A Dynamic Light Scattering Study of a 2311 Base Pair DNA Restriction Fragment

Susan S. Sorlie and R. Pecora*

Department of Chemistry, Stanford University, Stanford, California 94305.
Received August 18, 1987; Revised Manuscript Received October 28, 1987

ABSTRACT: Dynamic light scattering time correlation functions from a solution of monodisperse, blunt-ended 2311 base pair (molecular weight 1.5 million daltons, contour length 7860 Å) DNA restriction fragment were measured at 20 °C, in 100 mM NaCl, 10 mM Tris·HCl, and 1 mM EDTA (pH 8) at DNA concentrations ranging from 90 to 450 μg/mL. The correlation functions were analyzed by using CONTIN, a constrained inverse Laplace transform program. At scattering angles of 16° and 22° the correlation functions are consistent with single exponential decays, representing the translational motion of the DNA. At angles of 57° and above, the correlation functions exhibit two or more relaxation modes, reflecting the internal dynamics in addition to the translational diffusion of the DNA. The distribution of decay times is generally consistent with theoretical simulations based on the Rouse–Zimm free-draining model for a flexible coil. The first moments of the distribution of relaxation times were found both from second-order cumulant fits and from the first moments of the relaxation time distribution function obtained from CONTIN analyses. Although the numerical values differed slightly, the first moments appeared to vary as the cube of the scattering vector length at the higher scattering angles studied. The mutual diffusion coefficient extracted from the low-angle data was found to increase with concentration and the K_d obtained from the slope is consistent with predictions for the hard-sphere model for particle interactions. It is also in accord with literature values on DNAs with molecular weights of the same order of magnitude. At 73°, the internal motion decay times are still separate from the diffusion decay time and the apparent diffusion coefficient obtained from the slowest decay time decreases with concentration. The K_d value of −1.2 obtained is comparable to the value of −1.8 calculated by Batchelor for the concentration dependence of the self-diffusion constant for a hard sphere. At the highest concentration, the relaxation time distributions given by CONTIN showed increased contributions from slower relaxation times showing the increased effect of intermolecular interactions. The general results are consistent with transient electric birefringence experiments performed previously on the same fragment.

I. Introduction

The theory of the dynamics of a Gaussian coil in dilute solution in the free-draining limit was developed by Rouse[1] and later put in more convenient mathematical form by Zimm.[2] Zimm, in addition to the free-draining case, also treated the case with hydrodynamic interactions.[2] The mathematical simplicity of the free draining Rouse–Zimm model enables it to be used to easily compute quantities directly measurable by a number of experimental techniques, including the time correlation function measured in polarized dynamic light scattering (DLS). The application of this model to polarized DLS was first given by Pecora[3] and later by Saito and Ito[4] and DuBois-Violette and de Gennes.[5] The polarized DLS time correlation function, as predicted by these theories,[3-5] contains contributions from the macromolecular translational diffusion and the long-range internal modes of motion of the polymer. The relative contributions of the translational motion and the various internal modes depend on the product of the scattering vector length, q, and the radius of gyration of the polymer, R_g. When $qR_g \ll 1$, translational diffusion is the dominant contribution, and the time correlation function is predicted to decay as a single exponential. As qR_g becomes larger, the various internal modes begin, in addition, to contribute significantly to the correlation function, giving rise to multiexponential time correlation functions. In the limit $qR_g \gg 1$, the contribution from the translation is small and the resulting time correlation function is best expressed by the functions given by DuBois-Violette and de Gennes.[5]

Although this free-draining model is mathematically convenient, it contains assumptions that appear to render it inapplicable to dynamics of a linear DNA fragment of the type considered in this article. The 2311 base pair DNA is not infinitely flexible; it should exhibit a restoring force to bending. It should also to some extent exhibit hydrodynamic interactions[2] and excluded volume effects.[6,7] There may even be effects in the long time range due to chain torsional motions.[8] Theories which incorporate some of these effects have been formulated, although none of these theories incorporate all of them in a consistent manner. Often within each theory a large number of parameters are floated or the equations obtained are mathematically complex. This makes it difficult to calculate the form of the distribution of decay times for the internal motions. Thus, for these presumably more realistic models, a direct comparison with the distribution of decay times obtained from DLS is very difficult.

The long-range internal motions of the polymer can also be studied by techniques other than DLS, such as transient electric birefringence (TEB).[9] TEB should be more sensitive to the internal motions (including rotation) than DLS since translational diffusion does not directly contribute to the TEB decay. In addition, qR_g must be >1 for internal/rotational motions to significantly contribute to the polarized DLS time correlation function. A major difficulty in the interpretation of TEB data from DNA solutions arises because the basic mechanism for interaction of the DNA with the orienting field is not known. Theories which assume that the basic mechanism of the orientation is a simple rotation of a statistical segment due to the torque exerted by the external electric pulse on the polarizability anisotropy of the segment (assumed to be constant) are not in accord with experimental results.[10] Thus, the Baur–Stockmayer theory for electric birefringence which utilizes this assumption about the orientation mechanism and the Rouse–Zimm dynamical model gives, as these authors emphasize, unphysical results.[11] Given the absence of even a simplified model, which might be expected to include the main features of the results, it is very difficult to quantitatively interpret the results of TEB experiments on DNA. A combination of DLS and TEB techniques should help overcome some of these difficulties. In this paper, we report the results of DLS experiments on a 2311 base pair (contour length 7860 Å, molecular weight 1.5 million) fragment of DNA. TEB has previously

0024-9297/88/2221-1437$01.50/0

been performed on this same fragment, at different ionic strengths and concentrations.[12]

We study DNA because it is an important macromolecule in itself and because, unlike most synthetic polymers, it is relatively easy to prepare in monodisperse form. In addition, a homologous series of different molecular weights may be prepared. The ability to produce monodisperse samples is extremely important in testing theories for long-range internal motions. For instance, Nash and King have shown that even a small degree of polydispersity can interfere with accurate internal mode decay time determinations from DLS experiments.[13] Thus, although there are complications from electrostatic effects, DNA restriction fragments can serve as useful model systems for semistiff polymer chains in solution, and experiments on them may be used to test dynamical theories of semistiff polymers and to obtain information on parameters appearing in such theories.

In DLS, TEB, and several other techniques, the measured quantity is a complex decay representing the resultant decay of different superimposed modes of motion. Inverting this measured decay to obtain the distribution of decay times requires data of extremely high signal to noise ratio because the mathematical process of inverting the data are ill-conditioned. The usual procedure has been to fit the measured correlation function to a "theoretical" form (one exponential, sum of exponentials, etc.) or to measure moments of it. The most popular technique is to measure the average decay time by the method of cumulants.[14] Thus, the majority of information obtained has been quantities derived from the first moment of the time correlation function, quantities variously called the "apparent diffusion coefficient" or "first cumulant" and given as a function of scattering angle.[15-17] In some cases, especially in studies of DNA, the "apparent diffusion coefficient" appears to be derived from the time constant of a single exponential fit to nonexponentially decaying correlation functions.[18-21] The numerical values of the cumulants or apparent diffusion coefficients derived from such studies are critically dependent on the details of the fit procedure.[15,20] The development of a sophisticated data analysis program, CONTIN,[22-25] has made it possible to reliably obtain more information about the distribution of relaxation times in the measured decays than is possible from an analysis of the moments.[12,26-28]

As described in section II below, CONTIN chooses the distribution of decay times, which best represents the data, using parsimony. This means that if there is not enough information in the data to resolve two closely spaced decay times, CONTIN will give a distribution function with only one peak. To identify a distribution function with two peaks, it is necessary to obtain better data. This may be done by, for instance, increasing the accumulation time for a correlation function or, more commonly, for DNA solutions, attempting to reduce the already small levels of dust in the solution. Small residual levels of dust are the major obstacle to obtaining good data from DNA solutions.

In studying internal motions, we are often working at CONTIN's resolution limit. CONTIN usually cannot separate decay times if they are less than about a factor of 2 apart with the signal to noise levels commonly obtained. Working with a distribution of peaks at or near this resolution limit creates a problem. When the distribution function is very complex, CONTIN does not always accurately represent the exact times or amplitudes of the peaks. For example, if two relaxation times are very close together, they will smooth even closer together. If they are too close, they will smooth into one peak. If the times are a little farther apart than the resolution limit, they tend to smooth even farther apart. Thus it is difficult to compare the distribution function obtained from CONTIN with the decay times obtained from a theory (for instance, the Rouse–Zimm model). One way of overcoming this difficulty is to simulate a correlation function from the theory by adding a typical amount of noise to the theoretical function and then analyzing it using CONTIN to obtain a theoretical "simulated" distribution function. If the resulting distribution function is similar to that obtained from experiment, then the theory is deemed to be consistent with experiment. (Many other theories may of course also be consistent with the results.)

In general, one can be confident that the distribution of decay times obtained by using CONTIN represents the decays in the data with a resolution not greater than what would be expected given the uncertainty of the data. A data analysis program that gives one broad distribution when two narrow distributions are actually present in a decay time distribution is preferable to a program that obtains a solution with two peaks when there is not enough information contained in the data to determine the existence of a second peak.

A distribution function for the decay times of the internal modes of a polymer is obtained from experimental DLS data by using CONTIN. This function may be compared with theoretical "simulated" distribution functions obtained from time correlation functions predicted by a particular model and analyzed by using CONTIN. These distribution functions are usually easier to compare with each other than the respective time correlation functions as time correlation functions that appear to be almost identical often show quite different distributions of relaxation times. In addition to facilitating comparison of experiment and theory, it is usually easier to detect and interpret changes in the distribution function when scattering angle, concentration, or other quantities are varied than to interpret changes in the original time correlation functions.

Section II contains a description of the methods used for preparing samples, a brief description of the DLS apparatus, and a description of the data analysis methods employed. In section III, the application of the Rouse–Zimm free draining model to DLS is discussed. Also a brief description of the method of cumulants is given. Section IV gives first the relaxation time distributions obtained from CONTIN for the 2311 bp fragment at several scattering angles. A comparison of the experimental results with the relaxation time distributions obtained from CONTIN analyses of the free draining Rouse–Zimm model is made. Next the first moments of the relaxation time distribution functions obtained from both CONTIN analyses and cumulant fits is given. The results are compared with those obtained by other workers, and the difficulties and limitations of both types of analysis are discussed. Then a comparison of the DLS results with TEB results on the same DNA fragment is given. Finally, concentration effects on the DLS relaxation time distributions and on the apparent DNA diffusion coefficients are described.

II. Methods

Sample Preparation. The plasmid used here, pLH2311, was constructed from pRI25 in our laboratory.[29] The plasmid was used to transform *E. coli* strain HB101 from which it was purified by using a slightly modified method[29] of Marko, Chipperfield, and Birnboim.[30] The plasmid was then digested with PvuII from New England Biolabs in MSB (50 mM NaCl, 10 mM Tris·Cl (pH 7.5), 10 mM $MgCl_2$ and 1 mM dithiothreitol) for approximately

2 h. This was followed by a phenol/chloroform extraction and three ether extractions. The plasmid solution was ethanol precipitated twice and resuspended in 100 mM NaCl and 10 mM Tris·HCl (pH 8.0) to a final concentration ranging from 90 to 450 μg/mL. The solution was filtered twice, first using a Schleicher and Schuell SS009 centrifuge filtration unit with a OE66 0.2-μm pore size filter. In the second filtration the solution was filtered through a Gelman ACRO LC13 0.2-μm pore size filter into a clean square cuvette. The solution was examined for dust through a 10× microscope during illumination with 0.2-W of laser light. When clean, the sample was allowed to equilibrate for 20 min at 20 °C before making measurements. All measurements were performed at 20 °C.

Experimental Apparatus. The light scattering apparatus has been described elsewhere[31] and only a brief discussion will be given here. The light source was a 4-W Spectra Physics 165 argon ion laser operating at 488 nm. The laser power utilized in the experiments was between 0.2 and 1.5 W. The homodyne autocorrelation function was measured by using a Brookhaven BI2020 correlator at scattering angles of 123, 106, 90, 73, 57, 22, and 16°. In general, 64 data channels plus 8 delay channels were used to determine the low-angle correlation functions and 128 data channels and the 8 delay channels were used at angles of 90° or greater. Correlation functions were determined with as many as 512 data points by using the multiplex option on the Brookhaven correlator. However, difficulties in splicing the data together and the long-time required to obtain a sufficiently high signal to noise, usually negated the advantage of the increased number of channels. The first data point from each data set was not used. Typical signal to RMS noise ratios were between 1000 and 2000. The relative difference between the calculated base line and the delay channels from the correlator was generally between 0.005 and 0.05%. At the lowest two angles the difference was generally higher, up to (but usually less than) 0.2%. At these two angles small amounts of dust or stray light were not as important as at the higher angles where they could adversely affect the resolution of the multiple relaxation modes. The data were electronically transferred from the Brookhaven correlator to a Universe 68 microcomputer manufactured by Charles River Data systems. The square root of the data minus the base line (what is often called $g^1(t)$) was then analyzed by using versions of the Fortran programs CONTIN[22-25] and DISCRETE.[32,33] These programs were modified from the original versions written by Provencher in order to efficiently run on the Universe 68 microcomputer. The second-order cumulant fit was made by using the standard software package with the Brookhaven BI2020 correlator.

Data Analysis. CONTIN, using a constrained regularization method, performs a nonlinear least-squares fit to numerically solve the inverse Laplace transform (ILT) of the correlation function. The determination of the ILT of the correlation function is an ill-posed problem in that there are an unbounded number of solutions all of which fit the data to the precision of the data. Highly oscillatory solution often fit the data better than smoother solutions because they in fact fit the noise. To overcome this, CONTIN incorporates parsimony and statistical prior knowledge. This means that the smoothest solution, usually the one with the fewest peaks, which is still consistent with the data is used. CONTIN uses an F-test and confidence region to determine the correct amount of smoothing. Only solutions with nonnegative components are considered. This last constraint is not a physical necessity. It, however, greatly reduces the number of possible solutions. Recently, Aragón[34] was shown the depolarized DLS time correlation function, from the Aragón–Pecora[35] wormlike chain model, exhibits exponentials with negative amplitudes. No calculations have been performed as yet on the polarized DLS time correlation function using this model in its general form.

CONTIN does not require the number of exponentials (or peaks in the distribution of decay times) to be input in advance and, in general, does not overestimate the number of relaxation processes occurring in the solution. For example, if there is a small contribution from dust, afterpulsing, or an impurity, allowing the data to be fit to an arbitrary distribution of decay times is better than constraining the fit to a set number of exponentials. Also if the solution is very polydisperse, programs that do not penalize highly oscillatory solutions will often use two or more peaks to fit one broad peak.

Occasionally a small peak at the fast end of the window is observed. CONTIN seems to use this small peak to obtain a better fit. However, unlike the unphysical peaks given by some programs that do not incorporate parsimony, it is usually less than one percent of the total scattering intensity. In simulated correlation functions without added Gaussian noise it has not been observed by us. It tends to occur in solutions where noise has been added with little or no smoothing and tends to go away as the smoothing increases. Thus, observation of how the amplitude of this peak changes with smoothing aids in determining whether or not it is real. In the few cases when it has actually been observed in simulated correlation functions, it did not seem to appreciably alter the position of the other peaks.

DISCRETE, which was also written by Provencher, fits the data to a discrete sum of exponentials. However, since DISCRETE does not use parsimony to determine the best fit to the data, it is possible to obtain significant (larger than a few percent) peaks that do not correspond to decay processes which occur in a solution of macromolecules. Also DISCRETE fits a finite sum of exponentials (up to five). If the data contain a wide distribution of relaxation processes, one would expect CONTIN to better fit the data. In our experience studying a wide variety of macromolecular systems, when theoretical predictions indicate that only a small number of relaxation processes will contribute to the DLS time correlation function and the samples were clean, CONTIN and DISCRETE give virtually indistinguishable results. This was the case in our DNA experiments at all angles less than 90 degrees. For higher angles a wide range of relaxation times are predicted to be present (see sections III and IV below).

For most of the CONTIN analyses the amount of smoothing was within a 90% confidence region although less smoothed than that recommended by Provencher. It has been shown in our laboratory by Bott[36] that using threshold statistics to determine the amount of smoothing is comparable to Provencher's analysis when the amount of smoothing was closer to the undersmoothed end of the confidence region. (A probability one to reject of 0.1–0.2.) Bott also pointed out that threshold statistics are based on more accurate assumptions than the F-test and thus would be expected to be a better method for determining the amount of smoothing necessary. Instead of using threshold statistics, we have used the F-test with a smaller amount of smoothing (probability one to reject of 0.1–0.2).

Other precautions were taken to ensure that each major peak corresponded to actual motions in the polymer and not to some artifact in the experiment or data analysis. Each measurement at a particular angle was performed

several times on a given sample and was repeated on at least one other sample prepared at a different time. Similar results were obtained for each sample. Also, measurements performed at different scattering angles were qualitatively consistent with each other and with theoretical expectations.

Sample Purity. To determine if the samples contained impurities that would interfere with the observation of internal motions, the dynamic light scattering correlation function was measured at 16 or 22 degrees. It was then analyzed by using both CONTIN and DISCRETE. If either analysis gave more than one peak in a hydrodynamic radius, R_h (which is related to the decay time by eq 9), range between 20 and 1800 Å, it was considered unsuitable. Only two data sets had a small peak above 1800 Å that was presumed to be from dust. The "size" polydispersity is reflected in the moments of the CONTIN distribution function. It is equal to $\langle R_h \rangle \langle 1/R_h \rangle$. The value of the size polydispersity was typically 1.008 at low angles. We found that it was very important to painstakingly check the purity of our fragments by observing the DLS at low angles. We often found that samples which showed no detectable traces of impurities on electrophoresis gels (agarose and poly(acrylamide)), and whose relative absorbance at 260 vs 280 nm was not unusually high, still sometimes exhibited more than one peak of considerable amplitude at low angles. We found this to be the case not only when we used our usual method of preparation but also when we used more traditional methods such as cesium chloride density gradient ultracentrifugation.[37] Unfortunately, the DLS decay times observed for these impurities were close enough to the decay times for internal motions observed at higher angles that these impurities were not tolerable. All data taken with these samples was discarded.

After data collection, the DNA concentration was measured by UV absorption at 260-nm wavelength. The ratio of the absorbances at 260 and 280 nm, used to check for possible contamination with protein or phenol, was usually approximately 0.54. The DNA fragment was checked for degradation by agarose gel electrophoresis. At 20 °C, the DNA showed no detectable degradation over a period of 1–2 days.

III. Theory

Rouse–Zimm Simulated Correlation Functions. The Rouse–Zimm[2] model in the free-draining limit describes the motion of a completely flexible Gaussian polymer chain in a continuum solvent. The chain consists of n beads connected by $n - 1$ bonds. The forces allowed are an "entropic" restoring force on each bead, linear in the separation distance between nearest-neighbor beads, a frictional force on each bead proportional to its velocity, and a stochastic Brownian force exerted on each bead by the surrounding medium. Hydrodynamic interactions between different beads are considered negligible in this version of the model. Pecora[3] has calculated the dynamic form factor $S(x,t)$ for this model in the limit of large n as a function of the square of the normalized scattering vector, x, where

$$x = q^2 R_g^2 \tag{1}$$

Here q is the scattering vector length and R_g is the root-mean-square radius of gyration of the coil. The dynamic form factor may be written in the form

$$S(x,t) = e^{-q^2 D t} \sum_{n=0}^{\infty} S_n(x,t) \tag{2}$$

where D is the diffusion coefficient. The $S_n(x,t)$ $\exp[-q^2 D t]$ are called the "DLS relaxation modes" and the $S_n(x,t)$ are given by

$n = 0$

$$S_0(x,t) = \frac{\pi}{x} e^{-(1/6)n} \left[\operatorname{erf}\left(\frac{x^{1/2}}{2} \right) \right]^2$$

$n \geq 1$

$$S_n(x,t) = \frac{4^n}{n!}\left(\frac{1}{2}\right)^{2n-2} \frac{x^n}{\pi^{2n}} e^{-(2/3)x} \sum_{\substack{k_1=1 \\ k_2=1 \\ \vdots \\ k_n=1}}^{\infty} \left[\frac{e^{-t\left(\frac{1}{\tau_{k_1}} + \frac{1}{\tau_{k_2}} + \ldots + \frac{1}{\tau_{k_n}} \right)}}{k_1 k_2 \ldots k_n} \times \sum_{\substack{\text{all } 2^n \\ \text{combinatns} \\ \text{of } \pm}} [I(k_1 \pm k_2 \pm \ldots k_N)]^2 \right] \tag{3}$$

In eq 3, τ_k is the normal mode decay time and $I(k)$ is related to the real part of the complex error function

$I(k) =$

$$\pm\left(\frac{\pi}{x}\right)^{1/2} \exp\left(\frac{x}{4} - \frac{k^2\pi^2}{4x}\right)\left[\operatorname{Re}\, \operatorname{erf}\left(\frac{x^{1/2}}{2} + \frac{ik\pi}{2\sqrt{x}} \right) \right] \tag{4}$$

Note that in eq 2, the translational diffusion coefficient contributes directly in the first exponential factor and thus appears in each term in the series that constitutes the polarized DLS time correlation function. The first term in the series ($n = 0$) contains no contribution from the τ_k, and hence, the corresponding contribution to the time correlation function decays solely by translational diffusion. This contribution is called the "pure translational mode" below.

The τ_k in eq 3 are twice Zimm's normal mode relaxation times.[2] They are given by

$$\tau_k = 2R_g^2/\pi^2 D k^2 \quad k = 1, \ldots, N \tag{5}$$

It should be emphasized that these normal mode relaxation times are different from what we call the DLS relaxation times. The DLS relaxation times are the decay times of the individual exponentially decaying contributions to the DLS time correlation function

$$S_n(x,t)e^{-q^2 D t} \tag{6}$$

According to the Rouse–Zimm theory these DLS relaxation modes exhibit reciprocal relaxation times that are linear combinations of the reciprocals of both the normal mode relaxation times and the decay time from translational motion.

For this reason the time constants of the DLS relaxation modes are not as widely separated from each other as are those of the normal modes. This makes the DLS relaxation modes more difficult to resolve than if the decay constants of the DLS relaxation modes were simply those of the normal modes. The first normal mode decay time, given in eq 5, is a factor of 4 times larger than the second and a factor of 9 times larger than the third. Because the decay constants of the DLS relaxation modes are linear combinations of those from the normal modes and that from translational motion, the first internal mode decay time is less than a factor of 2 larger than the second and less than a factor of 3 larger than the third. The pure translational decay time is less than a factor of 2 larger than the first DLS decay time at low angle and is slightly

greater than a factor of 2 larger at the highest angle where measurements were performed.

Pecora calculated $S(x,t)$ for n = 0–2 for values of x ranging from 1 to 7 and included only some of the first few normal modes.[3] Later Perico, Piaggio and Cuniberti[38] calculated $S(x,t)$ for n = 0–6 and x = 1–10 and for higher orders in k. We here present an extension of this theoretical work to include higher terms in both n and k for particular values of x relevant to our studies of the 2311 bp DNA fragment.

Cumulants. The logarithm of the normalized polarized DLS time correlation function can be expanded in terms of cumulants

$$\ln S(x,t) = -K_1 t + K_2\frac{t^2}{2} - K_3\frac{t^3}{6} + \ldots (-1)^N K_N\frac{t^N}{N!} \quad (7)$$

where K_N is the Nth cumulant. Koppel[14] has shown that the first three cumulants are equal to the first three moments of the decay time distribution function. Higher order cumulants are related to the higher order moments, although some are combinations of moments. When the distribution function is composed of a single relaxation time, only the first cumulant is nonzero. As the number of relaxation times increase, the number of nonzero terms in the expansion necessary to describe the distribution function increases. The first cumulant from a multiexponential decay can be obtained by reducing the number of data points used (short times) and fitting the data to a polynomial in t. In the q region where the correlation function is multiexponential, the parameters obtained are, however, very sensitive to the fitting function (number of terms in the polynomial) and the number of experimental points included.[15,20] For our DNA data, we have compared the first cumulant obtained from a simple second-order cumulant fit using all the experimental data points except the first one with the first moment of the relaxation time distribution function. The first moment of the time distribution function can be easily calculated from the first negative moment from the size distribution function obtained from CONTIN. In principle, the first cumulant and the first moment from the second-order fit and the CONTIN analysis, respectively, should be equal at all values of q, but in practice, due to the limited expansion used for the cumulant method, they are equal at low q and deviate from one another as q is increased. We have observed experimentally how these quantities change with q and polymer concentration.

IV. Results and Discussion

Distribution Functions. In order to apply the Rouse–Zimm calculations to our 2311 base pair DNA restriction fragment, we need the diffusion coefficient, D, the radius of gyration, R_g, and $I(k)$ in eq 4.

The diffusion coefficient is related to the decay constant, γ, or decay time, τ, obtained at low angle by

$$D = \frac{\gamma}{q^2} = \frac{1}{\tau q^2} \quad (8)$$

The average of all the data taken at the two lowest angles at concentrations ranging from 90 to 260 μg/mL is D_0 = 4.56×10^{-8} cm^2/s ± 2.8%. This is about 3% faster than the value obtained by extrapolating to zero concentration (see Figure 6). We used this average value in the calculations and not the extrapolated zero concentration value for several reasons. First the internal mode data obtained at the higher angles could not be easily extrapolated to zero concentration and probably does include some effects from neighboring molecules. Second, the difference between the two numbers was within the scatter in the data points (see Figure 6 below). Finally the effect on the positions of the DLS relaxation modes was minimal and the difference would not appreciably affect the amplitudes of those modes.

The radius of gyration for our DNA was obtained from total intensity light scattering measurements performed on the same DNA restriction fragment at the same salt concentration by Lewis,[39] who obtained a value of R_g = 1044 Å.

The integral, $I(k)$, in eq 4 was solved numerically by using Simpson's rule. The number of divisions was large enough so that the solution was not altered appreciably by increasing the number of divisions.

At the values of x corresponding to the scattering angles where DLS measurements were performed, the decay times for the DLS relaxation modes and their corresponding amplitudes can be calculated as was shown in the previous section. At low values of x most of the contribution to the correlation function comes from the n = 0 term in eq 2 and corresponds to pure diffusion. As x is increased, many other modes contribute small amounts to the correlation function. Thus it becomes increasingly more cumbersome to calculate all the modes which contribute to the time correlation function. For our calculations, at x = 10, or an angle of 123°, 13.7% of the total scattering amplitude was not included in the calculation. At x = 6 or 57° this number drops to 0.4%. However, most of the DLS internal relaxation modes that are not included decay very rapidly, so that they would not be observed on the time scale of the experiment. Those that are contributing on the time scale of the experiment would be of small enough amplitude that they are only a small percentage of the overall decay. Thus the overall decay calculated closely represents the decay expected from the theory with all the internal motions included.

The calculated decay times and their amplitudes were used to simulate a correlation function for all scattering angles where measurements were performed. The simulated correlation function was then analyzed by using both CONTIN and DISCRETE. Even with high-precision calculated data, CONTIN was unable to resolve all of the decay times predicted from theory. This was expected since many of them were only a few microseconds apart and even "high-precision" data are not precise enough to resolve modes this close together. In addition, only a finite grid of points was calculated in the ILT. DISCRETE, on the other hand, could only resolve two or three relaxation times even at high values of x where CONTIN could resolve up to six. The simulated correlation function from CONTIN is the light, solid line shown in Figures 1–4. The dashed line in Figure 3 and 4 represents the same correlation function except that Gaussian noise has been added to the correlation function in order to show how CONTIN smoothes the data.

Shown in Figure 1 is a typical CONTIN analysis of DLS data from a solution of 2311 base pair DNA at a scattering angle of 22°. Similar data was obtained at 16.4°. The CONTIN analysis gives the fraction of the scattered intensity corresponding to an "apparent" hydrodynamic radius. The "apparent" hydrodynamic radius is related to the decay time through the relation

$$R_h = \frac{kTq^2\tau}{6\pi\eta} = 2.52\tau \sin^2(\theta/2) \quad (9)$$

where k is Boltzmann's constant, T the absolute temperature, and η the solvent viscosity. The second equality in eq 9 substitutes values of these parameters and the light

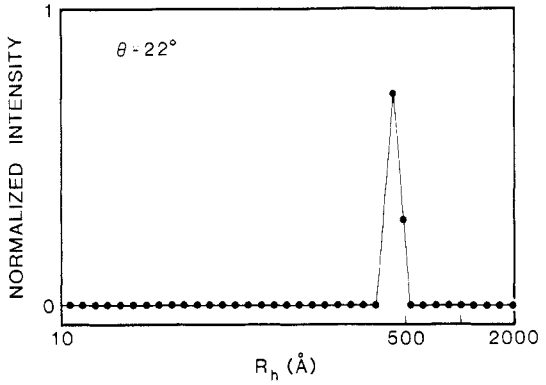

Figure 1. CONTIN analysis of the experimental correlation function at a scattering angle of 22°. The one peak seen here is from translational motion. Its position in hydrodynamic radius is 466 Å which corresponds to a diffusion coefficient of 4.60×10^{-8} cm^2/s.

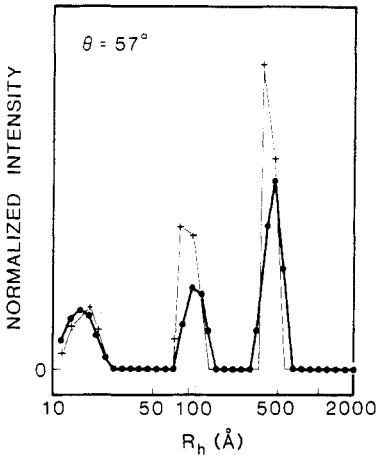

Figure 2. CONTIN analysis of the correlation function at 57°. The (●) represents the experimental distribution function. The (+) represents the distribution function from a simulated correlation function of the Rouse–Zimm model for the Gaussian coil in the free draining limit. The peak positions and their percent contribution to the correlation function are given in Table I.

wavelength appropriate to our experimental conditions, so that easy conversion between R_h (Å) and τ (μs) may be made. By using an "apparent" hydrodynamic radius for each relaxation process observed, we are in effect scaling these decay times by q^2 and thus removing the q dependence of the diffusion contribution and making the DLS internal relaxation times appear to be dependent on q^2.

Figure 1 shows only one peak at 22° indicating that only translational motion contributes to the correlation function at this angle. This is consistent with eq 3 which shows that the DLS internal relaxation modes ($n \geq 1$) contribute approximately 0.5% of the total scattering intensity. The average of the values obtained from both CONTIN and DISCRETE for the hydrodynamic radius is 470 ± 14 Å. This is about 20% lower than that obtained from sedimentation.[39,40] A difference between DLS and sedimentation results has also been observed by Thomas, Allison, and Schurr[19] for λ and ϕ 29 DNA. The diffusion coefficient reported by Lewis and Pecora[39,42] for the same restriction fragment as studied here is somewhat smaller than that found in our more extensive studies.

Figure 2 gives the CONTIN analysis of data taken at 57°. The dotted curve is the experimental distribution function, while the curve with crosses represents the theoretical distribution. At this scattering angle there are two peaks in addition to translation. The first peak, at 474 Å, is still assumed to come only from translational motion. Now,

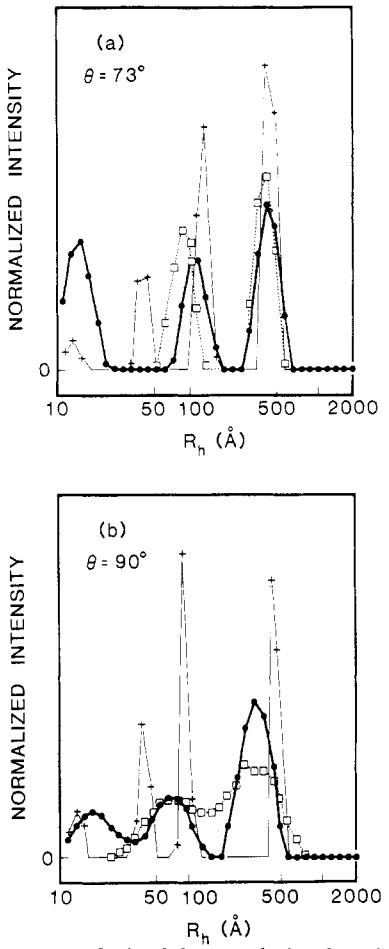

Figure 3. CONTIN analysis of the correlation function at 73° and 90°. The (●) represents the experimental data. The (+) represents the distribution function for the simulated correlation function from the Rouse–Zimm model for the Gaussian coil in the free draining limit. The (□) represents the distribution function from a simulated correlation function with approximately 0.1% Gaussian noise added to the function. The peak positions and their amplitudes are given in Table I. It should be noted that the abscissa is logarithmic; thus, the peaks at low hydrodynamic radius have less intensity than might be apparent at first glance.

however, it contributes only 87.6% of the total scattering amplitude. The next two peaks are at 113 and 15 Å and contribute 11% and 1.4%, respectively, to the total amplitude. The larger of these two peaks is presumed to be from a combination of the internal motions of the coil and translational diffusion. The smaller one could be from internal motions (with the usual contribution from translational diffusion), afterpulsing, some artifact of the data analysis, or even a combination of these. Since it is only a few percent of the total scattering amplitude, it is not considered significant. The curve with crosses that represents the theoretical distribution function is in good agreement with the experimental curve, even without adding noise to the theoretical correlation function. The largest peak is at 462 Å and has an amplitude of 87%. The next peak is at 95 Å and contributes 11.6%. Finally the smallest peak is at 15 Å and contributes 1.3% to the total amplitude.

Figure 3a shows the CONTIN analysis of data measured at 73°. The squares in this figure represent the theoretical distribution with a small amount of Gaussian noise added (less than 0.1%). This curve shows the effect of smoothing on the analysis. The experimental analysis gives three peaks, at 435, 116, and 14 Å. Their amplitudes are 82%, 16%, and 2.2%, respectively. The theoretical analysis has

four peaks, at 461 (75.6%), 137 (21%), 44 (3.2%), and 12 Å (0.2%). When noise is added to the theoretical analysis, it smoothes to two peaks at 410 (84%) and 92.3 Å (16%). It is interesting to note that the experimental solution seems to lie in between the theoretical solution with and without noise added. It is possible that slightly better agreement with experiment could be obtained by decreasing the amount of noise added to the simulated correlation function. Note also that according to the theoretical prediction the position of the second peak for the simulation with added Gaussian noise is not simply related to the first normal mode decay time. This peak represents the smoothing together of several closely spaced exponentials in eq 2.

Figure 3b gives the CONTIN analysis at 90°. In the experimental analysis the largest peak, at 318 Å, has moved to a position that is much smaller than what would be expected for the pure translation term. The theory shows that, as the scattering angle increases, the contribution to the relative amplitude from the internal motions increases. The analysis without added noise still gives a peak at 464 Å (84.4%) and with other peaks at 92 (10.8%), 41 (4%), and 12 Å (0.4%). When this is smoothed by the addition of noise, the first peak moves to a smaller hydrodynamic radius, 343 Å (84%). The position of this first peak is very similar to the position of the first peak from the experiment at 318 Å (88%). The next peak of the theoretical analysis with noise is at 79 Å (16%). The experimental analysis gives 69 (10%) and 18 Å (2%).

Figure 4a shows the analysis of data taken at 106°. The experimental analysis gives peaks at 296 (88.1%) and 58 Å (9.6%) and a small dust peak at 2000 Å (2.3%). The theoretical analysis gives peaks at 451 (54.1%), 192 (33%), 84 (10.1%), 42 (2.5%), and 17.5 Å (0.3%). Adding noise, this smooths to two peaks at 267 (92%) and 54.4 Å (8%). The agreement between the experimental and the theoretical function is quite good. If the amplitudes of each of the peaks in the noiseless data were changed only slightly, one might expect that the way the data would be smoothed to change, and thus it might not agree with the experimental analysis as well. For example, if the second largest peak were a bit smaller, one might expect that it could be resolved from the largest peak even when the data was smoothed.

Finally Figure 4b shows the CONTIN analysis of the data taken at 123 degrees. Theory tells us that there are many DLS modes contributing to the correlation function at this angle. The theoretical analysis without noise gives six peaks at 442 (47.6%), 203 (34.8%), 95 (13%), 56 (3.5%), 29 (0.9%), and 16 Å (0.2%). When noise is added, this smooths to three peaks at 323 (74%), 97 (25%), and 26.9 Å (1%). The experimental analysis gives peaks at 268 (91%), 69 (5.4%), and 16 Å (3.6%). The agreement between the experimental and theoretical analysis is not as good. This is probably because at very high angle the sample times are very short and thus the number of photons per sample time and the resulting signal-to-noise ratio are relatively low. In order to obtain signal-to-noise ratios in the correlation function comparable to those obtained at lower angles, it is necessary to run for very long periods of time, which is not always practical. It is interesting to note that, although the CONTIN analysis of the theoretical correlation function does not agree very well with experiment at this angle, the DISCRETE analysis agrees much better. The theoretical analysis using DISCRETE with added noise gives three peaks at 345 (71%), 108 (26%), and 32.5 Å (3%). The experimental DISCRETE analysis also gives three peaks at 338 (76%), 105 (20%), and 17 Å (4%).

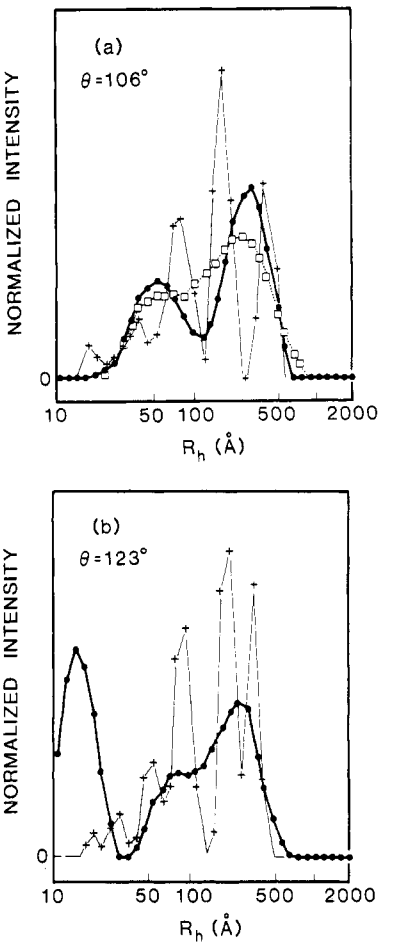

Figure 4. CONTIN analysis of the correlation function at 106° and 123°. The (●) represents the experimental data; the (□) and the (+) represent the simulation from the model with and without added Gaussian noise, respectively. At high angles the distribution function is complex, and there is not enough information in the data to resolve all the peaks. When Gaussian noise is added to the simulated distribution function, one can see how CONTIN smooths the data when there is not enough information present to resolve everything. The simulated distribution function with noise at 123° is not shown.

We conclude that, at this angle and concentration, there is not enough information in the data to resolve anything more than a broad single peak. However, since DISCRETE does not use the principle of parsimony, it can find a solution that seems to be correct even though there is not enough information in the data to be certain that it is correct.

It should be pointed out that at this high scattering angle q^{-1} is 332 Å. This is less than a persistence length in the polymer coil. The Rouse–Zimm theory in the form used in our simulations assumes that the coil, composed of n beads connected by $n - 1$ bonds with harmonic spring constants, is infinitely flexible; i.e., the number of beads goes to infinity and the distance between adjacent beads remains Gaussian. Thus, it would seem highly improbable that probing distances less than a persistence length in the polymer would give results that could be described by this model. For this reason, it would be very interesting to compare these results with simulated correlation functions that include stiffness. At present, however, no theories exist which incorporate stiffness which are simple enough to allow for easy comparison. It would also be interesting to compare the DLS results with those of a Langevin dynamics simulation of the DLS time correlation function for this fragment. Langevin dynamics simulations for fluorescence depolarization, depolarized zero angle DLS,

Table I
The Peak Positions and Their Amplitudes from the Distribution Function: Observed and Predicted for the Gaussian Coil Model in the Free Draining Limit

				predicted			
		obsd		without added noise		with added noise	
angle, deg	peak no.	size, Å	% amp.	size, Å	% amp.	size, Å	% amp.
57	1	474	88	462	87		
	2	113	11	95	11		
	3	15	1.4	15	1.3		
73	1	435	82	461	76	410	84
	2	116	16	137	21	92	16
	3	14	2.2	44	3.2		
	4			12	0.2		
90	1	318	88	464	84	343	84
	2	69	10	92	11	79	16
	3	18	2	41	4		
	4			12	0.4		
106	1	296	88	451	54	267	92
	2	58	10	192	33	54	8
	3			84	10		
	4			42	2.5		
	5			18	0.3		
123	1	268	91	442	48	338	76
	2	69	5.4	203	35	105	20
	3	16	3.6	95	13	17	4
	4			56	3.5		
	5			29	0.9		
	6			16	0.2		

and TEB have been performed for models of shorter DNA fragments than that studied here.[43,44]

The results from the CONTIN analysis for the experimental data as well as for the simulated data with and without Gaussian noise, as functions of scattering angle, are summarized in Table I.

Cumulants. The first cumulant and the first moment as a function of the product of the scattering vector and the low-angle hydrodynamic radius, qR_h, obtained from the two types of data fits is shown in parts a and b of Figure 5. Each fit is performed at a number of different concentrations. Figure 5a shows the first cumulant obtained from the second-order cumulant fit to all the data. The open circles represent measurements performed at the highest concentration (450 μg/mL) while the closed circles represent measurements performed at lower concentrations, (90–270 μg/mL). Figure 5b shows the first moment of the decay time distribution function. This is proportional to the average of the inverse of the size distribution function obtained from CONTIN. The open and closed circles are the same as given above. The reduced first cumulant and first moment from the two fits should be equivalent. At present consider only the open circles since this represents work done at low concentrations (90–270 μg/mL). The closed circles are discussed under concentration effects below. The ordinate is the log of the first cumulant divided by q^2D_0, where D_0 is the diffusion coefficient obtained at low angle. The abscissa is log (qR_h), where R_h is the hydrodynamic radius obtained at low angle.

Initially the slope is approximately zero. Between angles of 40° and 60°, or, on the abscissa scale, –0.1 and –0.2, the slope deviates from zero and begins to increase. Assuming that the data increases linearly, we can perform a linear regression to obtain the slope. For the first cumulant, in Figure 5a, all the data beyond log (qR_h) = –0.15 were included, except for the data at the highest concentration. A slope of 0.85 was obtained. For the first moment, in Figure 5b, all the data beyond log (qR_h) = –0.3 was included. Again the higher concentration data points were ignored. A slope of 0.99 was obtained. Thus, the first cumulant varies with the 2.85 and third power of q for the second-order and first moment cumulant fits, respectively.

The first cumulant has been shown to scale as q^4 for the Rouse–Zimm coil in the free draining limit and as q^3 in the non-free draining limit.[5] In both cases, this assumes that the chain is long enough that the contribution from diffusion is negligible or that qR_g is much greater than one and furthermore that qa, where a is the Kuhn statistical length, must be smaller than one. The statistical length may be estimated for the DNA fragment by assuming that the relation between R_g, a, and the contour length, L, is that of a Gaussian coil

$$a = 6R_g^2/L \tag{10}$$

A value of 832 Å is obtained. In our experiments, qR_g ranges from 0.68 to 3.1 and qa ranges from 0.55 to 2.5. Thus it was unexpected that both the second-order cumulant fit and the moments analysis would give a q dependence in the range predicted by the models.

The q dependence obtained from both the cumulant and first moment analyses suggest that the motions of the polymer could be better described by the model in the non-free-draining limit. The CONTIN distribution analysis results from the previous section suggest that the first cumulant at high q should scale, as predicted in the free-draining limit, with the fourth power of q.[5] Transient electric birefringence measurements performed on this fragment (although under slightly different conditions, see below) also suggest that the first cumulant should be described better by the model in the free-draining limit than the non-free-draining limit. The hydrodynamic radius predicted in the non-free-draining limit, calculated by using the radius of gyration[2,46]

$$R_h = 0.663R_g \tag{11}$$

is R_h = 692 Å. We obtained a value 470 Å from the experiments. From the radius of gyration and the contour length (or the statistical segment length), we estimate the number of statistical segments, $n \cong 10$. Finally, using the theoretical relation between R_h and n and the frictional coefficient in the free-draining limit, we may estimate the hydrodynamic interaction parameter,[2] h

$$h = 6R_h/a\sqrt{12\pi n} \tag{12}$$

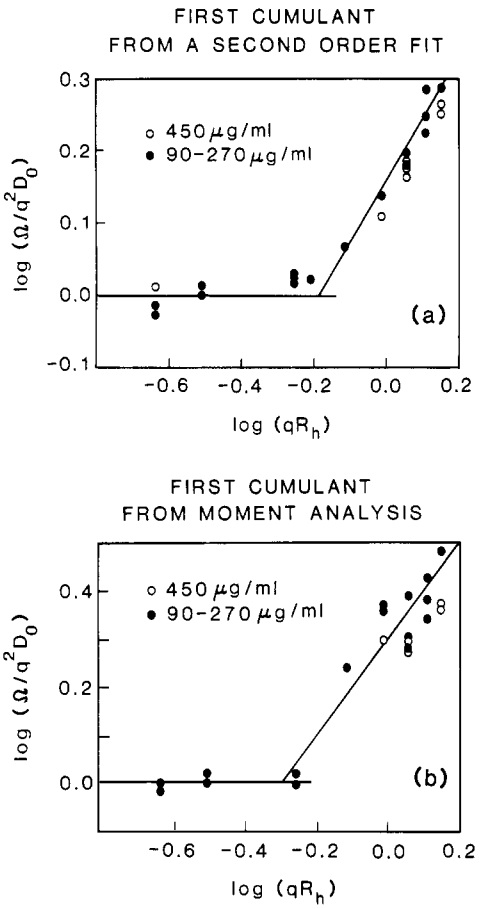

Figure 5. log of the reduced first cumulant as a function of the log of the reduced scattering vector length. The (○) represents work performed at low concentrations (90–270 μg/mL), while the (●) represents work performed at the highest concentration (450 μg/mL). See text. The reduced first cumulant, in Figure 5a, was obtained from a second-order cumulants fit. The reduced first cumulant, in Figure 5b, was obtained from the first moment of the distribution function from CONTIN. At high q, the slope is 2.99 for the CONTIN analysis and 2.85 for the second-order cumulants analysis.

A value of 0.17 is obtained, suggesting that the free-draining limit would be more appropriate for describing the motions of a 2311 bp DNA fragment. Neither the cumulant nor moment data analysis techniques appear to be consistent with this.

There are a number of possible explanations for this apparent discrepancy. First it is likely that the slopes from both the first cumulant and the first moment analysis are too low since qR_g is not very large and we could be in a transition region where the slope is changing from zero to four. Second, it is possible that these two data analysis techniques do not give accurate values for the slope. A possible source of difficulty with the CONTIN moment analysis is that small peaks at the fast end of the window from afterpulsing, artifacts, etc., which do not have much of an effect on the position of the other peaks, have a large effect on the first moment. This is because the first cumulant is related to the average of the inverse radius. Thus a peak with a small radius will have a large effect on the first moment.

Obtaining accurate first cumulants from the second-order cumulant analysis also, as mentioned above, presents major difficulties. This is because it is likely that too few terms were included in the time expansion of the correlation function (eq 7). A general discussion of the severe problems in doing a higher order cumulant fit is given by Akcasu et al.[15]

Han and Akcasu showed that for solutions of polystyrene of very high molecular weight, the first cumulant from a complicated higher order cumulant–shape function analysis scales with the third power of q.[41] Soda and Wada also found a q^3 power dependence for linear ColE1 DNA.[20] This DNA is approximately three times the size of our fragment. However, the transition region between the first cumulant depending on q^2 and q^3 occurs around $qR_h = 0.7$ for both DNA fragments. This corresponds to −0.2 on the abscissa in Figure 5 and a $q^2 = 0.7 \times 10^{14}$ m^{-2} in Figure 8 of Soda and Wada.[20] Adam and Delsanti obtained a 2.85 power dependence on q for polystyrene, MW = 8.4×10^6 and 24×10^6, in benzene.[47] Other authors have also seen this third-order power dependence on q.[16] These results are consistent with our results obtained for both the moment and first cumulant analysis. This suggests that the q dependence we obtained would not change as q is increased; i.e., we are not in a transition region. However, it is also possible that because these polymers have much longer contour lengths or are more flexible than our 2311 bp DNA fragment, hydrodynamic interactions are more important for them. In order to determine if the first cumulant scales with the third or fourth power of q, it would be necessary to use more esoteric data analysis techniques such as higher order cumulant–shape function fits and to increase the range of q used.

Comparison with Transient Electric Birefringence. Using a CONTIN analysis of the TEB decay curves, Lewis, Pecora, and Eden[12] measured the first few decay times for the internal motions of several linear DNA restriction fragments, including the 2311 bp fragment studied here. For the 2311 base pair fragment the first two decay modes measured were assumed to correspond to the first two normal modes for a flexible coil. The ratio of the slowest to the next slowest decay time was between three and four, values consistent with the Rouse–Zimm model with draining between the free-draining and non-free-draining limits. For the shorter fragments studied, the ratio of the first two times was larger than four as one would expect for chains whose dynamics were influenced by stiffness. The actual values of the decay times for the 2311 base pair fragment were much closer to those predicted in the free-draining limit than the non-free-draining limit.[42] This comparison of decay times, rather than ratios of decay times, appears to be a more sensitive test because the estimated values of the first normal mode decay times in the free-draining and non-free-draining limits for this fragment differ by a factor of 3.

There are several differences between DLS and TEB measurements that make a direct comparison of the results difficult. First, in TEB experiments, the molecules are perturbed by using an electric field pulse. To avoid excess heating of the solution, migration of the ions, etc., it is necessary to perform the experiments in low-salt solutions. For TEB, the experiment was performed at salt concentrations of less than 3 mM NACl. For DLS, the experiments were performed at 100 mM NaCl. The persistence length at high salt is possibly a factor of 2 smaller than at low salt concentrations.[7,21] Thus at high salt the plasmid is more flexible. Second, as discussed in section III above, DLS theory predicts that the reciprocal DLS relaxation times are each represented by a sum of the reciprocal relaxation times for the normal modes (as well as by translational diffusion),[2] so that the directly measured decay times are not the normal mode relaxation times of the theoretical models. Thus, even given the assumption that the TEB relaxation times are each one-half those given by eq 5, it would not be possible to directly compare

them with the DLS relaxation modes except in the simplest cases. An advantage of DLS over TEB in studying long-chain polymers is that it is a relatively noninvasive technique. No external fieelds except the relatively weak fields in the laser beam need be applied. As a consequence, modeling the DLS experiment is mathematically a more easily defined problem than for TEB, which requires the treatment of the difficult problem of the decay of polymers aligned by action of an external electric field.[10,48] DLS can also be performed, in principle, in both low- and high-salt solutions. It must, however, be kept in mind that DLS usually requires higher concentrations of macromolecules than TEB to obtain good signals and that at low-salt concentrations intermacromolecular electrostatic interactions become important at lower macromolecular concentrations. Thus, if one wishes to study single-chain properties of polyelectrolytes by DLS, one often cannot do so at very low-salt concentrations. Finally in TEB experiments, the amplitudes of the decay times have been shown[12] to depend on the duration of the orienting field. There is no quantitative model for the amplitudes, so they cannot at present be compared with theoretical simulations. It should also be noted that translational diffusion does not directly contribute to the TEB decay curve. Thus the decay curves contain information only about the internal motions (or perhaps overall rotation).[34,35] In DLS the translational diffusion of the entire coil contributes a pure diffusion peak (see eq 2 and 3) and also contributes a term to all of the other relaxation modes. Because the pure translational diffusion peak is close in decay time to the decay times of the internal motions, it affects the measurement of their decay times and amplitudes. This limits the resolution of the DLS experiment which would be necessary to obtain the detailed information that is obtainable from TEB. Because the two techniques complement each other, it is important to show that the results obtained from DLS are consistent with those obtained by TEB.

For the 2311 bp restriction fragment, the decay times from transient electric birefringence have been compared to those predicted by the Rouse–Zimm Gaussian coil model in both the free-draining and non-free-draining limits.[42] The first experimental decay time measured by TEB was 1376 μs. This was assumed to contain contributions from the first internal mode, τ_1. The first normal mode decay time was predicted, using eq 5 and R_g obtained at the lower salt concentration, to be 1168 μs in the free draining limit. In the non-free draining limit

$$\tau_1 = (23.6)\eta R_g^3/kT(4.04) \qquad (13)$$

It has a value of 3560 μs. These values differ from those reported by Lewis and Pecora[42] by a factor of 2 since the values reported by them were calculated on the basis of Zimm's normal mode[2] decay times while ours are calculated on the basis of Pecora's normal mode[3] decay times given by eq 5.

At an angle of 57°, the free draining Rouse–Zimm model for the DLS correlation function for a polymer with the radius of gyration of the 2311 bp fragment gives a simple two exponential form

$$S(x,t) = S_0(x)\exp(-q^2Dt) + S_2(x)\exp[-(q^2Dt + 2t/\tau_1)] \qquad (14)$$

Higher order terms are negligible, since they contribute at most 2% at this angle. The first normal mode decay time, τ_1, can be calculated from the second DLS decay time observed experimentally. A value of 532 μs was obtained from the DLS experiment. The model predicts values of 484 and 1644 μs in the free-draining and non-free-draining cases, respectively. These numbers suggest that the experimental results could best be described by a model with a small value for the hydrodynamic interaction parameter.

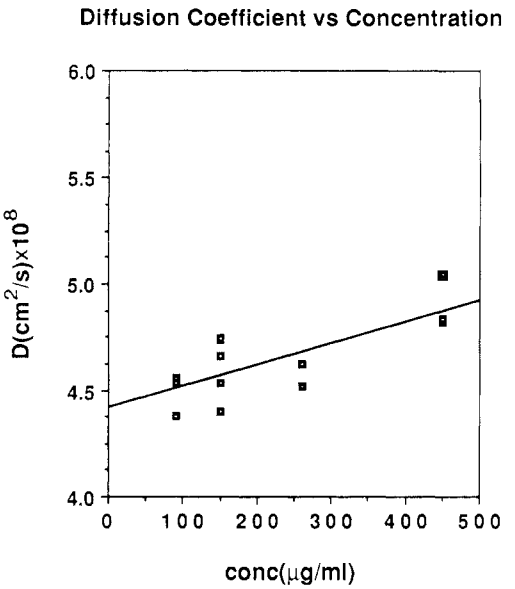

Figure 6. Diffusion coefficient increases with concentration in accordance with predictions based on the hard sphere interaction potential. The diffusion coefficient is obtained from measurements performed at low scattering angle (16° and 22°). Internal motion contributions are assumed to be insignificant at these scattering angles.

CONTIN, as discussed in section I, often moves two peaks that are hard to resolve farther apart than they really are. Since the experimental value of τ_1 was obtained by using CONTIN, it is necessary to show that it was not appreciably altered by CONTIN to give misleading results. If we analyze the simulated correlation function predicted in the free draining case by CONTIN and use eq 14, we obtain a value for the first normal mode decay time of 417 μs. Similarly, we would expect the apparent value in the non-free-draining case to decrease slightly when analyzed by CONTIN. However, the value for τ_1 obtained experimentally is still much closer to the value obtained from the model in the free-draining limit as was observed by TEB.

The results from TEB and DLS are in good agreement with each other and with the Rouse–Zimm model, despite the fact that the experiments were performed under different salt concentrations and thus the DNAs should exhibit different flexibilities. It would be interesting to compare DLS results for the 1010 bp DNA fragment to the TEB results for the 2311 bp fragment because at the different salt concentration ranges used in these two techniques, they would most likely have a similar number of Kuhn statistical segments. It should be pointed out, however, that there is some controversy over how much the DNA flexibility changes with salt concentration.[21]

Concentration Effects. At the highest concentration of 450 μg/mL, the ratio of the volume occupied by the DNA molecule to the total volume available for each molecule is 0.85. The volume occupied by the DNA was estimated, assuming a spherical shape, from the volume determined by the radius of gyration. We have assumed that for this concentration range, intermolecular interactions are weak and the diffusion coefficient depends linearly on concentration. Figure 6 gives a plot of the diffusion coefficient as a function of concentration. The diffusion coefficient was obtained from DLS experiments performed at low angle where the internal motions did not contribute significantly to the correlation function. The autocorrelation functions were collected at concentrations

Table II
Diffusion and Virial Coefficients

DNA sample	mol wt × 10^{-6}	$10^8 D_0$, cm²/s	$K_d = K_t - K_f$	K_t	K_f
pLH2311					
linear	1.53	4.42	1.30 ± 0.34		
ColE1[a]					
linear	4.35	2.16	1.25 ± 0.20		
ColE1[b]					
linear	4.3	1.98	1.12 ± 0.03[f]	6.23	2.57 ± 0.72
relaxed	4.15	2.45	1.42 ± 0.05[f]	9.75	6.45 ± 0.87
supercoil	4.59	2.89	1.38 ± 0.09[f]	2.86	6.63 ± 1.8
Fd DNA[c]					
single stranded	1.87	6.63	1.2 ± 0.4	7.6 ± 3.9[e]	6.7 ± 0.8
hard sphere model[d]			1.45	8.0	6.55

[a] Calculated from the diffusion vs concentration plot reported by Soda and Wada.[20] [b] Calculated from the plots and data reported by Voordouw, Kam, Borochov, and Eisenberg.[50] K_d was calculated from the diffusion vs concentration plot, K_f was calculated from sedimentation vs concentration plot, and K_t was calculated from the thermodynamic second virial coefficient. [c] Values reported by Newman, Swinney, Berkowitz, and Day.[51] [d] Theoretical values of Batchelor.[53] [e] Value reported by Berkowitz and Day.[52] [f] The reported error is the error in the slope after data from each concentration has been averaged. Thus, they are smaller than they would be if the error was calculated by using the raw data.

between 90 and 450 μg/mL. A linear least-squares fit to the data in Figure 6 was performed. Both the slope and the intercept of this line are important. The diffusion coefficient at infinite dilution is obtained from an extrapolation to zero concentration. The value obtained is $D_0 = 4.42 \times 10^{-8}$ cm²/s. This corresponds to a hydrodynamic radius of $R_h = 484$ Å. The slope of the line in Figure 6 is proportional to the virial concentration coefficient, K_d, given by

$$D(c) = D_0(1 + K_d\phi_v + ...) \quad (15)$$

ϕ_v is the volume fraction of the DNA fragment and is related to the partial molar volume of the DNA, V_h

$$\phi_v = NCV_h/MW \quad (16)$$

Here N is Avogadro's number, C is the concentration, and MW is the molecular weight. We have assumed that V_h is equal to the volume determined by the hydrodynamic radius. The value obtained for K_d is 1.3. It is a combination of two terms. These terms arise from a power series expansion of the friction coefficient and the partial derivative of the chemical potential with concentration. K_f is the contribution from the friction coefficient expansion in ϕ_v, and K_t is the contribution from the thermodynamic expansion.[49] The relationship between the three is

$$K_d = K_t - K_f \quad (17)$$

Table II gives the values of K_d, K_f, and K_t for our 2311 bp DNA fragment and values calculated or obtained from the results of other workers.[20,50-52] The value of K_d was calculated from the slope of the diffusion coefficient versus concentration plots from low-angle measurements. The K_f value can be obtained from the initial slope of a plot of the sedimentation coefficient versus volume fraction

$$s(c) = s(0)(1 - K_f\phi_v + ...) \quad (18)$$

K_t is calculated from the second osmotic virial coefficient, B_2

$$K_t = 2(MW)B_2C/\phi_v \quad (19)$$

The second osmotic virial coefficient was obtained from static light scattering measurements.

The value for K_d remains relatively constant independent of whether the DNA is single stranded, linear, relaxed circle, or supercoiled. It also remains constant over the molecular weight range of these studies. The values of K_f and K_t vary much more. The K_t values, which were obtained from the slope of the $q = 0$ line on a Zimm plot, are difficult to obtain accurately. Thus, the apparent variation in K_t is probably due to measurement errors. Most of the values obtained for K_f seem to fall very close to the value of 6.55 predicted by the hard-sphere model. The value of K_f for the linear ColE1 DNA measured by Voordouw et al.[50] is significantly lower than the others.

The value of $2\pi/q$ at the angle where the diffusion coefficients, shown in Figure 6, were measured is approximately 12 600 Å. The radius of gyration for the DNA is 1044 Å. Thus it is reasonable to assume that at this scattering angle the DNA molecules "see" each other as spheres and not as flexible coils with internal motions. Batchelor[53] has calculated values for K_d, K_f, and K_t based on the hard-sphere model, obtaining values of $K_d = 1.45$, $K_f = 6.55$, and $K_t = 8$. These values are in good agreement with the experimental values given in Table II. Numerous other calculations have been performed which incorporate other potentials.[54-59] Calculations have, for instance, been performed for spheres with both long-range repulsive interactions and short-ranged attractive interactions.[49] Our diffusion data are consistent with both of these calculations in the limit of small interaction parameters. More precise data on DNA under different solution conditions must be obtained before the applicability of these more refined models to DNA can be determined.

The CONTIN analysis at concentrations between 150 and 270 μg/mL and angles below ninety degrees shows only a very small concentration dependence. Above 90° there is possibly a larger concentration dependence. At these angles where the distribution function is very complex, increasing the concentration can have a large effect on the resolution since it increases the signal-to-noise ratio. Thus the "concentration dependence" observed above 90° in these lower concentration solutions is more likely to be a difference in the signal-to-noise ratio. This seems particularly true in light of the fact that only a small concentration effect is observed at lower angles where one might expect to observe a concentration effect first.

At all angles a concentration effect is evident when the lower concentration (150–270 μg/mL) results are compared to those at the highest concentration (450 μg/mL). This effect is much larger than the spread between data sets collected at different times which varied from less than a percent at 57 and 73° to approximately 10% at the higher angles up to 123°. At a scattering angle of 123° the data were difficult to reproduce precisely, and thus by itself there is not significant information to determine if there is a concentration dependence at this angle.

Figure 7 shows the CONTIN analysis at 73, 90, and 123°. The solid circles give the analysis at lower concentrations, and the squares give the analysis at the higher concen-

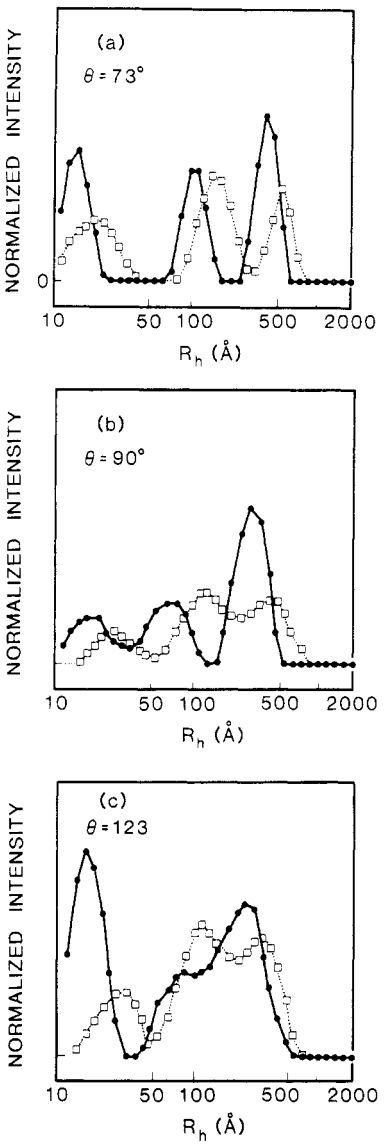

Figure 7. CONTIN analysis of the correlation function at 73°, 90°, and 123°. The (●) represents experimental data at the lower concentrations (150–250 μg/mL). The (□) represents experimental data at the higher concentration (450 μg/mL).

tration. We can see that at all three angles the peak positions move to larger R_h. This corresponds to a slowing down of the polymer motions. At an angle of 73° the slowest peak moves from 435 Å at 150 μg/mL to 537 Å at 450 μg/mL. It also contributes less to the scattering intensity at the higher concentration (70% vs 82%). The middle peak seems to increase in contribution suggesting that at a higher concentration the longest internal motions of the coil become more significant. This is in direct contrast to that observed at lower angles. Figure 6 shows that at low angle the diffusion coefficient of the polymer increases with concentration. At the highest concentration and at low angle the value obtained for the hydrodynamic radius was 435 Å. At the same concentration but at higher angles, the value for the hydrodynamic radius obtained from the slowest peak in the CONTIN analysis is 537 Å. This value could also contain unresolved contributions from the internal motions which would decrease the R_h in comparison to that observed in the absence of internal motions. Thus the hydrodynamic radius at high angle increases by at least 15% over that observed at low angles and the lower concentrations.

One possible explanation for the discrepancy is that at high angles we are observing self diffusion, while at the low angle we observe mutual diffusion. Assuming that the observed decrease in diffusion coefficient is linear in concentration, a value for $K_d = -1.22$ is calculated by using eq 15. Batchelor[53] calculated a value for $K_d = -1.83$ for the self-diffusion of particles in a hard-sphere potential. The difference between these numbers could be explained by unresolved contributions from the internal modes that increase the apparent diffusion coefficient as well as by deviations of DNA from the hard-sphere model used in the calculations.

It is also likely that the apparent diffusion coefficient obtained from the high-angle measurements is neither the true self-diffusion coefficient nor the mutual diffusion coefficient but a q-dependent one measuring the decay of a fluctuation with some collective character. At 450 μg/mL, the average interparticle spacing is approximately 2400 Å. Fluctuations on a scale of approximately $2\pi/q$ are measured, which at 73° scattering angle corresponds to a length of about 3000 Å.

Finally it is also possible that the apparent decrease in the diffusion coefficient is due to an artifact in the data analysis method. The internal motions seem to be slowing down and contributing more to the correlation function. It is possible that this has affected the data analysis by forcing the peak due to translational motion to higher hydrodynamic radius than would really be observed in the absence of the internal motions. Much more extensive experiments are needed to sort out all these possibilities.

In the cumulant plots in parts a and b of Figure 5 the higher concentration points seem to lie beneath the other data points. This corresponds to a slowing down of the motions of the plasmid as concentration is increased and is consistent with what we have already observed from the distribution function from CONTIN.

V. Conclusion

We have presented experimental relaxation time distribution functions that describe the translational and long-range internal motions of a linear DNA fragment in dilute solution. These distribution functions have been compared with simulated distribution functions predicted by the Rouse–Zimm model in the absence of hydrodynamic interactions. The agreement is surprisingly good, in spite of the simplicity of the model. The first internal mode relaxation time extracted from the lowest angle data showing two peaks is consistent with that predicted from the free draining model and is in agreement with that extrapolated from TEB measurements on the same fragment. In addition cumulants obtained from second-order cumulant fits and first moments of the distribution function obtained from CONTIN have at higher values of q been shown to vary approximately as q^3. Apparent diffusion coefficients obtained from the CONTIN fits at both low and high scattering angles have been studied as functions of concentration.

These studies, in addition to the intrinsic importance of the results, illustrate the power and some of the limitations of using the CONTIN data analysis method to study these systems as well as the deficiencies of the cumulant method. DLS studies on smaller DNA fragments (with laser wavelengths extending into the UV) combined with more sophisticated computer simulations using the data analysis and comparison methods given here should prove to be very valuable in understanding the dynamics of DNA in particular and semiflexible molecules in general.

Acknowledgment. This work was supported by NSF Grant CHE-85 11178 to R.P. and by the NSF MRL pro-

gram through the Center for Materials Research at Stanford University. We are grateful to Dr. Roger J. Lewis who initiated the DNA program in our laboratory and provided invaluable advice and encouragement during this study.

References and Notes

(1) Rouse, P. E., Jr. *J. Chem. Phys.* **1955**, *21*, 1272.
(2) Zimm, B. H. *J. Chem. Phys.* **1956**, *24*, 269.
(3) Pecora, R. *J. Chem. Phys.* **1965**, *43*, 1562; **1968**, *49*, 1032.
(4) Saito, N.; Ito, S. J. *J. Phys. Soc. Jpn.* **1968**, *25*, 1446.
(5) de Gennes, P.-G. *Physics* **1967**, *3*, 37. Dubois-Violette, E.; de Gennes, P.-G. *Physics* **1967**, *3*, 181.
(6) Borochov, N.; Eisenberg, H.; Kam, Z. *Biopolymers* **1981**, *20*, 231.
(7) Borochov, N.; Eisenberg, H.; Kam, Z. *Biopolymers* **1981**, *20*, 2671.
(8) Yamakawa, H. *Annu. Rev. Phys. Chem.* **1984**, *35*, 23.
(9) Eden, D.; Elias, J. G. In *Measurements of Suspended Particles by Quasi-Elastic Light Scattering*, Dahneke, B. E., Ed.; Wiley: New York, 1983; p 401.
(10) Porschke, D. *Annu. Rev. Phys. Chem.* **1985**, *36*, 159.
(11) Stockmayer, W. H.; Baur, M. E. *J. Am. Chem. Soc.* **1964**, *86*, 3485.
(12) Lewis, R. J.; Pecora, R.; Eden, D. *Macromolecules* **1986**, *19*, 134.
(13) Nash, P. J.; King, T. A. *Polymer* **1985**, *26*, 1003.
(14) Koppel, D. E. *J. Chem. Phys.* **1972**, *57*, 4814.
(15) Akcasu, A. Z.; Benmouna, M.; Han, C. C. *Polymer* **1980**, *21*, 866.
(16) Wilcoxon, J.; Schurr, J. M. *Biopolymers* **1983**, *22*, 849.
(17) Maeda, T.; Fujime, S. *Macromolecules* **1984**, *17*, 1157, 2381.
(18) Lin, S. C.; Schurr, J. M. *Biopolymers* **1978**, *17*, 425.
(19) Thomas, J. C.; Allison, S. A.; Schurr, J. M.; Holder, R. D. *Biopolymers* **1980**, *19*, 1451.
(20) Soda, K.; Wada, A. *Biophys. Chem.* **1984**, *20*, 185.
(21) Schurr, J. M.; Schmitz, K. S. *Annu. Rev. Phys. Chem.* **1986**, *37*, 271.
(22) Provencher, S. W.; Hendrix, J.; De Maeyer, L.; Paulussen, N. *J. Chem. Phys.* **1978**, *69*, 4237.
(23) Provencher, S. W. *Makromol. Chem.* **1979**, *180*, 201.
(24) Provencher, S. W. *CONTIN User's Manual*, Technical Report EMBL-DA02; European Molecular Biology Laboratory: Heidelberg, 1980.
(25) Provencher, S. W. *Comput. Phys. Commun.* **1982**, *27*, 213, 229.
(26) *Particle Size Distribution: Assessment and Characterization*; Provder, T., Ed; American Chemical Society: Washington, DC, 1987.
(27) Lewis, R. J.; Huang, J. H.; Pecora, R. *Macromolecules* **1985**, *18*, 944.
(28) Langowski, J.; Giesen, U.; Lehmann, C. *Biophys. Chem.* **1986**, *25*, 191.
(29) Lewis, R. J.; Huang, J. H.; Pecora, R. *Macromolecules* **1985**, *18*, 1530.
(30) Marko, M. A.; Chipperfield, R.; Birnboim, H. C. *Anal. Biochem.* **1982**, *121*, 382.
(31) Maier, K. R. Ph.D. Thesis, Stanford University, Stanford, CA, 1986.
(32) Provencher, S. W. *Biophys. J.* **1976**, *16*, 27.
(33) Provencher, S. W. *J. Chem. Phys.* **1976**, *64*, 2772.
(34) Aragón, S. R. *Macromolecules* **1987**, *20*, 370.
(35) Aragón, S. R.; Pecora, R. *Macromolecules* **1985**, *18*, 1868.
(36) Bott, S. E. Ph.D. Thesis, Stanford University, Stanford, CA, 1984.
(37) Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning, A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1982.
(38) Perico, A.; Piaggio, P.; Cuniberti, C. *J. Chem. Phys.* **1975**, *62*, 2690.
(39) Lewis, R. J. Ph.D. Thesis, Stanford University, Stanford, CA, 1985.
(40) Kovacic, R. T.; van Holde, K. E. *Biochemistry* **1977**, *16*, 1490.
(41) Han, C. C.; Akcasu, Z. A. *Macromolecules* **1981**, *14*, 1080.
(42) Lewis, R. J.; Pecora, R. *Macromolecules* **1986**, *19*, 2074.
(43) Allison, S. A. *Macromolecules* **1986**, *19*, 118.
(44) Lewis, R. J.; Allison, S. A.; Eden, D.; Pecora, R., submitted for publication in *J. Chem. Phys.*
(45) Schaefer, D. W.; Han, C. C. In *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy*; Pecora, R., Ed.; Plenum: New York, 1985.
(46) Yamakawa, H. *Modern Theory of Polymer Solutions*; Harper and Row: New York, 1971.
(47) Adam, M.; Delsanti, M. *Macromolecules* **1977**, *10*, 1229.
(48) Lewis, R. J.; Pecora, R.; Eden, D. *Macromolecules* **1987**, *20*, 2579.
(49) Pusey, P. N.; Tough, R. J. A. In *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy*; Pecora, R., Ed.; Plenum: New York, 1985.
(50) Voordouw, G.; Kam, Z.; Borochov, N.; Eisenberg, H. *Biophys. Chem.* **1978**, *8*, 171.
(51) Newman, J.; Swinney, H. L.; Berkowitz, S. A.; Day, L. A. *Biochemistry* **1974**, *13*, 4832.
(52) Berkowitz, S. A.; Day, L. A. *Biochemistry* **1974**, *13*, 4825.
(53) Batchelor, G. K. *J. Fluid Mech.* **1976**, *74*, 1.
(54) Yamakawa, H. *J. Chem. Phys.* **1962**, *36*, 2995.
(55) Imai, S. *J. Chem. Phys.* **1969**, *50*, 2116.
(56) Pyun, C. W.; Fixman, M. *J. Chem. Phys.* **1964**, *41*, 957.
(57) Altenberger, A. R.; Deutch, J. M. *J. Chem. Phys.* **1973**, *59*, 894.
(58) Akcasu, A. Z.; Benmouna, M. *Macromolecules* **1978**, *11*, 1193.
(59) Akcasu, A. Z. *Polymer* **1981**, *22*, 1169.

Scattering Theory and Properties of Block Copolymers with Various Architectures in the Homogeneous Bulk State

H. Benoit

Institut Charles Sadron (CRM-EAHP), 6 rue Boussingault, 67083 Cedex, France

G. Hadziioannou*

IBM Research, Almaden Research Center, 650 Harry Road, San Jose, California 95120-6099. *Received July 13, 1987; Revised Manuscript Received October 13, 1987*

ABSTRACT: We have developed a general theory to predict the scattering profiles in the homogeneous state and the phase separation behavior of multiblock copolymers having various architectures such as "linear", "comb", and "star". We found that the scattering intensity versus angle in the homogeneous state has a universal slope at small angles independent of the architecture of the block copolymers. The position of the maximum intensity is almost independent of the number N of sequences, and its amplitude rapidly reaches a plateau value, indicating that the microphase separation for $N \geq 20$ is independent of N. Detailed calculations concerning the molecular dimensions and conformations and the thermodynamic behavior of the three types of the multiblock copolymer architectures have been performed based on fundamental variables such as the radius of gyration of each block, the composition, the degree of polymerization, the number of sequences, and the polymolecularity in molecular weight and composition.

Introduction

Block copolymers at the homogeneous state have been the subject of extensive theoretical studies,[1-12] but usually the authors have restricted their interest to copolymers made only of a few blocks.

In this paper we would like to generalize these known

0024-9297/88/2221-1449$01.50/0