. It denotes a change in the macromolecule-small-molecule interaction compared with normal linear DNA, which seems therefore to be linked to the change in tertiary structure. One consequence of this is that the partial specific volume, density increment and other such properties of superhelical DNA should be different from normal linear DNA in this buffer, and hence adoption of well-established values with respect to normal DNA for such parameters may affect the accuracy of hydrodynamic experiments with supercoiled DNA. The refractive-index increment, however, will almost definitely remain the same, since conformational changes such as denaturation appear to have a negligible effect on this parameter (Krasna, 1970).

Dean & Lebowitz (1971) have obtained results suggesting that in superhelical DNA of superhelix density -0.036 there is rupture of Watson-Crick hydrogen bonds, i.e. considerable perturbation of the secondary structure. Although our DNA has a superhelical density of -0.025 and hence considerably less torsional strain in maintaining the Watson-Crick base-pairing, the same type of event may be occurring. This can easily be imagined as leading to a decrease in the small-molecule solution structure owing to increased exposure of the hydrophilic bases (Sinanoglu & Abdunur, 1964), and hence to a decrease in the virial coefficient.

Clearly further studies, by the scattering techniques, by varying solution conditions, observing any conformational transitions and alterations in the second virial coefficient, should enable further resolution of the superhelix structure.

Transitions in tertiary structure and the tertiary structure itself in this type of DNA are reflections of the secondary structure and its alterations, leading to a magnification of the effect of ionic strength, temperature and other changes in conditions on the secondary structure. This means that these effects are far more easily studied by a number of conventional techniques, e.g. light-scattering, ultracentrifugation, viscosity, than in normal linear DNA. This should lead to further elucidation of the stability of the DNA helix and the possible winding and unwinding mechanisms and conditions *in vivo*.

This type of tertiary structure also occurs in a number of oncogenic viruses, e.g. SV40, polyoma, papilloma. The elucidation of any peculiarities in this type of structure is therefore obviously a prerequisite to any large understanding of the modes of action and replication of these viruses.

Conclusion

Our conclusion is therefore that superhelical DNA at a superhelix density of -0.025 , is best fitted by a Y-shape model, possibly with equal arm-lengths spaced at 120° intervals.

We acknowledge the help of Dr. K. Sargeant of the Microbiological Research Establishment, Porton, for the provision of bacteriophage-infected cells. We are also grateful to Professor J. N. Davidson, F.R.S., and Professor R. M. S. Smellie for their interest and for providing facilities with the aid of a grant from the Cancer Research Campaign. The Fica 50 scattering instrument was obtained with a grant from the Science Research Council. D. J. J. is in receipt of a Medical Research Council research studentship.

References

- Bauer, W. & Vinograd, J. (1968) *J. Mol. Biol.* **33**, 141-171
- Bauer, W. & Vinograd, J. (1970a) *J. Mol. Biol.* **47**, 419-435
- Bauer, W. & Vinograd, J. (1970b) *J. Mol. Biol.* **54**, 281-298
- Campbell, A. M. & Jolly, D. J. (1972) *Biochem. J.* **127**, 39p
- Campbell, A. M. & Lochhead, D. S. (1971) *Biochem. J.* **123**, 661-663
- Cassassa, E. F. & Eisenberg, H. (1964) *Advan. Protein Chem.* **19**, 287-393
- Cohen, G. & Eisenberg, H. (1966) *Biopolymers* **4**, 429-440
- Cohen, G. & Eisenberg, H. (1968) *Biopolymers* **6**, 1077-1100
- Dawson, J. R. (1970) *Fed. Proc. Fed. Amer. Soc. Exp. Biol.* **29**, 870
- Dean, W. W. & Lebowitz, J. (1971) *Nature New Biol. (London)* **231**, 5-8
- Douthart, R. J. & Bloomfield, V. A. (1968) *Biopolymers* **6**, 1297-1309
- Flory, P. J. (1953) *Principles of Polymer Chemistry*, Cornell University Press, Ithaca
- Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules*, pp. 4, 383, Interscience, New York
- Follet, E. A. C. & Crawford, L. V. (1967) *J. Mol. Biol.* **28**, 455-459
- Fuller, F. B. (1971) *Proc. Nat. Acad. Sci. U.S.A.* **68**, 815-819
- Geiduschek, E. P. & Holtzer, A. (1958) *Advan. Biol. Med. Phys.* **6**, 431-551
- Goring, D. A. I., Senez, M., Melanson, B. & Huque, M. M. (1957) *J. Colloid Sci.* **12**, 412-416
- Gray, H. B. (1967) *Biopolymers* **5**, 1009-1019

- Harpst, J. A., Krasna, A. I. & Zimm, B. H. (1968) *Biopolymers* **6**, 585-594
- Hinton, D. M. & Bode, V. C. (1971) *Fed. Proc. Fed. Amer. Soc. Exp. Biol.* **30**, 1095
- Hirschman, S. Z. & Felsenfeld, G. (1966) *J. Mol. Biol.* **16**, 347-358
- Hudson, B., Upholt, W. B., Devinny, J. & Vinograd, J. (1969) *Proc. Nat. Acad. Sci. U.S.* **62**, 813-820
- Kleinschmidt, A. & Zahn, R. K. (1959) *Z. Naturforsch. B* **14**, 770-779
- Krasna, A. I. (1970) *Biopolymers* **9**, 1029-1038
- Kratohvil, J. P., Dezelic, G., Kerker, M. & Matijevic, E. (1962) *J. Polym. Sci.* **57**, 59-78
- Maestre, M. F. & Wang, J. C. (1971) *Biopolymers* **10**, 1021-1030
- Maron, S. H. & Lou, R. L. H. (1954) *J. Polym. Sci.* **14**, 29-36
- Paoletti, J. & Le Pecq, J.-B. (1971) *J. Mol. Biol.* **59**, 43-62
- Revet, B. M. J., Schmir, M. & Vinograd, J. (1971) *Nature New Biol. (London)* **229**, 10-13
- Rush, M. G., Kleinschmidt, A. K., Hellman, W. & Warner, R. C. (1967) *Proc. Nat. Acad. Sci. U.S.* **58**, 1676-1683
- Sinanoglu, O. & Abdalnur, S. (1964) *Photochem. Photobiol.* **3**, 333-342
- Sinsheimer, R. L. (1959) *J. Mol. Biol.* **1**, 43-53
- Smit, E. M. & Borst, P. (1971) *FEBS Lett.* **14**, 125-129
- Tanford, C. (1961) *Physical Chemistry of Macromolecules*, p. 151, Wiley, New York
- Upholt, W. F., Gray, H. B. & Vinograd, J. (1971) *J. Mol. Biol.* **62**, 21-38
- Wang, J. C. (1969) *J. Mol. Biol.* **42**, 25-39
- Waring, M. (1970) *J. Mol. Biol.* **54**, 247-279
- Zimm, B. H. (1948) *J. Chem. Phys.* **16**, 1093-1116