Theory of DNA Condensation: Collapse Versus Aggregation

CAROL BETH POST* and BRUNO H. ZIMM, *Department of Chemistry, B-017, University of California at San Diego, La Jolla, California 92093*

Synopsis

In an unfavorable solvent environment, DNA (and other polymers) undergo a conformational transition to a collapsed form, accompanied by a dramatic reduction in the effective volume of the molecule. Solvent conditions leading to the collapse are the same as those that cause aggregation. We give here a thermodynamic description of the collapse and its relations to aggregation (or precipitation). This is formulated in terms of the Flory-Huggins theory of the thermodynamics of polymer solutions. The results show that it is possible for three different states of DNA to be stable under different conditions: (1) the extended random coil, (2) the collapsed coil, and (3) a concentrated phase of aggregated random coils. The collapsed coil is predicted to be stable against aggregation only at high dilutions, of the order of parts per million. For DNA the transition between the extended coil and the collapsed coil is predicted to be discontinuous, in the sense that intermediate states are not present, because of the relatively high stiffness of the chain. The transition should appear diffuse because of the small size of the single molecule in comparison to macroscopic systems.

INTRODUCTION

In an unfavorable solvent environment there is a tendency for the segments of a polymer chain to associate with other polymer segments, reducing their interaction with the solvent. These associations can be either *intra*molecular, leading to a collapsed structure of a single polymer chain, or *inter*molecular, leading to an aggregated polymer phase. As a result of the intramolecular associations, the collapsed-chain radius of gyration, R_g, is substantially decreased relative to that of the extended chain in a good solvent or in an aggregated phase. Both internal condensation, a monomolecular phenomenon, and external condensation, the association of many molecules, produce states of high segment density, conditions with a lower solution free energy due to the increased number of like–like contacts and the decreased number of the less favorable contacts between polymer and solvent. The higher segment density is achieved in the former case by reducing the solvent volume occupied by a single polymer molecule (smaller R_g), and in the latter case by an increase in the number of molecules per volume of solution. It is to be expected that aggregation becomes more favorable at higher polymer concentrations.

* Present address: Inorganic Chemistry Laboratory, Oxford, England, OX1 3QR.

Biopolymers, Vol. 21, 2123–2137 (1982)

CCC 0006-3525/82/112123-15$02.50

We use the term *condensation* to refer to the formation of a polymer-rich state of high segment density; therefore, condensation applies both to aggregation of many molecules and to compaction of a single molecule. The term *collapse*, on the other hand, is reserved for the single-molecule phenomenon of a large reduction in R_g.

In an earlier publication[1] we described the collapse transition as derived from the free energy function, ΔG, for the mixing of an isolated polymer with solvent molecules. Following the Flory polymer solution theory, ΔG was expressed in terms of χ, the parameter of interaction between the polymer and solvent, and α, the linear expansion parameter ($\alpha \equiv \langle R_g \rangle / \langle R_{g_0} \rangle$, with R_{g_0} the unperturbed radius of gyration in a theta-solvent). The equilibrium configuration of the polymer was determined by the minimization of ΔG with respect to α for a given χ. The collapse transition is the point at which the global minimum changes from a value of α close to one to a value of α much less than one.

It is not, of course possible to measure the behavior of a single polymer molecule experimentally; therefore, one must consider the consequences of *inter*molecular associations on the free energy of the system, recognizing the possibility of an aggregated polymer state. We now extend the theory to describe a solution of many polymer molecules by including the effect of polymer concentration on the mixing process. How the collapse transition relates to the equilibrium between macroscopic phases of the solution is the subject of this paper. A brief summary of the results has already appeared.[2] Lifshitz et al.[3] and Swislow et al.[4] have independently arrived at similar conclusions concerning the phase diagram of polymer solutions.

The free energy of mixing of a polymer solution is found by adding two previously known expressions for ΔG—the single molecule mixing free energy[1] and Flory's free energy expression for the mixing of several molecules.[5] Though the equations presented here are applicable to all polymer solutions, the numerical calculations refer to DNA solutions. The phase diagram for a DNA solution is calculated as a function of χ and DNA concentration from ΔG. In addition to the well-known coexistence curve that separates the concentrated phase from the dilute solution phase of expanded coils (Ref. 5, Chap. 12), the phase diagram contains boundaries defining the equilibrium domain of the collapsed state. Within the region of immiscibility, the phase concentrated in polymer molecules may be in equilibrium with either the dilute phase of extended coils or the dilute phase of collapsed coils. In going from a good to a poor solvent, two transitions are possible in a DNA solution: (1) at higher DNA concentrations, the transition from a solution of extended random coils to an aggregated precipitate, and (2) at lower DNA concentrations, the transition from the extended coil to the collapsed coil in solution. A study to distinguish the collapse region of the phase diagram from the aggregation region using light scattering from high-molecular-weight DNA is reported in an accompanying paper.[6]

We also include a calculation of the effects of fluctuations on the value of $\langle \alpha^2 \rangle$, the average of equilibrium. Our earlier work[1] was concerned with only the value of α corresponding to the minimum in ΔG of mixing for a single molecule. This work, however, focuses on the properties of a solution of polymer molecules; thus, the average configuration, which is experimentally observable, is of interest. In single molecules, fluctuations from the minimum of ΔG are sizeable; they affect the transition curve, making it noticeably diffuse.

THE FREE-ENERGY FUNCTION

The change in free energy on dissolution of a population of polymer molecules is obtained by adding two previously known expressions for mixing on a lattice. The polymer concentration dependence is introduced by adding the Flory-Huggins equation for a polymer mixture (Ref. 5. Chap. 12) to the single-molecule mixing expression presented earlier.[1] Several accounts of the Flory-Huggins theory have appeared.[5,7-9] Flory's own comprehensive description[5] is perhaps the most lucid and includes a thorough discussion of the limitations of the lattice model. A detailed explanation of Flory's theory will not be given here, except for some remarks on the approximations in the model that are relevant to this work.

Hence, there are two parts to the ΔG of mixing. The first part, the Flory-Huggins equation, is the external free energy of mixing, which is concerned with the placement of solvent and of disoriented polymer molecules in a common system. The external free energy includes the increase in entropy as a result of a larger available volume, plus that part of the heat of mixing due to the interaction of two polymer molecules. As before,[1] the free energy of interaction is expressed in terms of the unitless parameter χ. The external free energy can be thought of as arising from the mixing of chains whose *inter*molecular collisions consist of a sum of independent single segment–segment contacts. The second part, the single molecule equation for dissolution, is the internal free energy of mixing. This ΔG of mixing involves the effects of dilution on an individual polymer molecule, that is, the change in the configurational statistics of a single chain in response to the solvent influence on the *intra*molecular potentials.

The total mixing free energy for the polymer and solvent mixture is

$$\Delta G = \Delta G_{\text{ext}} + n_2 \Delta G_{\text{int}}$$

$$\frac{\Delta G}{kt} = n_1 \ln v_1 + n_2 \ln v_2 + \chi n_1 v_2 + n_2 \left[N \left((\chi - 1) + \frac{B_2 \omega}{2^{3/2} \alpha^3} + \frac{B_3 \omega^2}{2 \cdot 3^{5/2} \alpha^6} \right) + \frac{3}{2} (\alpha^2 - 1) - \ln \alpha^3 \right] \quad (1)$$

where n_1 is the number of solvent molecules and v_1 is the solvent volume fraction. The quantities n_2 and v_2 are the same for the polymer. N is the ratio of the molecular volume of the polymer to the solvent, so that $N =$

V_p/V_1. The expansion parameter, α, is the ratio of R_g to the unperturbed radius of gyration, R_{g_0} (i.e., the radius of gyration when the net excluded-volume effect is zero). The quantities B_2, B_3, and ω, as defined previously,[1] are given below:

$$B_2 = \tfrac{1}{2} - \chi$$

$$B_3 = 1 + 12\chi^2/q - 16\chi^3/q^2$$

$$\omega = \left(\frac{9}{\pi\langle h_0^2\rangle}\right)^{3/2} V_p$$

with $\langle h_0^2\rangle$ the unperturbed end-to-end length and q the lattice coordination number. The first three terms in Eq. (1) represent the external free energy. The increase in entropy due to the external arrangement of whole polymer molecules and solvent molecules is expressed in the familiar form of the logarithm of a concentration. The third quantity is a function of the interaction parameter, χ, and contains the free energy due to the formation of segment–solvent contacts at the expense of solvent–solvent contacts and *inter*molecular segment–segment contacts. These three terms make up the expression that Flory derived (Ref. 5, p. 509) for the free energy of mixing as a function of polymer concentration. The square brackets of Eq. (1) contain the terms for the internal free energy of mixing per polymer chain given by Eq. (5) of Ref. 1. As discussed in detail in Ref. 1, they account for the change in free energy associated with the mixing of segments within the domain of one polymer. The contributions from the heat and hard-core repulsion are given in a series expansion of the segment volume fraction *within the polymer domain*, i.e., ω/α^3. The elastic nature of the chain is expressed by the last terms, which are functions of only α.

The Interaction Parameter χ

The difference between the interaction free energies of like and unlike species is expressed in terms of the interaction parameter χ. In Eq. (1), χ is used for both the external and internal mixing processes. As such, the contact between segments as a result of a collision between two polymer molecules is assumed to be equivalent in free energy to the binary contacts between segments within one molecule. This is reasonable, since a contact between segments belonging to two different chains should be similar to a contact between segments from one chain. The molecular details and the form of the potential of interaction are not specified.

A brief discussion of the meaning of χ is needed to clarify the link between χ and experimental parameters. The complete theoretical basis of χ as it was in the early developments of the lattice model is given by Flory (Ref. 5, Chap 12). In short, χ was assumed to be proportional to the difference in the energy of contacts between unlike and like species. This was written mathematically in terms of h_{ij}, the energy associated with a contact

between species i and species j, as

$$\chi \propto h_{12} - \tfrac{1}{2}(h_{11} + h_{22}) \tag{2}$$

with 1 and 2 referring to solvent and segment, respectively. An increase in h_{ij} is equivalent to increasing the repulsive energy or decreasing the attractive energy between i and j. Following the first developments of the Flory theory, it was discovered that experimentally determined values of χ show a dependence on concentration, temperature, and molecular weight. Therefore, a more appropriate formulation for the interaction parameter is a free-energy function with an entropic component as well as an enthalpic one.[9-11] Nonetheless, the heat terms of Eq. (2) are still associated with χ.

Let us then consider h_{12} to be the solvent–segment free energy of interaction. An actual solvent of a DNA solution is usually composed of many components that together make up the environment surrounding the DNA segments. Miscibility is determined by the segment interaction with the environment as a whole; thus, h_{12} represents here the overall free energy of contact between a segment and all the solvent components.

Clearly, χ changes whenever there is a change in any one of the terms h_{12}, h_{11}, or h_{22}. Examination of Eq. (2) finds that χ increases (DNA is less soluble) either when h_{12} increases or when h_{11} or h_{22} decreases. For example, neutralization of the phosphate charges with spermidine reduces the repulsion between DNA segments, decreasing the segment–segment interaction parameter, h_{22}. An increase in the segment–solvent interaction parameter, h_{12}, can be produced by the addition of nonsolvents such as poly(ethylene oxide) (PEO) or ethanol, since such a change in the environment makes the segment interaction with the solvent less favorable. Classification of the types of interactions based on the molecular details of the interaction is not important here; neither the form of the potential nor the specific mechanism of DNA condensation needs to be addressed. A change in any one of the h_{ij} terms in Eq. (2) alters χ. The theory is applicable to all DNA solutions, since χ is an overall measure of the favorability of the intereaction of the DNA with the solvent relative to the other interactions.

Hence, χ is best considered to be an empirical parameter that becomes $\frac{1}{2}$ at the theta point. In synthetic polymer solutions, experiments have shown that χ varies somewhat with the concentration[5,9-11]; but since there are no data on this matter with DNA solutions, we must assume χ to be constant. Although χ appears only in the term of the first power of the polymer concentration in ΔG_{ext} of Eq. (1), χ appears in both the second and third terms of the expansion in the segment volume fraction ΔG_{int}. Since it is essential for the prediction of the collapsed state to keep the third term in the series, we retain χ in this term, as discussed in Ref. 1.

We note that the theory presented here is limited; however, the expressions are not intended to be accurate but simply to describe semiquantitatively the behavior of DNA solutions.

THE PHASE DIAGRAM

The attractive potential between segments leads to a phase separation when χ is greater than $\frac{1}{2}$. Whether the association of segments occurs within a single molecule, resulting in collapsed structures, or by interaction of several molecules, resulting in aggregation, depends on the polymer concentration. The ΔG of mixing from Eq. (1) was used to determine the phase equilibrium of DNA solutions in the following way.

The chemical potentials of the two components of the mixture, μ_1 and μ_2, can be obtained by differentiating ΔG with respect to n_1 and n_2, respectively:

$$\frac{\mu_2 - \mu_2^0}{kT} = \ln v_2 - (N-1)(1-v_2) + \chi N(1-v_2)^2 + N\left((\chi - 1) + \frac{B_2\omega}{2^{3/2}\alpha^3} + \frac{B_3\omega^2}{2 \cdot 3^{5/2}\alpha^6}\right) + \frac{3}{2}(\alpha^2 - 1) - \ln \alpha^3 \quad (3)$$

$$\frac{\mu_1 - \mu_1^0}{kT} = \ln(1-v_2) + (1 - 1/N)v_2 + \chi v_2^2 \quad (4)$$

where μ_1^0 and μ_2^0 are the chemical potentials of the pure phases.

The curves for μ_1 and μ_2 versus v_2 are single-valued functions for $\chi < \frac{1}{2}$; thus, only one phase occurs. For larger values of χ, the curves are multiple-valued; thus coexisting phases are possible. The concentration of polymer in the two phases in equilibrium is then determined by equating the polymer and the solvent chemical potentials in the polymer-dilute phase to their respective potentials in the polymer-rich phase. Representing the concentrated polymer phase by a prime and the dilute phase by the absence of a superscript, we write

$$\mu_1 = \mu_1' \quad (5)$$

$$\mu_2 = \mu_2' \quad (6)$$

Thus, for a χ greater than $\frac{1}{2}$, the phase compositions are determined by the two values of the polymer volume fraction, v_2 and v_2', which simultaneously satisfy conditions (5) and (6). The coexistence curve in the phase diagram of χ plotted against composition was determined by using a series of Newton-Raphson iterations to solve for v_2 and v_2'.

Three χ-versus-composition phase diagrams are plotted in Fig. 1. Each curve represents a different DNA molecular weight, with the appropriate $\langle h_0^2 \rangle$ and N values given in the figure caption. The concentration units along the abscissa are μg/mL, converted from the less familiar units of v_2, mL/mL, by the density of DNA, 1.8 g/mL.[12] For each molecular weight, the boundaries separate three regions—extended random coils in solution, collapsed DNA in dilute solution, and concentrated precipitate ("aggregated"). These regions are labeled for the smallest molecular weight only. The horizontal line on the dilute solution side marks the value χ_{col}, the value of χ for which the free energies of the extended and collapsed coils are equal.

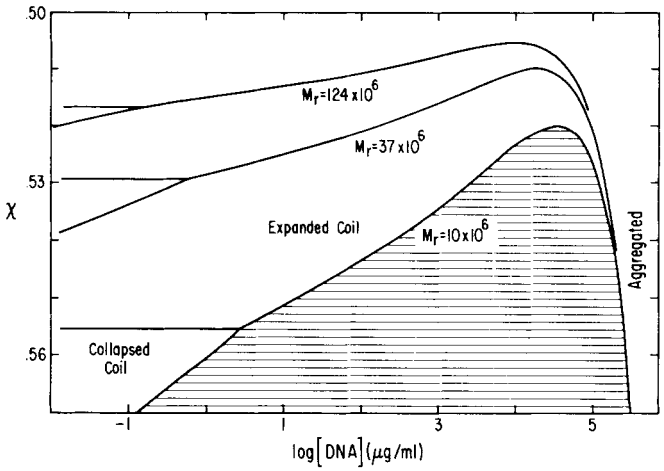

Fig. 1. Phase diagram of χ versus composition for DNA of different molecular weights. For $M_r = 124 \times 10^6$, $\langle h_0^2 \rangle^{1/2} = 2.5 \times 10^{-4}$ cm and $N = 3.1 \times 10^4$; for $M_r = 37 \times 10^6$, $\langle h_0^2 \rangle^{1/2} = 1.3 \times 10^{-4}$ cm and $N = 9300$; and for $M_r = 10 \times 10^6$, $\langle h_0^2 \rangle^{1/2} = 6.2 \times 10^{-5}$ and $N = 2500$. The two-phase region is marked with horizontal lines.

As discussed below in detail, this value of χ marks the center of a diffuse transition from the extended to the collapsed state, the relative populations in the two minima of the free-energy function changing as χ is changed. The procedure for determining χ_{col} was reported in Ref. 1.

The shaded area marks the two-phase region, with the compositions of the coexisting phases being the values that lie on the curve (as just discussed). The maximum in the coexistence curve indicates the solution composition and χ value critical for phase separation of the two-component system. This χ value is designated by χ_{sep} to distinguish it from the critical-point value discussed in Ref. 1.

For values of χ less than χ_{sep}, υ_2 is a single-valued function of μ_1 and μ_2, and the solution is a homogeneous mixture of solvent and extended, randomly coiled DNA at all concentrations. For solvent conditions in which like contacts are favored, χ is greater than χ_{sep}; υ_2 is a multiple-valued function of μ_1 and μ_2, and there is a region of immiscibility. At high DNA concentration, the phase separation results in aggregation. On the other hand, at low DNA concentration, as χ is increased there is first a transition from the extended polymer to the densely packed monomolecular collapsed state, with progression through the collapse region eventually leading again to aggregation. With decreasing DNA concentration, the collapse region widens in the vertical direction, indicating a larger range in χ within which the DNA is collapsed before aggregation appears. Two equilibria are possible when the mixture is heterogeneous: (1) the concentrated DNA phase may coexist with a dilute solution of extended coils when χ is between χ_{sep} and χ_{col}, and (2) the concentrated DNA phase may coexist with a dilute solution of collapsed DNA for χ greater than χ_{col}.

The maximum of the coexistence curve is found by the usual procedure of setting both the first and second derivatives with respect to v_2 of either μ_1 or μ_2 equal to zero and solving for χ_{sep} and $v_{2,sep}$. The results are given by Flory (Ref. 5, Chap. 13) and are as follows:

$$v_{2,sep} = 1(1 + \sqrt{N}) \tag{7}$$

$$\chi_{sep} = (1 + \sqrt{N})^2/2N \tag{8}$$

It is seen that the position of the maximum in the χ-composition diagram is a function of N only. The effects of N on the phase diagram will be discussed more fully in the following section.

χ_{sep} and χ_{col} are functions of molecular weight. At infinite molecular weight, $\chi_{sep} = \chi_{col} = 1/2$. For finite M_r, both are slightly greater than $1/2$, with χ_{col} increasing somewhat more rapidly with a decrease in M_r than does χ_{sep}. Poorer solvent conditions (larger χ) are required for phase separation of lower-molecular-weight polymers.

At smaller values of M_r the collapse region extends to higher polymer concentrations. The phase diagram in Fig. 1 shows that the DNA concentration at which the horizontal line for χ_{col} intersects the coexistence curve, outlining the region of immiscibility, is approximately 2–3 μg/mL for M_r of 10×10^6, and only 0.6 μg/mL for M_r of 37×10^6. The collapse region also spans a broader χ range for a lower molecular weight. The collapse of smaller polymers should therefore occur over a larger range in solvent conditions (such as collapsing-reagent concentration or temperature) before the polymer precipitates.

Examination of Eq. (3) for the case of pure polymer ($v_2 = 1$) shows that this expression for $\mu_2 - \mu_2^0$ cannot be exact. For the pure polymer state, μ_2 should equal μ_2^0; however, when χ is set equal to zero and v_2 and α are set equal to one, the terms in the square brackets of Eq. (3) are nonzero. The magnitudes of the two terms in ω are negligible, but the term $(-N)$ is not. $N(\chi - 1)$ comes from integration of the single-molecule ΔG function as originally formulated by Flory (Ref. 5, Chaps. 12 and 14). This problem is left unresolved, since the $N(\chi - 1)$ term has no concentration dependence and does not alter the solution equilibrium between two phases.

Effects of N on the Phase Diagram

Some comments about the parameter N are in order. As the ratio of the molecular volumes of polymer (V_p) and solvent (V_1), N can change significantly with V_1. In the lattice model, V_1 equals V_s, the segment volume. The proper choice for V_s is sometimes not apparent. It is possible in some synthetic polymer solutions to apply the lattice model without complication and assign values for V_s equal to the molecular volume of the solvent. In these polymer–solvent systems, the monomer unit and the solvent are structurally similar, with the molecular volume of the solvent equivalent to 1 or to 3 monomer units.[13] The connection between the experimental system and the lattice model is obvious and allows a reasonable choice for

V_s. Still, in some such systems where the monomer unit and solvent molecules are similar, discrepancies between theory and experiment indicate that a better agreement could be obtained if a segment size larger than the monomer unit were chosen.[14] In the case of DNA, it is uncertain what structural criterion should be used to define a segment. Without an empirical determination for N, its most appropriate value remains unknown.

The phase diagrams in Fig. 1 were calculated using V_s equivalent to six base pairs; hence, the segment unit is approximately spherical to make at least some tie of DNA to the lattice model. The V_s reported in Ref. 1 is different, since the work reported there was completed before this tie was considered.

The differences in the phase diagram that result from a change in N are illustrated in Fig. 2, with V_s corresponding to one base pair for $N = 5.6 \times 10^4$ and to six base pairs for $N = 9.3 \times 10^3$, with M_r and $\langle h_0^2 \rangle$ held constant. The choice of N does not alter the overall picture of the phase behavior, only the details of the shape. The main effect of a different N is practically just a change in the χ scale, for which there is no consequence without an empirical determination of χ for DNA.

THE COLLAPSE TRANSITION IS A DIFFUSE TRANSITION

The two minima in ΔG representing the expanded and collapsed forms are within a few multiples of kT of each other over an appreciable range of χ for molecules of finite size. We assume that the system distributes itself between the two minima according to a Boltzmann distribution. (This problem has apparently not been addressed in previous theoretical papers[1,3,15] concerning the single-molecule phase transition; see also other references given in Ref. 1.)

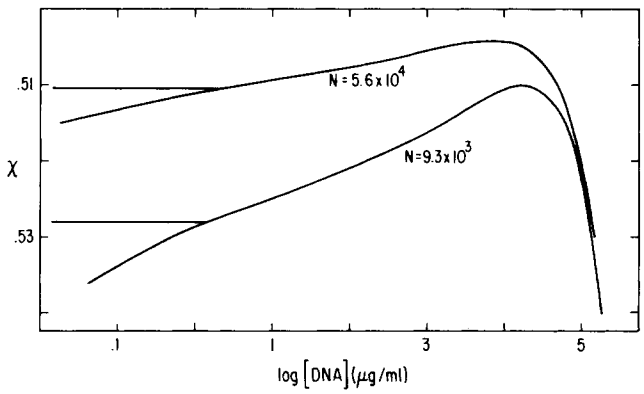

Fig. 2. Effect of N, the ratio of the molecular volume of the polymer to the segment, on the χ versus composition phase diagram. Both coexistence curves were calculated for DNA with $M_r = 37 \times 10^6$ and $\langle h_0^2 \rangle^{1/2} = 1.3 \times 10^{-4}$ cm.

Since most experimental methods measure some average property of the solution, a procedure for averaging various functions of α is needed. Light scattering, for example, depends on R_g^2 of the scattering particle; thus, the quantity $\langle \alpha^2 \rangle$ is of importance. The average is calculated as a function of χ, using the Boltzmann distribution, according to the following equation:

$$\langle \alpha^2 \rangle = \frac{\int \alpha^2 e^{-\Delta G/kT}\, d\alpha}{\int e^{-\Delta G/kT}\, d\alpha} \tag{9}$$

ΔG is the single-molecule free energy of mixing in Eq. (5) of Ref. 1 [also the internal free energy in Eq. (1)]. In a similar fashion, $\langle \alpha \rangle$ and $\langle 1/\alpha \rangle$ can be easily determined. The effects of averaging are illustrated for a DNA solution in Fig. 3, with $\langle \alpha^2 \rangle^{1/2}$ plotted against χ. The results were obtained by numercial integration of Eq. (9). For purposes of comparison the nonaveraged values of $\alpha_{\min}$, corresponding to the global minimum in ΔG, are also plotted. The value $\langle \alpha^2 \rangle^{1/2}$ is larger than $\alpha_{\min}$ when α is close to one, because of a broad minimum in ΔG for the expanded state. In contrast to the sharp discontinuity in $\alpha_{\min}$ at χ_{col} (indicated by the dashed line), the averaging process produces a smooth transition to the collapsed state, with a broader range in χ over which $\langle \alpha^2 \rangle^{1/2}$ decreases from a value greater than one to approximately 0.1. In addition, the effect of averaging on the collapse transition is more apparent at smaller molecular weight.

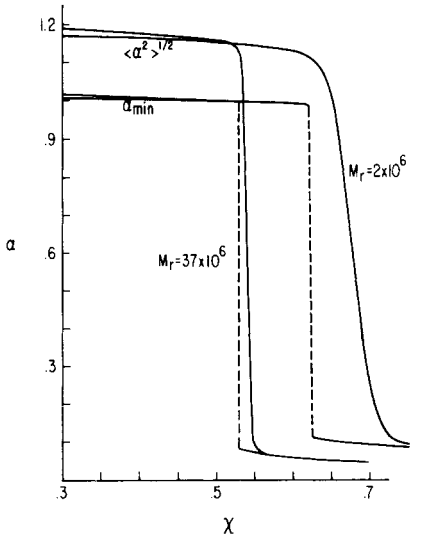

Fig. 3. Value of α, the linear expansion parameter, as a function of χ. Values of $\alpha_{\min}$ correspond to the global minimum in ΔG_{int}. Average values, $\langle \alpha^2 \rangle^{1/2}$, are calculated assuming a Boltzmann distribution. The two left-hand curves were calculated with $M_r = 37 \times 10^6$, $\langle h_0^2 \rangle^{1/2} = 1.3 \times 10^{-4}$ cm, $N = 9300$; the others with $M_r = 2 \times 10^6$, $\langle h_0^2 \rangle^{1/2} = 2.4 \times 10^{-5}$ cm, $N = 500$.

An accurate numerical prediction for the radius of gyration of collapsed DNA cannot be made from the theoretically determined $\langle\alpha\rangle$ for three reasons. First, the Flory model assumes a Gaussian-chain polymer, whereas the best model for DNA is a wormlike chain. Moreover, the densely packed collapsed state of DNA is almost certain to possess some degree of helical organization,[16–18] which is not considered in the theory. Second, since α is the ratio of the radius of gyration to its unperturbed value in a theta solvent, it is necessary to know R_{g_0}. R_{g_0} can be calculated from the persistence length; however, the persistence length for DNA is uncertain, since measured values vary widely.[19–22] Third, the value of χ is unknown. Nevertheless, the rapid change of α on collapse should be observable.

DISCUSSION

Addition of the single-molecule ΔG, determining the equilibrium configuration of an isolated polymer molecule as a result of mixing with solvent molecules, to the Flory-Huggins ΔG, containing the polymer concentration dependence of mixing, leads to a χ-versus-composition diagram that discriminates three states: extended random coils in concentrated solution, extended random coils in dilute solution, and collapsed coils in dilute solution. The first state can be in equilibrium as a precipitate in contact with either of the other two. The collapse of DNA does not occur in the concentrated phase. It is assumed that individual chains remain close to their unperturbed dimensions in the concentrated polymer phase[23] (also Ref. 8, p. 137), since the preference for segment–segment contacts can be satisfied by joining with other chains, thus avoiding the decrease in entropy imposed by the collapsed state.

Effect of Collapse on μ_1

No terms appear in μ_1 [Eq. (4)] that correspond to the internal collapse terms in μ_2 [Eq. (3)]. Although compaction of the polymer reduces μ_2 in dilute solution when χ is greater than $\frac{1}{2}$, the chemical potential of the solvent is practically unaffected by collapse of the polymer. This situation does not contradict the Gibbs-Duhem equation, which is a relationship between the derivatives of the chemical potentials with respect to polymer concentration (i.e., the slope of the chemical potential plotted as a function of v_2). ΔG_{int} has no concentration dependence; therefore, the internal terms change the magnitude of μ_2 in the dilute region but do not alter the slope of the curve of μ_2. (In principle, concentration effects are present in ΔG_{int} arising from interactions between different molecules, but these are not important at the concentrations of concern here.)

Role of ΔG_{int} in the Mixing Free Energy

We wish to examine the extent to which ΔG_{int} and ΔG_{ext} determine the phase equilibrium by considering their relative contributions to μ_2 for the dilute and concentrated phases. Only the terms containing v_2 and α in Eq. (3) are of concern. Within the collapse region of the dilute phase, α is much smaller than one, and the terms from ΔG_{int} are large. Furthermore, for small v_2, the only term from ΔG_{ext} that is significant to μ_2 is the logarithm. As expected, then, in dilute solution the interaction between segments occurs internally (intramolecularly), and there is practically no external interaction between two polymer molecules. For the expanded polymer α is close to one, and the terms from ΔG_{int} do not differ substantially from zero. As v_2 increases, the terms from ΔG_{ext}, and thus the external interactions, become dominant. In conclusion, it is found that ΔG_{int} affects only the composition of the dilute polymer phase, whereas in the DNA-rich phase the theory is essentially that known as the Flory-Huggins theory. We note that the Flory-Huggins formula fits experimental data for polymer solutions surprisingly well and is generally accepted as an empirically useful formula,[9] in spite of the inaccuracies of the lattice model from which it was derived.

Relation to Previous Work

By now there ia a considerable literature on the theory of the collapse transition, most of which is summarized in the excellent review by Lifshitz et al.[3] (This is perhaps the proper place to point out how much of the subsequent work was anticipated in the remarkable paper by Ptitsyn et al.,[15] a paper of which we were unaware at the time of our previous publication.[1]) We limit ourselves here to a few points that do not seem to have been widely studied, or about which there seems to have been some confusion.

Polymer Concentration

Although most previous theoretical work has dealt with single polymer molecules, these studies have limited applicability to real systems composed of many molecules. Our results suggest that the single-molecule collapsed state is thermodynamically stable against aggregation only at high dilution, at least with the simple polymers of the type we have considered. (Presumably a polymer could be constructed with a specific structure that would prevent aggregation, even after collapse; globular proteins must be such structures.) Experimental results of Swislow et al.[4] on polystyrene and our own work on DNA in an accompanying paper[6] confirm this prediction. The experimental phase diagram of Swislow et al. is, in fact, remarkably like our predicted one.

Many of the experimentally observed "collapsed forms" of DNA were found to be aggregates of several molecules[16,17,24,25] (for example, the

striking globule photographed by Lerman[17]). The question arises, Were these forms thermodynamically unstable and trapped in the process of separating into a macroscopic precipitate?

Chain Stiffness

The effect of chain stiffness on the nature of the transition does not always seem to have been appreciated, even though it is implicit in the work of Ptitsyn et al.[15] and in that of de Gennes,[26] and was discussed in our previous paper[1] and in the review by Lifshitz et al.[3] With chains of more than a certain degree of stiffness, that is, chains whose parameter y is less than 0.0227,[1,15] there is a transition between two distinct "states" representing different minima of the free energy. With chains of less than this degree of stiffness there is only a gradual change of the position of a single minimum of the free-energy function as the solvent power or temperature is varied. DNA apparently belongs to the former class and polystyrene, a rather flexible chain, to the latter.

Sanchez[27] studied the theory of the transition at a particular value of y equal to 0.1005 = 19/189 [Ref. 27, Eqs. (51b), (61) and Fig. 2]. One would expect a rather gradual transition at such a high value of y, and this is what he found.

Baumgärtner[28] and Webman et al.[29] have carried out computer simulations of chains of beads with Lennard-Jones interactions and with various degrees of flexibility. Baumgärtner's chains were completely free-jointed, so their degree of flexibility was high, corresponding to $y = 0.5$ or greater; he found a gradual transition, as we would expect. Webman et al. varied the flexibility but still found only a gradual transition. However, the value of y that best fits their results on their least flexible chain is 0.045, which is still greater than the critical value. On the other hand, Ptitsyn et al., simulating chains on a lattice, found a situation where two disconnected regions of configuration space, corresponding to collapsed and expanded chains, respectively, were both substantially occupied in the condensation region.

Order of Phase Transition

Lifshitz et al.[3] classify collapse transitions into two types. If there are two minima in the free-energy function at the transition region, corresponding to stable and metastable states ($y < 0.0227$), they call the collapse transition "first-order." If there is only one minimum in the free energy, and hence no metastable state $y > 0.0227$), they call the transition "second-order." The term "second-order transition" has been used somewhat differently in the thermodynamic literature; Ehrenfest[30] originally defined a first-order transition as one in which an extensive variable, such as the volume, had a discontinuity as a function of an intensive variable, such as pressure, while in a second-order transition a first derivative, such as the

compressibility, had the discontinuity (see also Mayer and Streeter[31]). These definitions are not easy to apply to the collapse transition because the finite size of the molecules causes the transitions to be diffuse, but the so-called first-order collapse transition is clearly analogous to a first-order macroscopic phase transition of a gas to a liquid at low temperature. On the other hand, the "second-order" collapse transition corresponds to the continuous passage from gas to liquid at a temperature above the critical point, a process that is not usually considered to be a phase "transition" in the Ehrenfest sense at all. Obviously there is a problem of nomenclature here that is outside the scope of the present paper to resolve. The important thing from the physical point of view is that the two types of collapse processes proceed by different mechanisms.

We thank Professor John C. Wheeler for many discussions about the theory of phase transitions. This work was supported by a Public Health Service grant, GM-11916, and C.B.P. was the recipient of an IBM Graduate Fellowship.

References

1. Post, C. B. & Zimm, B. H. (1979) *Biopolymers* **18,** 1487–1501.
2. Post, C. B. & Zimm, B. H. (1980) *Biophys. J.* **32,** 448–450.
3. Lifshitz, I. M., Grosberg, A. Yu. & Khokhlov, A. R. (1978) *Rev. Mod. Phys.* **50,** 683–713.
4. Swislow, G., Sun, S. T., Nishio, I. & Tanaka, T. (1980) *Phys. Rev. Lett.* **44,** 796–798.
5. Flory, P. J. (1953) *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, N.Y.
6. Post, C. B. & Zimm, B. H. (1982) *Biopolymers*, **21,** 2139–2160.
7. Yamakawa, H. (1971) *Modern Theory of Polymer Solutions*, Harper & Row, New York.
8. Morawetz, H. (1975) *Macromolecules in Solution*, 2nd ed., John Wiley, New York.
9. Casassa, E. F. (1976) *J. Polym. Sci., Symp.* **54,** 53–83.
10. Flory, P. J. (1970) *Discuss. Faraday Soc.* **49,** 7–29.
11. Koningsveld, R., Stockmayer, W. H., Kennedy, J. W. & Kleintjens, L. A. (1974) *Macromolecules* **7,** 73–79.
12. Cohen, G. & Eisenberg, H. (1968) *Biopolymers* **6,** 1077–1100.
13. Shultz, A. R. & Flory, P. J. (1952) *J. Am. Chem. Soc.* **74,** 4760–4766.
14. Zimm, B. H. (1946) *J. Chem. Phys.* **14,** 164–179.
15. Ptitsyn, O. B., Kron, A. K. & Eizner, Y. Y. (1968) *J. Polym. Sci., Pt. C* **16,** 3509–3517.
16. Shapiro, J. T., Leng, M. & Felsenfeld, G. (1969) *Biochemistry* **8,** 3219–3232.
17. Lerman, L. S. (1973) *Cold Spring Harbor Symp. Quant. Biol.* **38,** 59–73.
18. Weiskopf, M. & Li, H. J. (1977) *Biopolymers* **16,** 669–684.
19. Harrington, R. E. (1978) *Biopolymers* **17,** 919–936.
20. Borochov, N., Eisenberg, H. & Kam, Z. (1981) *Biopolymers* **20,** 231–235.
21. Hagerman, P. J. (1981) *Biopolymers* **20,** 1503–1535.
22. Rizzo, V. & Schellman, J. A. (1981) *Biopolymers* **20,** 2143–2163.
23. Flory, P. J. (1949) *J. Chem. Phys.* **17,** 303–310.
24. Dore, E., Frontali, C. & Gratton, E. (1972) *Biopolymers* **11,** 443–459.
25. Allison, S. A., Herr, J. C. & Schurr, J. M. (1981) *Biopolymers* **20,** 469–488.
26. de Gennes, P. G. (1975) *J. Phys. Lett. (Paris)* **36,** 55–57.

27. Sanchez, I. C. (1979) *Macromolecules* **12,** 980–988.
28. Baumgärtner, A. (1980) *J. Chem. Phys.* **72,** 871–879.
29. Webman, I., Lebowitz, J. L. and Kalos, M. H. (1981) *Macromolecules* **14,** 1495–1501.
30. Ehrenfest, P. (1933) *Leiden Comm. Suppl.* 756.
31. Mayer, J. E. & Streeter, S. F. (1939) *J. Chem. Phys.* **7,** 1019–1025.

Received January 12, 1982
Accepted April 23, 1982