Extension Behavior of Helicogenic Polypeptides

Arnaud Buhot*,†

Theoretical Physics, University of Oxford, 1 Keble Road, Oxford, OX1 3NP, UK

Avraham Halperin*

UMR 5819 (CEA, CNRS, UJF), DRFMC/SI3M, CEA-Grenoble, 17 rue des Martyrs, 38054 Grenoble Cedex 9, France

Received September 17, 2001; Revised Manuscript Received December 6, 2001

ABSTRACT: The force laws governing the extension behavior of homopolypeptides are obtained from a phenomenological free energy capable of describing the helix−coil transition. Just above the melting temperature of the free chains, T^*, the plot of force, f, vs end-to-end distance, R, exhibits two plateaus associated with coexistence of helical and coil domains. The lower plateau is due to tension induced onset of the helix−coil transition. The higher plateau corresponds to the melting of the helices by overextension. Just below T^*, the fR plot exhibits only the upper plateau. The fR plots, the helical fraction, the number of domains and their polydispersity are calculated for two models. In one the helical domains are viewed as rigid rods while in the second they are treated as wormlike chains.

I. Introduction

Single-molecule biomechanical experiments allow to measure the forces generated by biomolecules and their response to applied forces.[1] For "passive" biopolymers, which do not produce active movement against load, the resulting force-extension curves provide a probe of the nonlinear elasticity of the chains and its relationship to configurational changes. Such studies were carried out for single-stranded and double-stranded DNA,[2−5] for the muscle protein Titin,[6−8] the extracellular matrix protein Tenascin,[9] and the polysaccharides Dextran[10] and Xanthan,[11] as well as the synthetic polymer poly(ethylene glycol).[12] Theoretical modeling of the elastic behavior of these polymers is often hampered by lack of detailed knowledge concerning the configurations involved and the relevant interactions. A direct and detailed confrontation between experiment and theory is however possible for the case of homopolypeptides capable of forming an α-helix. In this case the molecular configurations are well understood. Furthermore, one can relate the elastic force law to the known Zimm−Bragg parameters that characterize the helix−coil transition of the free polypeptides. As we shall discuss, the extension force law of a long polypeptide just above the melting temperature of the helix, T^*, exhibits two plateaus. Both are associated with a coexistence of helical and coil domains. The low-tension plateau is due to helix formation induced by the chain extension and the accompanying loss of configurational entropy. The second, high-tension plateau is due to the extension induced melting of the helices when the end-to-end length exceeds the length of the fully helical chain. For $T < T^*$, only the second plateau survives, but the force law for weak extensions exhibits a "quasi-plateau", i.e., a regime with a weak slope that is not associated with a coexistence. This last regime is due to the facile alignment of the long persistent regions in the helix. In the following, we analyze the force laws, their dependence on temperature, and the corresponding population distribution of helical and coil segments.

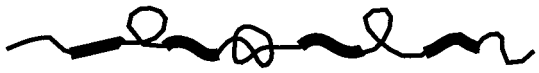

Figure 1. Schematic picture of the stretching of a polypeptide comprising different helical domains (thick lines) and coil domains (thin lines).

While we are unaware of single-molecule experiments on this system, similar behavior was observed experimentally on fibers of collagen during the 1950s.[13]

Two different treatments of this problem were recently advanced.[14,15] In the present article, we present a unified analysis of the problem, tracing the physical origins of the disagreements between the force laws obtained in the two earlier publications and analyzing their range of validity. Our treatment is distinctive in that it is based on free energy argument instead of the customary transfer matrix approach.[16−18] This free energy argument allows us to recover the results of the transfer matrix method but is physically transparent and relatively simple mathematically. Within this approach we consider the extension of a very long homopolypeptide capable of forming α-helices. The discussion is limited to quasi-static extension assuming that the rate of equilibration of the configurations of the chain is much faster than the rate of extension. Two models are explored. In one, the helical domains are viewed as rigid rods while in the second, they are modeled as semiflexible chains. For each model, we present a rigorous derivation of the corresponding force law allowing for the polydispersity of the helical and coil domains and the associated mixing entropy (see Figure 1). In addition, we explore two useful approximations allowing for simplified calculations of the force law. In one approximation, the mixing entropy of the helices and domains is neglected and the plateau is associated with coexistence of one helical domain with one coil domain (see Figure 2). Within this "diblock" or "$S_{mix} = 0$" approximation, the coexistence regimes involve a first-order phase transition. Despite this erroneous conclusion, this approximation correctly identifies the important length and force scales in the problem.[19] The second, mean field approximation, neglects the direct coupling of the distribution of domains

† Current address: UMR 5819 (CEA, CNRS, USF), DRFMC/SI3M, CEA-Grenoble, 17 rue des Martyrs, 38054 Grenoble Cedex 9, France.

10.1021/ma011631w CCC: $22.00
Published on Web 03/09/2002

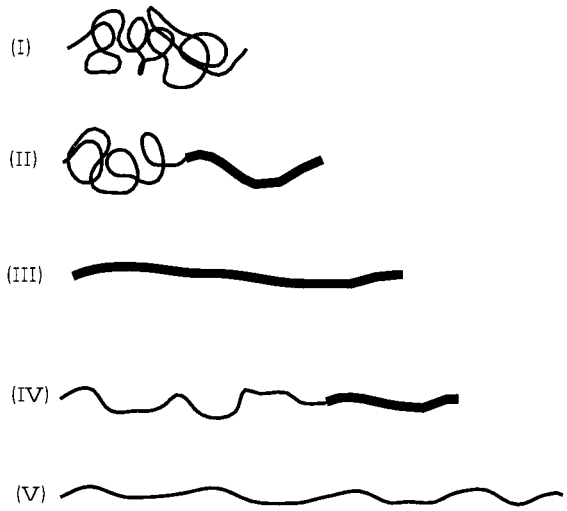

Figure 2. Schematic pictures of the stretching of a polypeptide just above T^* within the $S_{mix} = 0$ approximation. (I) For small extension, the chain is a weakly extended coil, (II) for $r > r_{1L}$, a helical domain coexists with a weakly extended coil domain, (III) at $r = \gamma$, the chain is fully helical, (IV) a helical domain coexists with a strongly extended coil domain for $r > \gamma$, and eventually, (V) for $r > r_{2U}$, the chain is again fully in a coil state but in a highly stretched configuration.

sizes and the applied tension. As we shall see, this approximation is actually exact when the helical domains are modeled as long semiflexible chains.

The problem and the analysis are of interest from a number of perspectives. Homopolypeptides are the simplest representatives of helicogenic molecules. In this case it is not necessary to allow for the sequence heterogeneity which affects the elasticity of heteropolypeptides. Nor is it necessary to consider the long range interaction that occur in multiple helices. As a result, the number of parameters invoked in the theory is smaller and the mathematical analysis is simpler. The force laws obtained are determined by known Zimm−Bragg parameters with no adjustable parameters. Accordingly this system enables a direct comparison between experiment and theory. In this juncture it is important to note the availability of synthetic methods for producing long polypeptides with a well-defined architecture.[20] Furthermore, the formalism described provides a convenient basis for the analysis of more complex helicogenic polymers such as DNA and collagen. From a polymer science point of view the elasticity of homopolypeptides is of interest because it reflects the effects of internal degrees of freedom arising from intrachain self-assembly. The study of these systems enables the exploration of the distinctive nonlinear force laws that are the signatures of such intrachain self-assembly. Finally, this elasticity plays a role in the stabilization of the helical state of grafted polypeptides.[21] In turn this is of importance for the function of fusogenic polypeptides and the design of surfaces that favor spreading and growth of cells.[22−25]

The paper is organized as follows. In section II, we introduce the free energy argument via an analysis of the helix−coil transition in free, undeformed homopolypeptides. A semiquantitative analysis of the effect of extension is presented in section III. Two approaches are used. One focuses on the intersections of the free energy curves of a pure coil and a pure helix as a function of the end-to-end distance, R. As we shall see, each intersection is a rough diagnostic for a transition between the two configurations and for the occurrence of a plateau in the force law. The second, more quantitative approach is the diblock or $S_{mix} = 0$ approximation. Within this approximation the homopolypeptide is assumed to consist of two domains, a helix and a coil, and the mixing entropy associated with the polydispersity of the helical and coil domains is set to zero, $S_{mix} = 0$. Finally, in section IV, we present a rigorous analysis of the force laws allowing for the polydispersity of the coexisting helices and coil domains. As stated earlier, we implement this analysis for two models. In one the helical domains are viewed as rigid rods and in the second as semiflexible chains described by the wormlike chain (WLC) model. In this section we also examine the validity of the mean field approximation introduced in ref 14. Within this approximation, the explicit coupling between the tension and the distribution of domain sizes is ignored. As we shall see, it recovers the exact result when the helices are semiflexible, i.e., when the persistence length of the helices is finite and the helical domains are large. The analysis of the rigid helices model recovers the results obtained by Tamashiro and Pincus using a transfer matrix formalism.[15] A comparison between the two models and an outline of some open problems are presented in the Discussion.

II. Helix-Coil Transition

It is helpful to first summarize the main features of the helix−coil transition of free polypeptides in the absence of imposed extension. We will consider only homopolypeptides, consisting of a single type of amino acid residues, to avoid complications introduced by sequence of different residues found in heteropolypeptides. Our discussion focuses on the case of long chains where the degree of polymerization of the polymer N is large, $N \gg 1$. To this end we first discuss the Zimm−Brag parameters and then introduce the free energy analysis of the transition. As we shall see, this is equivalent to the transfer matrix formalism[26] that is traditionally used to analyze the transition.[16−18]

A. Zimm—Bragg Parameters. Our discussion focuses on helicogenic polypeptides that are capable of assuming two configurations: a random coil and an α-helix. In the coil state the monomers are free to rotate thus leading to a small Kuhn length with typical value of $l_c \simeq 18$ Å corresponding to the length of few monomers. For simplicity we will identify the Kuhn length with the span of a single monomer, $a \simeq 3.8$ Å. The helical configuration involves hydrogen bonds between residues i and $i + 3$. This constrains the orientation of the three intermediate monomers. As a result the persistence length P of the helical configuration is large, of order $P \simeq 2000$ Å. At the same time the projected length of a monomer along the axis of the helix decreases to $a_h \simeq 1.5$ Å. The length of the polypeptide in a helical configuration is thus smaller than the span of a fully extended coil by a factor of $\gamma = a_h/a \simeq 2/5$.[16−18]

The homopolypeptides are modeled as a two-state system. Each monomer can exist either in a helical or a coil state. Choosing the coil state as a reference state, the excess free energy Δf of a monomer within a helical domain reflects two contributions. First is the change in enthalpy upon formation of an intrachain H bond, Δh. The second is associated with the loss of configurational entropy, Δs, in the helical state. Altogether

$$\Delta f \simeq \Delta h - T\Delta s \qquad (1)$$

The enthalpy term favors the helical state while the entropy term favors the coil state. Thus, at low temperatures, the free energy Δf is negative and the helical configuration is preferred while at high temperatures Δf is positive and the coil state is dominant. The two terms are comparable at $T^* \simeq \Delta h/\Delta s$ when $\Delta f \simeq 0$. T^* provides an approximate value for the melting temperature of the α-helix.

An additional parameter is needed to describe the helix−coil transition. The number of H bonds in a helical domain consisting of n monomers is $n - 2$. It is thus necessary to allow for the special state of the terminal monomers of the helical domain. These two monomers lose their configurational entropy with no gain due to the formation of H bonds. Each of the terminal bonds is thus assigned an excess free energy of

$$\Delta f_t \simeq - T\Delta s \tag{2}$$

with respect to the coil state. In the following we will use Δf_t measured with respect to the helical state, that is

$$\Delta f_t \simeq - \Delta h \tag{3}$$

As we shall discuss later, this definition is consistent with the known experimental results. Clearly, this definition is consistent with the traditional one when the coil state is used as a reference. Δf_t plays the role of an interfacial free energy associated with the boundary between helix and coil domains. This term makes the helix−coil transition cooperative since it favors the formation of large domains.

It is customary to describe the helix−coil transition in terms of the Zimm−Bragg parameters[27]

$$s = \exp(-\Delta f/kT) \tag{4}$$

$$\sigma = \exp(-2\Delta f_t/kT) \tag{5}$$

where k is the Boltzmann constant. The T dependent s represents the Boltzmann factor associated with adding one monomer to a helical domain. $s \simeq 1$ at the transition temperature T^* while $s > 1$ for $T < T^*$ and $s < 1$ for $T > T^*$. σ corresponds to the Boltzmann factor associated with the creation of a helical domain. The definition of Δf_t as $- T\Delta s$ suggests that σ is independent of T. The definition of Δf_t as $- \Delta h$ does not predict that σ is independent of T, but it yields an identical estimate for the numerical value of σ.[28] The experimentally measured σ, as obtained from plots of the helical content vs T, are in the range $10^{-3}-10^{-4}$, depending on the residue, and exhibit a weak T dependence. As stated earlier, in the following we will use $\sigma = \exp(2\Delta h/kT)$. As we shall discuss, a plot of the helical content vs s exhibits a sigmoid behavior and the width of the transition scales with $T^* \sigma^{1/2}$. Thus, the transition becomes sharper, more cooperative, as σ decreases.

B. The Helix—Coil Transition: A Free Energy Approach. Having defined the relevant molecular parameters, Δf and Δf_t or equivalently s and σ, we are in a position to analyze the helix−coil transition as it occurs in free chains upon changing the temperature T. The minimal description of the transition requires the specification of two properties, the total number of monomers in a helical configuration, $N\theta$ and the number of helical domains, Ny. While θ and y are sufficient for a discussion of the helix−coil transition of free chains, the analysis of the effect of extension requires additional information. In particular, the probability distribution of helical domains, $P_h(n)$, and of coil domains, $P_c(n)$, comprising of n monomers. To determine these parameters we minimize the appropriate phenomenological free energy. This method recovers the results obtained by the transfer matrix approach and is easily generalized to allow for the effect of applied tension.

The free energy allows for two principal contributions: (i) The first is the free energy of the monomers. All $N\theta$ helical monomers, including the terminal monomers of the domains are assigned an excess free energy of Δf. The $2Ny$ terminal monomers are assigned an additional free energy of $\Delta f_t = - \Delta h$. As a result their excess free energy with respect to the coil state is $\Delta f_t = - T\Delta s$. (ii) The second is a mixing entropy term associated with the different possible placements of $N\theta$ helical monomers and $2Ny$ terminal monomers constituting the domain boundaries. Since the helical and coil domains alternate, there is no entropy due to their mixing. Rather, the mixing entropy arises because of the polydispersity of the domains; i.e., helical and coil domains of different size n are considered distinguishable. Since there are Ny helical and Ny coil domains, the mixing entropy is

$$S_{mix} = - Nky\sum_{n=1}^{\infty}\{P_h(n) \ln P_h(n) + P_c(n)\ln P_c(n)\} \tag{6}$$

The free energy per monomer in the unperturbed chain is thus

$$F_0/NkT = -\theta \ln s - y \ln \sigma + y\sum_{n=1}^{\infty}P_h(n) \ln P_h(n) + y\sum_{n=1}^{\infty}P_c(n)\ln P_c(n) - \mu_1^h(\sum_{n=1}^{\infty}P_h(n) - 1) - \mu_1^c(\sum_{n=1}^{\infty}P_c(n) - 1) - \mu_2^h\left(\sum_{n=1}^{\infty}nP_h(n) - \frac{\theta}{y}\right) - \mu_2^c\left(\sum_{n=1}^{\infty}nP_c(n) - \frac{1-\theta}{y}\right) \tag{7}$$

The Lagrange multipliers, μ_1^h and μ_1^c, ensure the normalization $\sum_{n=1}^{\infty}P_h(n) = 1$ and $\sum_{n=1}^{\infty}P_c(n) = 1$. The two remaining Lagrange multipliers, μ_2^h and μ_2^c, impose the average sizes $k_h = \theta/y$ and $k_c = (1 - \theta)/y$ of the helical and coil domains. This description corresponds to a simplified version of the transfer matrix method involving a two by two matrix. In both cases, there is no constraint ensuring that helical domains incorporate at least three monomers.[29] Since we focus on the case of $\sigma \ll 1$ and $N \gg 1$, the weight of small domains is negligible and this description yields the correct results.

The probabilities $P_h(n)$ and $P_c(n)$ are determined by minimization of F_0 with respect to $P_h(n)$ and $P_c(n)$ (Appendix A) leading to

$$P_h(n) = \frac{y}{\theta - y}\left(\frac{\theta - y}{\theta}\right)^n \tag{8}$$

$$P_c(n) = \frac{y}{1 - \theta - y}\left(\frac{1 - \theta - y}{1 - \theta}\right)^n \tag{9}$$

Note that $P_h(n)$ and $P_c(n)$ decay exponentially with decay constants $\lambda_h = - 1/\ln(1 - y/\theta)$ and $\lambda_c = - 1/\ln[1 - y/(1 - \theta)]$. Thus, $P_h(n)$ and $P_c(n)$ are not sharply peaked at the mean values $k_h = \theta/y$ and $k_c = (1 - \theta)/y$. In particular, there is a significant population of small helices even when the average size of the helical domain

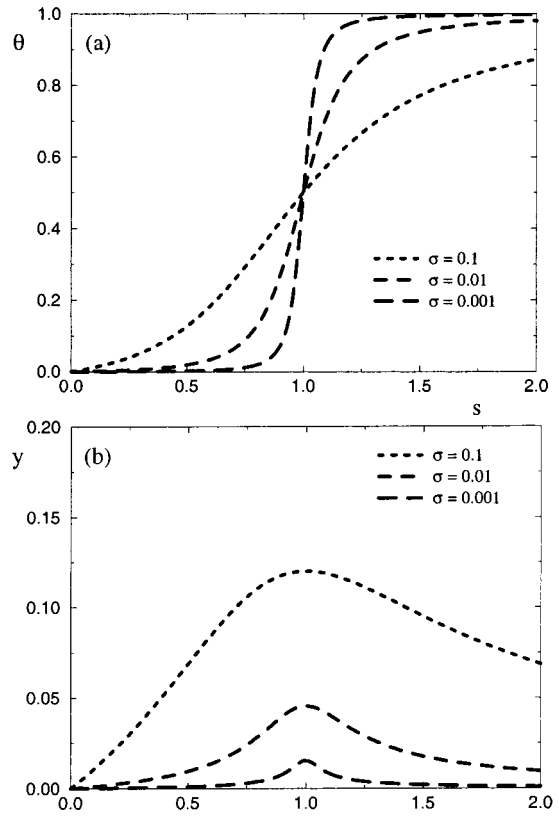

Figure 3. Helical fraction θ (a) and domain fraction y (b) plotted as a function of s for long polypeptides free of tension. In all cases, the helix−coil transition occurs in the vicinity of $s = 1$; however, the width of the transition region decreases with σ.

$k_h = \theta/y$ is large. In this limit, $\lambda_h \simeq k_h$. Similarly, when $k_c = (1 - \theta)/y$ is large, $\lambda_c \simeq k_c$.

Upon substituting $P_h(n)$ and $P_c(n)$ in F_0, as given by eq 7, we obtain F_0 as a function of θ and y only:[30]

$$\frac{F_0}{NkT} = -\theta \ln s - y \ln \sigma + (\theta - y) \ln \frac{\theta - y}{\theta} + y \ln \frac{y}{\theta} + (1 - \theta - y) \ln \frac{1 - \theta - y}{1 - \theta} + y \ln \frac{y}{1 - \theta} \quad (10)$$

The equilibrium conditions $\partial F_0/\partial\theta = 0$ and $\partial F_0/\partial y = 0$ then lead to

$$(1 - \theta)(\theta - y) = s\,\theta(1 - \theta - y) \quad (11)$$

$$y^2 = \sigma\,(\theta - y)(1 - \theta - y) \quad (12)$$

In turn, these yield (Appendix B) the expression for the equilibrium values of θ and y

$$\theta = \frac{1}{2} + \frac{1}{2}\frac{s - 1}{\sqrt{(s - 1)^2 + 4\,\sigma s}} \quad (13)$$

$$y = \frac{2\,\sigma s[(s - 1)^2 + 4\,\sigma s]^{-1/2}}{s + 1 + \sqrt{(s - 1)^2 + 4\,\sigma s}} \quad (14)$$

These expressions are identical to the ones obtained via the transfer matrix method.[18]

The behavior of θ and y as a function of s and σ is plotted in Figure 3. When $\sigma = 1$ we recover the Boltzmann distribution for independent two-level systems: $\theta = s/(1 + s)$. In marked contrast when $\sigma = 0$, the fraction of helical monomers is steplike: $\theta = 1$ for $s > 1$ and $\theta = 0$ for $s < 1$. The chain is either a pure helix or a pure coil. For intermediate values of σ, the plot of θ vs s yields a sigmoid curve with a transition region of width $\sigma^{1/2}$ centered around $s = 1$. The number of domains is maximal in the vicinity of $s \simeq 1$, and it decreases with σ. The parameter $s = \exp(-\Delta f/kT)$ serves as a measure of the temperature, T. For low temperatures, $s \gg 1$, and the polymer is mainly in a helical state. At high temperatures, $s \ll 1$, and the polymer is mostly in a coil state. The crossover occurs at $T^* = \Delta h/\Delta s$ for which $s = 1$. Near this crossover, $s - 1 \sim T^* - T$ and the width of the crossover regime is $\Delta T \sim T^*\,\sigma^{1/2}$.

III. Extension Behavior: Qualitative Analysis

The most notable feature of the force laws of homopolypeptides are the plateaus. In this section, we present qualitative and semiquantitative analysis of the plateaus. Each plateau in the force law corresponds to a coexistence between two states, a helix and a coil. The situation is somewhat analogous to a liquid−gas transition which leads to a plateau in the pressure−volume diagram. In our case, the tension f plays the role of the pressure and the end-to-end distance R is analogous to the volume. This analogy is however of limited validity since the homopolypeptides are one-dimensional systems with short-range interactions and thus incapable of undergoing a first-order phase transition. In such systems the boundary free energy is independent of the size of the domain. As a result, a first-order phase transition involving a coexistence of two domains is prevented by the mixing entropy.[31] Accordingly, the average domain sizes $k_c = (1 - \theta)/y$ and $k_h = \theta/y$, as obtained in the previous section, are independent of the size of the chain, N. Nevertheless, the main features of the system can be recovered if one ignores the role of the mixing entropy. As expected, within this rough approximation the plateaus are erroneously associated with first-order phase transitions. Despite this deficiency, this approximation recovers the correct length and force scales, but the force laws obtained are wrong in detail. Before we present this approximation we investigate the crossing of the free energy curves associated with the "pure" helix and coil states. This is a rough diagnostic for first-order phase transitions. In our case it provides an insight concerning the physical origin of the plateaus as well as a reasonable estimate for some of the length and force scales involved.

A. Free Energy Curve Crossing. The occurrence of the plateaus is signaled by the crossing of the free energy curves of the pure helix and the pure coil as a function of R. Each crossing point signals a coexistence between the two "phases" and thus a plateau in the force law.

With the unperturbed coil as a reference state, the free energy of the coil is simply its elastic free energy. In this section, we will utilize the fixed R ensemble and we thus denote the elastic free energy by $F_{el}(R)$. To allow for the finite extensibility of the chain, we use the freely jointed chain (FJC) model (Appendix C) for $F_{el}(R)$, and the free energy per monomer is

$$\frac{F_{el}(R)}{NkT} = rx - \mathcal{L}_{int}(x) \quad (15)$$

where $\mathcal{L}_{int}(x) = \ln[\sinh(x)/x]$, $r = R/Na$ is the reduced end-to-end distance of the polymer, and $x = fa/kT$ is the

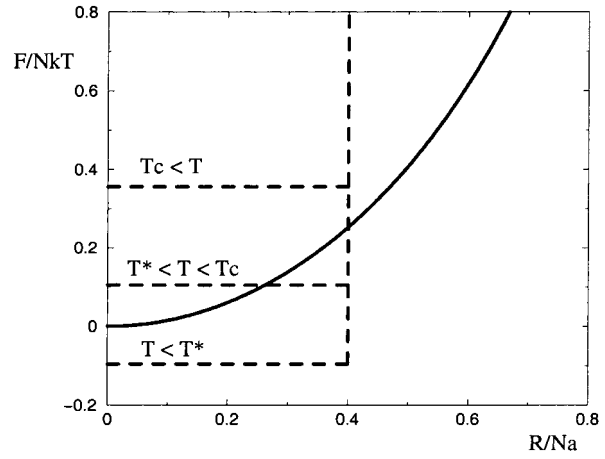

Figure 4. Occurrence of plateaus in the extension force laws signaled by crossing of the free energy curves corresponding to a polypeptide in a pure coil form (continuous curve) and in a fully helical form (dashed lines).

reduced force. r and x are related via $r = \mathcal{L}(x) \equiv \coth(x) - 1/x$ where $\mathcal{L}(x)$ is the Langevin function.

The free energy of the helix reflects two contributions. First is the excess free energy of a helix, $F_h = N\Delta f + 2\Delta f_t$. In the limit of an infinite chain, when the contribution of the terminal monomers is negligible, the free energy per monomer is $F_h/NkT = -\ln s$. At this stage we assume that the persistence length of the helix is infinite and that it may be viewed as a perfectly rigid and unextendable rod. Within this view, extending the helix beyond its natural length is associated with an infinite free energy penalty. This penalty is important because the helix is shorter than the fully extended coil. As noted earlier, the contribution of each helical monomer to the overall length of the helix is $a_h = \gamma a < a$. Altogether, $F_h/NkT = -\ln s$ so long as $r = R/Na \leq \gamma$ while in the range $r > \gamma$ it is $F_h/NkT = \infty$.

Three scenarios are now possible (Figure 4). At low temperatures, $T < T^*$ and $s > 1$. Accordingly, $F_h/NkT < 0$ while $r \leq \gamma$. In this regime the helix is more stable than the coil. When $r > \gamma$ the free energy of the helix diverges and the coil state becomes preferable. This curve crossing corresponds to a plateau involving the breakdown of the helix due to the applied tension. Above the melting temperature, for $T \gtrsim T^*$ and $F_h/NkT = -\ln s \gtrsim 0$ the unstretched polypeptide is in a coil state. Upon stretching F_{el}/NkT increases and eventually intersects with F_h/NkT. This intersection signals a plateau in the force law due to the formation of helical domains. At $r = \gamma$ the free energies F_{el}/NkT and F_h/NkT cross again because of the infinite free energy of a helix with $r > \gamma$. This second intersection signals a second plateau involving the melting of the helix and the formation of coil domains. Finally, for sufficiently high temperatures, $T > T_c$, no intersections occur and the polypeptide remains in a coil state for all r values. The threshold temperature for this regime, T_c, is defined by the equality $F_{el}(N\gamma a) = F_h$ where $F_{el}(N\gamma a)$ is the free energy of a stretched coil with $r = \gamma$.

B. The $S_{mix} = 0$ approximation. The intersection of the free energy curves provides a rough indicator for the occurrence of plateaus in the force law. However, it does not allow one to calculate an explicit force-extension diagram. A simple approach yielding the main features of the force law is the $S_{mix} = 0$ or diblock approximation.[14,32] Within this approximation the force law is calculated for a diblock copolymer with one helical domain and one coil domain. The size of the annealed domains varies with the extension and is determined by the equilibrium conditions. As is implied by its name, in this approximation S_{mix} is neglected. As expected, setting $S_{mix} = 0$ leads to a first-order phase transition associated with a coexistence of two "phases". Consequently, some features of the force law are wrong. In particular, this approximation yields a perfect plateau with $f(R) = \text{const}'$ while in reality f increases weakly with R. These features notwithstanding, this simple approximation does recover the correct length and force scales. Within this discussion, we will also briefly examine the effect of the persistence length of the helical domains. In particular, we compare the behavior of a helix endowed with an infinite persistence length to that of a helix with a finite one.

Within the $S_{mix} = 0$ or diblock approximation, the free energy F_{chain} of the chain is

$$F_{chain} = N\theta\Delta f + 2\Delta f_t + F_{el}(\theta, \gamma, R) \qquad (16)$$

Here, the first term allows for the excess free energy of the helical monomers. Since $S_{mix} = 0$, the helical and coil phases separate into two domains. The "interfacial" free energy of the single helical domain gives rise to the second, constant term. These two terms are supplemented by an elastic free energy F_{el} that depends on θ, γ and the end-to-end distance, R. Different versions of the $S_{mix} = 0$ are possible, depending on the choice of F_{el}. In the following, we focus on the simplest version where F_{el} is given by the Gaussian elasticity of an ideal random coil in the fixed R ensemble. Two additional assumptions are invoked: (i) The helical domain is perfectly rigid and unextendable and (ii) the helical domain is perfectly aligned with the applied force. These assumptions specify the elastic free energy per monomer

$$\frac{F_{el}}{NkT} = \frac{3(r - \gamma\theta)^2}{2 \quad 1 - \theta} \qquad (17)$$

This simple version of the diblock approximation permits an explicit analytical solution for the force law. The corresponding equilibrium condition is

$$\frac{\partial F_{chain}}{\partial \theta} = 0 = \frac{3}{2}\frac{(r - \gamma\theta)^2}{(1 - \theta)^2} - 3\gamma\frac{r - \gamma\theta}{1 - \theta} + \frac{\Delta f}{kT} \qquad (18)$$

Within this approximation, the equilibrium condition $\partial F_{chain}/\partial y = 0$ is an identity yielding no additional information. The solution of this quadratic equation in $X = (r - \gamma\theta)/(1 - \theta)$ is $X_\pm = \gamma(1 \pm \sqrt{1 - 2\Delta f/3\gamma^2 kT})$. This defines a critical Δf

$$\Delta f_c/kT_c = 3\gamma^2/2 \qquad (19)$$

corresponding to a critical $s_c = \exp(-\Delta f_c/kT_c)$ or T_c. While $\Delta f > \Delta f_c$, the equilibrium condition cannot be satisfied; that is, the polypeptide remains in a coil state irrespective of the imposed tension. A critical point occurs for $\Delta f = \Delta f_c$. When $0 < \Delta f < \Delta f_c$ two solutions exist, each defining a plateau in the force vs extension curve. The solution X_- corresponds to the lower plateau, obtained for $r < \gamma$, where

$$\theta = 1 - \frac{1 - r/\gamma}{\sqrt{1 - \Delta f/\Delta f_c}} \qquad (20)$$

This first regime disappears for $\Delta f < 0$ or $T < T^*$. The condition $\theta = 0$ defines a lower boundary of the plateau

$$r_{1L} = \gamma(1 - \sqrt{1 - \Delta f/\Delta f_c}) \tag{21}$$

r_{1L} increases as Δf increases until $r_{1L} = \gamma$ for $\Delta f = \Delta f_c$. For $r < r_{1L}$, the fraction of helices is $\theta = 0$ and the chain is in a purely coil form. The upper boundary of this plateau occurs at

$$r_{1U} = \gamma \tag{22}$$

when the chain is fully helical, $\theta = 1$. At intermediate extensions, $r_{1L} \le r \le r_{1U}$ the fraction of helical bonds increases as

$$\theta = \frac{r - r_{1L}}{r_{1U} - r_{1L}} \tag{23}$$

and $\theta = {}^1/_2$ occurs at the midpoint of the plateau $r_1(\theta = {}^1/_2) = (r_{1U} + r_{1L})/2$. The equilibrium free energy of the chain for $r < r_{1L}$ is $F_{chain}/NkT = 3r^2/2$ while for $r_{1L} \le r \le r_{1U}$ it is

$$\frac{F_{chain}}{NkT} = \frac{3r_{1L}^2}{2}\frac{r_{1U} - r}{r_{1U} - r_{1L}} + \frac{\Delta f}{kT}\frac{r - r_{1L}}{r_{1U} - r_{1L}} \tag{24}$$

The second plateau is described by the solution, X_+, leading to

$$\theta = 1 - \frac{r/\gamma - 1}{\sqrt{1 - \Delta f/\Delta f_c}} \tag{25}$$

The lower boundary of this plateau occurs at

$$r_{2L} = r_{1U} = \gamma \tag{26}$$

when $\theta = 1$. The upper boundary, when $\theta = 0$, occurs at

$$r_{2U} = \gamma(1 + \sqrt{1 - \Delta f/\Delta f_c}) \tag{27}$$

For $r_{2L} \le r \le r_{2U}$, the fraction of helical monomers decreases as

$$\theta = \frac{r_{2U} - r}{r_{2U} - r_{2L}} \tag{28}$$

and $\theta = {}^1/_2$ occurs at the midpoint of the plateau $r_2(\theta = {}^1/_2) = (r_{2U} + r_{2L})/2$. Notice that the width of the two plateaus shrinks as $\Delta f \to \Delta f_c$ or $T \to T_c$. For $r \ge r_{2U}$ the fraction of monomers in a helical state is again $\theta = 0$ and the corresponding free energy is $F_{chain}/NkT = 3r^2/2$. The equilibrium free energy of the chain in the range $r_{2L} \le r \le r_{2U}$ is

$$\frac{F_{chain}}{NkT} = \frac{3r_{2U}^2}{2}\frac{r - r_{2L}}{r_{2U} - r_{2L}} + \frac{\Delta f}{kT}\frac{r_{2U} - r}{r_{2U} - r_{2L}} \tag{29}$$

Note that only the lower plateau is obtained if $\gamma = 1$ is assumed.[14]

The $S_{mix} = 0$ approximation allowed us to obtain rough force laws corresponding to the three scenarios identified by the curve crossing argument. Thus, for $\Delta f > \Delta f_c$ or $T > T_c$, the polypeptide remains in a coil configuration for the whole range of extension. The "two plateau scenario" occurs for $0 < \Delta f < \Delta f_c$ corresponding to $s_c < s < 1$ or $T^* < T < T_c$. In this case we can distinguish five regimes (Figure 2).

(I) For low extensions the chain is in a coil state, $\theta = 0$, and the elastic behavior is Gaussian leading to $fa/kT = 3r$.

(II) In the range $r_{1L} \le r \le r_{1U}$, helix and coil domains coexist according to the lever rule $\theta/(1 - \theta) = (r - r_{1L})/(r_{1U} - r)$. This coexistence is associated with a plateau with a constant tension

$$\frac{f_{1co}a}{kT} = 3r_{1L} \tag{30}$$

(III) At $r_{1U} = r_{2L} = \gamma$ the force curve exhibits a discontinuity associated with a steplike increase to a second plateau. The magnitude of the step, δf is

$$\delta f\, a/kT = 6\gamma\sqrt{1 - \Delta f/\Delta f_c} \tag{31}$$

(IV) In the range $r_{2L} \le r \le r_{2U}$, helix and coil domains coexist according to the lever rule $\theta/(1 - \theta) = (r_{2U} - r)/(r - r_{2L})$. This coexistence is associated with a second plateau with a constant tension

$$\frac{f_{2co}a}{kT} = 3r_{2U} \tag{32}$$

(V) Finally, for $r > r_{2U}$, the chain is in a strongly stretched coil state obeying $fa/kT = 3r$.

When $\Delta f < 0$, or $T < T^*$, the unperturbed chain is in a helical state. Consequently, the lower plateau disappears. It is replaced by a "quasi-plateau" which is not associated with a helix−coil coexistence. Rather, it is due to the facile orientation of the helical chain which is modeled as a rigid rod. At $r = \gamma$, a jump occurs to the upper plateau.

The analysis presented above utilized the Gaussian elastic free energy in the fixed R ensemble to describe the coil. While this allows us to obtain explicit expressions for all the characteristics of the force curve, it also introduces errors. Clearly, this choice of F_{el} does not capture the behavior of the coil in the high extension regime, where finite extensibility begins to play a role. As a result, the values of r_{2U} are overestimated. This has two consequences. First, within this approximation the span of the two plateaus are equal $r_{1U} - r_{1L} = r_{2U} - r_{2L} = \gamma\sqrt{1-\Delta f/\Delta f_c}$ while for a more realistic choice of F_{el} the higher plateau is narrower. A second more important pathology is the disappearance of regime (V) in the $\Delta f < 0$ case.

To allow for the finite extensibility, it is necessary to use the freely jointed chain model (FJC) for the coil segment.[33] It is important to note that so long as $\gamma < 0.5$ and $\Delta f > 0$ the discrepancies between the two schemes are rather small (Figure 5). Another ingredient of our simplified analysis is the modeling of the helices as rigid and unextendable rods. An alternative description of the elastic properties of helices models them as wormlike chains, i.e., semiflexible chains capable of undergoing undulations (see Appendix D).[34] As we shall see later, this semiflexible approximation of the helices has a qualitative effect on the results of the rigorous analysis of the system. However, in the context of the $S_{mix} = 0$ approximation for $\Delta f > 0$, it mainly affects the transition between the two plateaus (Figure 5). When the helix is modeled as a rigid rod the transition occurs

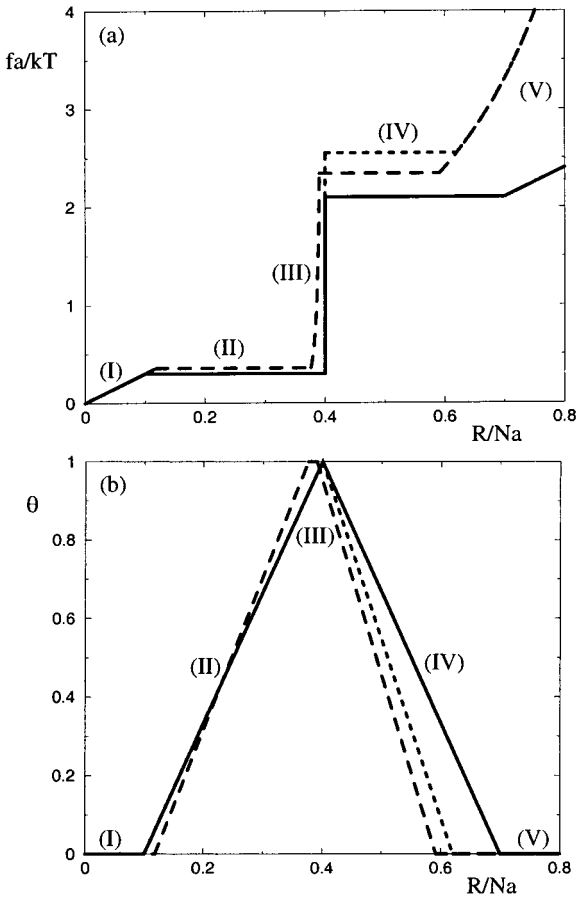

Figure 5. Plots of the reduced force $x = fa/kT$ (a) and the helical fraction θ (b) vs the reduced end-to-end distance, $r = R/Na$, as obtained via the $S_{mix} = 0$ approximation for the $s_c < s < 1$ scenario. Both figures correspond to $s = 0.9$, $\gamma = {}^2/_5$, and $n_h = 1000$. The continuous curve corresponds to the case of rodlike helices and Gaussian coils. The dashed lines depict the results obtained when the coils are described by the FJC model and the helices as rigid rods (short dashed line) or by the WLC model (long dashed line). The roman numerals identify the different regimes as depicted in Figure 2.

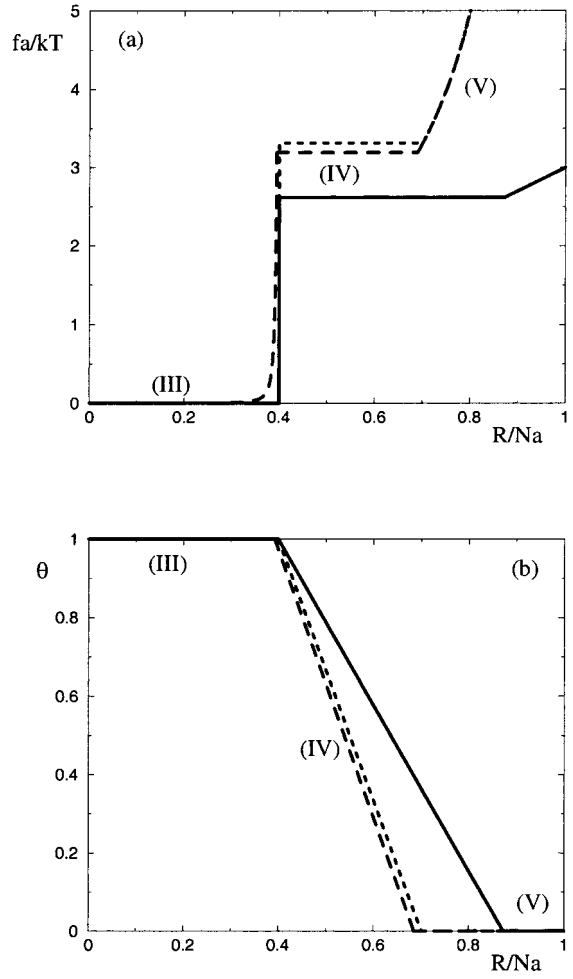

Figure 6. Same as previous figures for the $s > 1$ scenario. Both figures correspond to $s = 1.1$, $\gamma = {}^2/_5$, and $n_h = 1000$.

as a jump that takes place at $r = \gamma$. On the other hand, when the WLC model is used, the jump, while abrupt, takes place over a finite interval; i.e., $r_{1U} = r_{2L}$ is replaced by $r_{1U} < r_{2L} < \gamma$. For the $\Delta f < 0$ (Figure 6) the transition at the upper boundary of the "quasi-plateau" is more gradual when the elasticity of the semiflexible helix is allowed for. Finally, note that the diblock approximation is recovered from the more rigorous treatments presented below when the $N \to \infty$ limit is taken before the $\sigma \to 0$ limit. When the $\sigma \to 0$ limit is approached for a finite chain it leads to a single domain.

IV. Rigorous Analysis

The analysis presented in the previous section fails to allow for a number of important ingredients. The most important is the mixing free energy, $-TS_{mix}$, associated with the polydispersity of the domains. In this section we present an analysis that allows for the role of this mixing entropy. Because of S_{mix} the helix−coil transitions no longer take place as first-order phase transitions. In turn this has two important consequences: (i) The "perfect" plateaus, with constant forces f_{1co} and f_{2co}, are replaced by "imperfect" plateaus with a finite slope. (ii) The crossovers between the plateaus and the neighboring regimes become smoother.

The analysis is based on a generalization of the approach described in section II that is, all chain characteristics, $P_h(n)$, $P_c(n)$, θ, and y are obtained from the minimization of the appropriate free energy i.e., the equilibrium conditions. As in section III, the free energy is of the form

$$F_{chain} = F_0 + F_{el} \tag{33}$$

where, F_0 is the free energy of the unperturbed free chain and F_{el} accounts for the elasticity of the chain. However, in the following we utilize different expressions for these two terms. Since we aim to obtain $P_h(n)$ and $P_c(n)$ it is necessary to utilize F_0 as given by eq 7. The handling of the elastic free energy is significantly different. In section III, we used the Gaussian elasticity to describe the deformation behavior of the coils and assumed that the helices contribute to the elastic behavior as a geometric constraint. While this rough description is sufficient at the level of the $S_{mix} = 0$ approximation, a more refined treatment is needed for the detailed analysis to be undertaken below. In particular, it is necessary to allow for two additional features: (i) The finite extensibility of the chain and (ii) the gradual alignment of the helical domains with the applied tension. The freely jointed chain (FJC) model allows us to incorporate the finite extensibility of the coil domains. With regard to the helices we will explore two different descriptions that impose finite extensibility on the helices. First, we will model the helical

domains as rigid and unextendable rods and allow for their length dependent alignment by using the FJC model; i.e., the helical domains are considered as freely jointed rigid rods forming a chain. This amounts to the assumption of an infinite persistence length for each helical domain. This picture is reasonable so long as the size of the helical domains is small compared to the persistence length, P. In the second approach, each of the helical domains is viewed as wormlike chain (WLC) that is, semiflexible chain characterized by a finite persistence length and capable of undergoing undulations. It is assumed that the domains are large in comparison to P. An important difference between the FJC and the WLC models is that in the later the persistence length depends on the applied tension. Our approach utilizes the analysis of Marko and Siggia of the extension elasticity of the WLC.[35] The utilization of these models makes the constant f ensemble mathematically convenient. This choice of ensemble is also preferable because it corresponds better to the experimental situation encountered in single-molecule measurements. In each case, we first obtain the equilibrium values of $P_h(n)$, $P_c(n)$, θ, and y of a chain subject to a given tension, f, by minimizing $F_{chain}(f)$ with respect to $P_h(n)$, $P_c(n)$, θ, and y. This allows us to determine the equilibrium $F_{chain}(f)$ thus enabling us to calculate the fR diagram by using

$$R = -\partial F_{chain}(f)/\partial f \tag{34}$$

A. Rigid Helices. When the helical domains are viewed as rigid rods and the FJC is invoked, the reduced end-to-end distance $r = R/Na$ of the polypeptide is

$$r = y\sum_{n=1}^{\infty} P_c(n)\, n\, \mathcal{L}(x) + y\sum_{n=1}^{\infty} P_h(n)\, n\, \gamma\, \mathcal{L}(n\gamma x) \tag{35}$$

where $x = fa/kT$ is the reduced tension. The first part of this expression corresponds to the projection of the coil monomers along the direction of the tension. The second part reflects to the contribution of the helical domains to r. Because the individual monomers within the coil domains are aligned independently, the domain size does not play a role. The extension of each coil domain is extensive with respect to the number of monomers, n. Consequently, this contribution may be expressed in terms of the first moment of the probability distribution $k_c = (1 - \theta)/y = \sum_n nP_c(n)$. As a result the full probability $P_c(n)$ is irrelevant. In marked contrast, the monomers within the helical domains are not aligned individually but as part of the rodlike domains. The nonlinearity introduced by the $\mathcal{L}(n\gamma x)$ term in the expression for r couples f and $P_h(n)$. As a consequence, $P_h(n)$ will be explicitly modified by the applied tension while $P_c(n)$ depends on f only implicitly, via θ and y.

Eliminating $P_c(n)$ from eq 35, we obtain

$$r = (1 - \theta)\mathcal{L}(x) + y\sum_{n=1}^{\infty} P_h(n)\, n\gamma\, \mathcal{L}(n\gamma x) \tag{36}$$

This expression for $R(f)$ enables us to obtain the corresponding elastic free energy $F_{el}(f) = -\int_0^f R\, df'$ (Appendix C)

$$F_{el}/NkT = -(1 - \theta)\mathcal{L}_{int}(x) - y\sum_{n=1}^{\infty} P_h(n)\, \mathcal{L}_{int}(n\gamma x) \tag{37}$$

where $\mathcal{L}_{int}(x) = \ln[\sinh(x)/x]$.

Altogether the free energy per monomer is thus

$$F_{chain}/NkT = -\theta \ln s - y \ln \sigma - (1 - \theta)\mathcal{L}_{int}(x) - y\sum_{n=1}^{\infty} P_h(n)\, \mathcal{L}_{int}(n\gamma x) + y\sum_{n=1}^{\infty} P_h(n) \ln P_h(n) + y\sum_{n=1}^{\infty} P_c(n) \ln P_c(n) - \mu_1^h\Big(\sum_{n=1}^{\infty} P_h(n) - 1\Big) - \mu_1^c\Big(\sum_{n=1}^{\infty} P_c(n) - 1\Big) - \mu_2^h\Big(\sum_{n=1}^{\infty} nP_h(n) - \frac{\theta}{y}\Big) - \mu_2^c\Big(\sum_{n=1}^{\infty} nP_c(n) - \frac{1 - \theta}{y}\Big) \tag{38}$$

This free energy reduces to the one given by eq 7 in the case of $x = 0$ thus ensuring the recovery of the known features of the helix−coil transition in an unperturbed chain. Since the f dependence of F_{chain} is due only to F_{el}, eq 34 yields R as given by eq 36. As before, $P_h(n)$ and $P_c(n)$ are determined by minimizing $F_{chain}(f)$ with respect to these probabilities subject to two constraints: (i) the normalization of the probabilities $P_h(n)$ and $P_c(n)$ and (ii) the average sizes $k_c = (1 - \theta)/y$ and $k_h = \theta/y$ of the coil and helical domains. Upon substituting the resulting $P_h(n)$ and $P_c(n)$ in F_{chain}, we obtain $F_{chain}(f)$ as a function of θ and y. θ and y are then determined by the equilibrium conditions $\partial F_{chain}/\partial\theta = F_{chain}/\partial y = 0$. The details of the calculation involved are described in Appendix E. θ and y for a given tension f are

$$\theta = \frac{\sigma(1 - \theta - y)\, sA \sinh(\gamma x)}{\gamma x(1 - sAe^{\gamma x})(1 - sAe^{-\gamma x})} \tag{39}$$

$$y = \frac{\sigma(1 - \theta - y)}{2\gamma x} \ln\Big(\frac{1 - sAe^{-\gamma x}}{1 - sAe^{\gamma x}}\Big) \tag{40}$$

where

$$A(x,\theta, y) = \frac{(1 - \theta - y)\, x}{(1 - \theta)\sinh(x)} \tag{41}$$

The probabilities $P_c(n)$ and $P_h(n)$ are

$$P_c(n) = \frac{y}{1 - \theta - y}\Big(\frac{1 - \theta - y}{1 - \theta}\Big)^n \tag{42}$$

$$P_h(n) = \frac{\sigma}{y}(1 - \theta - y)\,(sA)^n \frac{\sinh(n\gamma x)}{n\gamma x} \tag{43}$$

$P_h(n)$, as $P_c(n)$, is essentially an exponentially decaying function. However, the characteristic decay constants are now dependent on f. The dependence is explicit for $P_h(n)$ while for $P_c(n)$ the f dependence is implicit, via θ and y.

In Figure 7, the equilibrium θ and y are plotted as functions of $x = fa/kT$ for different values of σ. Clearly, the results of the $S_{mix} = 0$ or diblock approximation are approached in the limit of strong cooperativity, $\sigma \to 0$. In particular, y approaches zero while θ exhibits a steplike behavior. Thus, $\theta \simeq 0$ outside the plateaus range, $x < x_{1co} = 0.30$, and $x > x_{2co} = 2.55$, while within the plateaus range, $x_{1co} < x < x_{2co}$ we find $\theta \simeq 1$ and $y \simeq 0$. Consequently, the mean size of the helical domains, $k_h = \theta/y$, becomes very large compared to P. In turn, this suggests that the validity of the model of helical domains as rigid rods is questionable. Clearly, the rigid helices model is meaningful only while the persistence length is much larger than k_h.

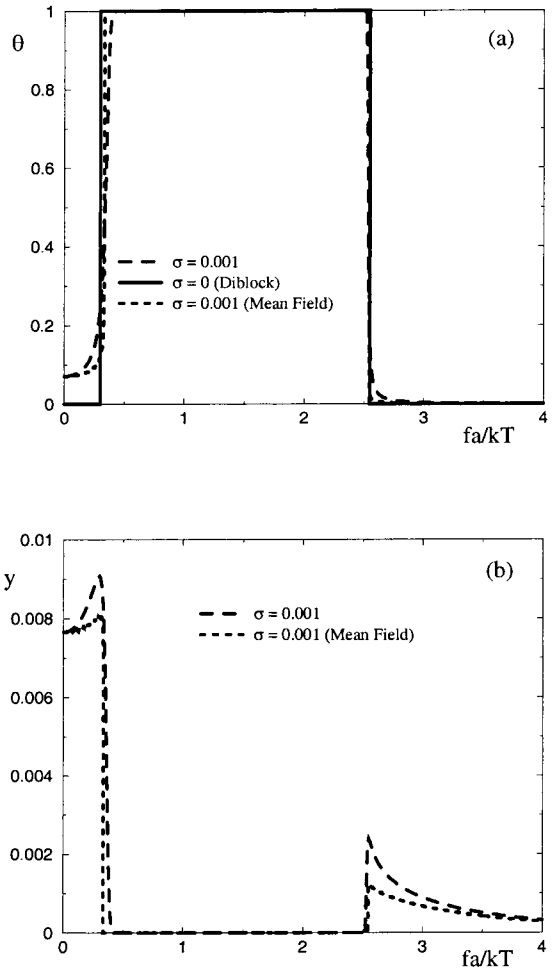

Figure 7. Plots of θ (a) and y (b) vs fa/kT as obtained from the rigorous solution (dashed lines), the $S_{mix} = 0$ or diblock approximation (continuous line), and the mean field approximation (dotted line) of the "rigid helices" model for $s = 0.9$ and $\gamma = {}^2/_5$.

Once the equilibrium characteristics of the chain for a given f are found, it is possible to obtain the force–extension diagram. The reduced end-to-end distance, r, as a function of the reduced tension, x, is (Appendix E)

$$r = (1-\theta)\mathcal{L}(x) + \theta\gamma\left[\coth(\gamma x) - \frac{sA}{\sinh(\gamma x)} - \frac{y}{\theta\gamma x}\right] \tag{44}$$

The corresponding force-extension diagram is plotted in Figure 8. Again, the results of the $S_{mix} = 0$ or diblock approximation are approached in the $\sigma \to 0$ limit. Note the "jump" in the force at $R = N\gamma a$.[36] At this point, between the two plateaus, the polypeptide is in a "purely" helical state, $\theta \simeq 1$ and $y \simeq 0$. This jump is another indication of difficulties due to the approximation of the helices as rigid rods. As we shall see in the next section, this unphysical feature disappears when the finite persistence length of the helices is allowed for.

It is of interest to compare the exact solution of the model to the mean field approximation (MF).[14,15] The MF approximation involves the removal of the explicit coupling between $P_c(n)$, $P_h(n)$, and f. As noted earlier, this decoupling is exact for $P_c(n)$. For $P_h(n)$ the decoupling is achieved by replacing $y\sum_n P_h(n)n\gamma\mathcal{L}(n\gamma x)$ in eq 36 by $\theta\gamma\mathcal{L}(\theta\gamma x/y)$. In effect, all helical domains are assigned an equal size of $k_h = \theta/y$, and the effect of their

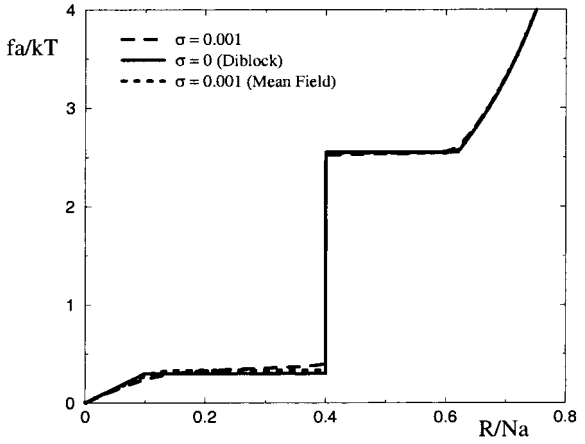

Figure 8. Extension force law as obtained from the rigorous solution (dashed lines), the $S_{mix} = 0$ or diblock approximation (continuous line), and the mean field approximation (dotted line) of the "rigid helices" model for $s = 0.9$ and $\gamma = {}^2/_5$.

polydispersity on F_{el} is ignored. The corresponding free energy is

$$\frac{F_{chain}^{MF}}{NkT} = -\theta \ln s - y \ln \sigma + (\theta - y)\ln\frac{\theta - y}{\theta} + y\ln\frac{y}{\theta} + (1-\theta-y)\ln\frac{1-\theta-y}{1-\theta} + y\ln\frac{y}{1-\theta} - (1-\theta)\mathcal{L}_{int}(x) - y\mathcal{L}_{int}(\theta\gamma x/y) \tag{45}$$

The equilibrium conditions $\partial F_{chain}/\partial\theta = \partial F_{chain}/\partial y = 0$ lead to

$$\frac{(1-\theta)(\theta - y)}{\theta(1-\theta-y)} = \frac{sx}{\sinh(x)}\exp\left[\gamma x\mathcal{L}\left(\frac{\theta\gamma x}{y}\right)\right] \tag{46}$$

$$\frac{y^2}{(\theta - y)(1-\theta-y)} = \frac{\sigma y}{\theta\gamma x}\sinh\left(\frac{\theta\gamma x}{y}\right)\exp\left[-\frac{\theta\gamma x}{y}\mathcal{L}\left(\frac{\theta\gamma x}{y}\right)\right] \tag{47}$$

The overall features of the MF solution are rather similar to the exact solution (Figure 8). The two primary differences are that (i) the MF solution indicates that the transition between the plateaus involves a second-order phase transition[37] in marked distinction to the exact solution and that (ii) because of point i, the plateaus do not exhibit an inflection point and the characteristic value of the slope, at $r = \gamma$, scales as σ^2 (see ref 15).

B. Flexible Helices. The assumption of perfectly rigid helical sequences implies an infinite persistence length. This assumption is reasonable so long as the size of the helical domains is small compared to the known values of their persistence length, P: A typical value is $P = 2000$ Å corresponding to $n_h = P/a_h \sim 10^3$ monomers. As we have seen, the rigid rod approximation for an infinite chain leads to helical domains that are larger. In turn, this leads to an unphysical jump, located at $r = \gamma$, in the force-extension curve.

Two approaches allow one to circumvent this difficulty. One approach treats each helical monomer as an extendable spring; i.e., it introduces an elastic penalty associated with the deviation of the length of the helical monomer from its equilibrium length a_h. This approach was utilized in the analysis of the extension behavior of DNA[38] and of poly(ethylene glycol).[39] In the following, we adopt a different approach that proved

successful in fitting the force curves of DNA. Within this approach the helical domains are considered as worm-like chains (WLC) instead of rigid rods (Appendix D).

In the WLC model, the helical domains are viewed as semiflexible chains undergoing undulations. The angular correlations along the chain decay because of the thermal undulations. Within this model the decay length of the angular correlations, λ, depends on the tension, $\lambda = \lambda(f)$. For weakly stretched chains, λ is identical to the persistence length, P, of the unperturbed chain, $\lambda(f) = P$ while for stronger tensions $\lambda(f) = P(kT/fP)^{1/2}$ and the correlations decay faster. A long helical domain of length $n\gamma a$ may thus be considered as a FJC of $n\gamma a/2\lambda(f)$ "rigid" segments of length $2\lambda(f)$ and the reduced end-to-end distance is

$$r = (1 - \theta)\mathcal{L}(x) + \theta\gamma\mathcal{L}(\alpha) \tag{48}$$

where $\mathcal{L}(x)$ is the Langevin function and $\alpha = \alpha(n_h\gamma x) = 2f\lambda(f)/kT$.[40] As before, the expression for r reflects two contributions. The first term corresponds to the alignment of the coil monomers with the applied tension using the FJC model. The second term describes the alignment of the semiflexible helical domains. In the following we assume that the helical domains are larger than $\lambda(f)$ that is, $\lambda(f) \ll n\gamma a$. The error introduced by this assumption diminishes as f increases and $\lambda(f)$ decreases. Within this approximation each $2\lambda(f)$ segment in a helical domain is aligned independently. This is in marked contrast to the rigid helices model where all the monomers within a helical domain align as a unit. As a result we may utilize $k_h = \theta/y = \sum_n nP_h(n)$ in order to obtain $y\sum_{n=1}^{\infty} P_h(n)\, n\gamma\, \mathcal{L}(\alpha) = \theta\gamma\mathcal{L}(\alpha)$. Consequently, r is independent of the probabilities $P_c(n)$ and $P_h(n)$. Expression 48 for $R(f)$ enables us to obtain (Appendix D) the corresponding elastic free energy $F_{el}(f) = -\int_0^f R\, df'$

$$F_{el}/NkT = -(1 - \theta)\mathcal{L}_{int}(x) - \frac{\theta}{4n_h}(4n_h\gamma x - \alpha)\mathcal{L}(\alpha) \tag{49}$$

In marked distinction to the "rigid helices" scenario, this elastic free energy is independent of the probabilities $P_h(n)$ and $P_c(n)$. Accordingly, $P_h(n)$ and $P_c(n)$ are not modified explicitly by the extension and obey eqs 8 and 9. In other words, the mean field approximation provides an exact solution of the problem. Altogether, the free energy of a polypeptide, using the WLC model for the helical domains, is

$$\frac{F_{chain}(f)}{NkT} = -\theta \ln s - y \ln \sigma + (\theta - y)\ln\frac{\theta - y}{\theta} + y\ln\frac{y}{\theta} + (1 - \theta - y)\ln\frac{1 - \theta - y}{1 - \theta} + y\ln\frac{y}{1 - \theta} - (1 - \theta)\mathcal{L}_{int}(x) - \frac{\theta}{4n_h}[4n_h\gamma x - \alpha]\mathcal{L}(\alpha) \tag{50}$$

This expression is similar to F_{chain} for a free, undeformed chain (eq 10). The two free energies differ in two respects: (i) Equation 50 contains an additive term $-\mathcal{L}_{int}(x)$ which is independent of θ and y. (ii) The force independent s in eq 10 is replaced by a force-dependent $\tilde{s}(x)$

$$\tilde{s}(x) = \frac{sx}{\sinh(x)}\exp[(\gamma x - \alpha/4n_h)\mathcal{L}(\alpha)] \tag{51}$$

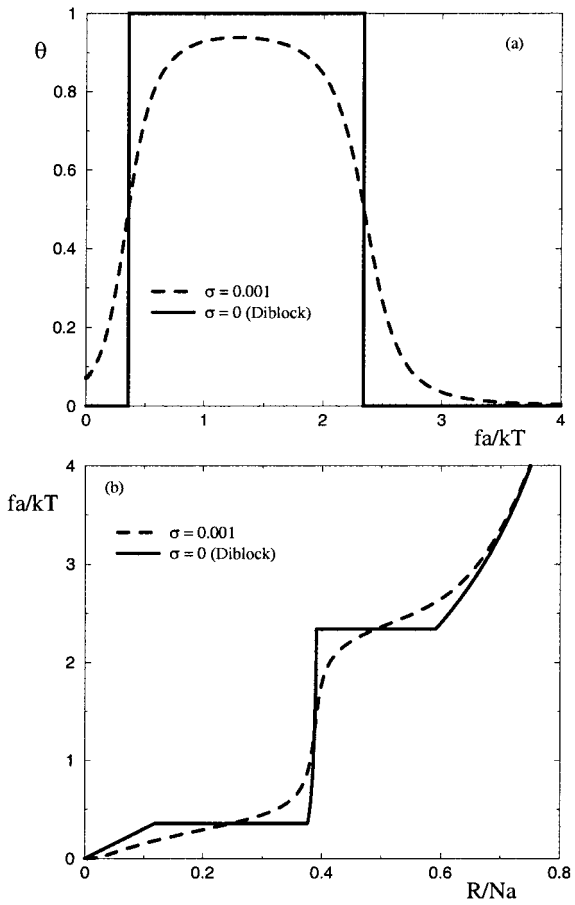

Figure 9. Plots of the helical fraction θ vs the reduced force fa/kT (a) and plots of the reduced force fa/kT vs $r = R/Na$ (b) as obtained from the rigorous solution (dashed lines) and the $S_{mix} = 0$ approximation (continuous line) of the "semiflexible helices" model for $s = 0.9$, $\gamma = 2/5$ and $n_h = 1000$.

Consequently, the equilibrium conditions $\partial F_{chain}/\partial\theta = F_{chain}/\partial y = 0$ lead to expressions of the same form as those found for the free chains but with $\tilde{s} = \tilde{s}(x)$ replacing s

$$\theta = \frac{1}{2} + \frac{1}{2}\frac{\tilde{s} - 1}{\sqrt{(\tilde{s} - 1)^2 + 4\sigma\tilde{s}}} \tag{52}$$

$$y = \frac{2\,\sigma\tilde{s}\,[(\tilde{s} - 1)^2 + 4\,\sigma\tilde{s}]^{-1/2}}{\tilde{s} + 1 + \sqrt{(\tilde{s} - 1)^2 + 4\,\sigma\tilde{s}}} \tag{53}$$

Substitution of the equilibrium values of θ and y for a given x in eq 48 yields the corresponding equilibrium end-to-end distance.

Plots of θ vs fa/kT and fa/kT vs R/Na, as obtained for the "semiflexible helices" scenario for different σ, are depicted in Figure 9. The replacement of the "rigid helices" by semiflexible ones leads to the disappearance of the "jump" in the force law. Another difference with respect to Figure 7a is that $\theta = 1$ is never attained. Again, we recover the appropriate diblock approximation in the limit of $\sigma \to 0$.

All the force-extension curves in Figure 9b cross at three different points. The external points correspond to a reduced force $x_\pm$ satisfying $\tilde{s}(x_\pm) = 1$. In turn, this condition is equivalent to the equilibrium condition $\partial F_{chain}/\partial\theta = 0$ for F_{chain} corresponding to the appropriate diblock approximation in which the coil is modeled as FJC and the helix as a WLC.[34] In particular, $x_\pm$

correspond to points in the plateaus where $\theta = {}^1/_2$, $r_\pm = [\mathcal{L}(x_\pm) + \gamma\mathcal{L}(\alpha_\pm)]/2$ and $\alpha_\pm \equiv \alpha(n_h\gamma x_\pm)$. The inner crossing point corresponds to x_0, for which $\tilde{s}(x)$ is maximal. x_0 satisfies the condition $\mathcal{L}(x_0) = \gamma\mathcal{L}[\alpha(n_h\gamma x_0)]$ i.e., to an equality between the projected length of the coil and the helical monomers along the force. From eqs 48 and 51, it is apparent that this condition is independent of σ. The plateaus depicted in Figure 9b exhibit a small but finite slope. Its characteristic value, at the inflection points $x_\pm$, scales with $\sigma^{1/2}$ (see ref 41).

V. Discussion and Conclusions

Our analysis of the extension behavior of polypeptides utilized two models. The two differ in the description of the helical domains. In one model, the helical domains are viewed as perfectly rigid and unextendable rods. The "rigid helices" model is reasonable so long as the length of the helical domains is small compared to the persistence length. In the second approach the helical domains are treated as long semiflexible polymers within the wormlike chain (WLC) model. This approach is preferable when the length of the helical domains is larger than the force-dependent decay length $\lambda(f)$. For each of the models, we presented a rigorous solution and two approximations. Within the diblock approximation one neglects the mixing entropy associated with the polydispersity of the domains. In the mean field approximation, there is no direct coupling between the distribution of sizes of the domains and the applied tension. In all cases considered, our analysis is based on the minimization of a phenomenological free energy. This approach is equivalent to the transfer matrix method, but it is physically transparent and mathematically simple.

The general features of the extension behavior of the polypeptides are independent of the model and the approximation scheme. Three different scenarios are possible for the extension of long polypeptides: (a) For $T > T_c$ the polypeptide remains in a coil configuration for all extensions. (b) In the range $T^* < T < T_c$, the chain is initially in a coil state, but it undergoes a helix−coil transition upon extension. The helical domains eventually melt upon stronger extension. Consequently, the force law exhibits two plateaus associated with a helix−coil coexistence. (c) For $T < T^*$ the unstretched polypeptide is a helix which melts upon stretching, thus leading to a force law with a single plateau.

In the second scenario, $T^* < T < T_c$, it is possible to distinguish between five regimes: (I) For weak extensions the polypeptide behaves as a random coil exhibiting a linear response. (II) Stronger extensions, and the associated loss of configurational entropy, favor the formation of helical domains. The associated coexistence of helical and coil domains gives rise to a lower plateau. (III) A steep crossover corresponding to the stretching of a stiff, mainly helical, polymer leads to (IV) a second, higher plateau due to the break-up of the helices when the imposed end-to-end distance is larger than the length of a fully helical chain. This plateau is also associated with a helix−coil coexistence. (V) When the helical content is reduced to zero, the chain exhibits a strong extension behavior of a random coil, with deviations from the linear response behavior due to finite extensibility effects.

In the third scenario, $T < T^*$, regimes I and II disappear and regime III extends down to $R = 0$. The small slope of the force law in this regime (Figure 10) is reminiscent of a plateau. However, this quasi-plateau is not associated with a helix−coil coexistence. Rather, it is due to the easy alignment of the persistent segments in the helix.

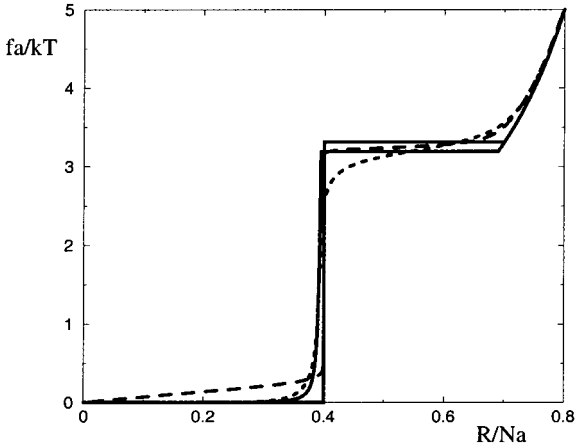

Figure 10. Plots of the reduced force fa/kT vs the reduced end-to-end distance R/Na for the $s > 1$ scenario ($s = 1.1$, $\gamma = 2/5$ and $n_h = 1000$). Long dashed line: "rigid helices" model ($\sigma = 0.01$). Dashed line: "semiflexible helices" model ($\sigma = 0.001$). Continuous lines: corresponding diblock approximations.

The two models, "rigid" and "semiflexible" helices, differ with regard to the precise details of the force diagram. Thus, region III occurs as a "jump" within the "rigid helices" model but is a steep yet gradual increase in the "semiflexible helices" model. The slope of the plateaus within the rigid helices model scales as $\sigma[\ln \sigma + \ln|\ln \sigma|]^2$ (see ref 15) while in the "semiflexible helices" model it scales as $\sigma^{1/2}$. Overall, the crossovers between the various regimes are sharper in the "rigid helices" model.

The diblock approximation, with the appropriate elastic term, allows to recover the main features, length and force scales, of each model. However, within this $S_{mix} = 0$ approximation, the transition takes place as a first-order phase transition. As a result it yields perfect plateaus, $f = f_{1co}$ or $f = f_{2co}$, and sharp crossovers. The mean field approximation, where the polydispersity is decoupled from the tension is exact for the "semiflexible helices" model. In the case of the "rigid helices" model this approximation, while it captures the overall features of the force law, also gives rise to a number of artifacts.[15]

Neither of the models considered captures the physics completely. The "rigid helices" model fails to describe the behavior of helical domain that are long compared to the persistence length. Within the "semiflexible helices" model all helical domains are assumed to be long in comparison to $\lambda(f)$. Accordingly, this model is likely to mishandle short helical domains that do behave as rigid rods. The importance of this deficiency diminishes, however, as f increases and $\lambda(f)$ decreases. A full description of the system requires a model combining the positive features of the two models described above.

While our analysis focused on the case of homopolypeptides forming α-helices, our approach can be extended to more complicated systems. In the case of heteropolypeptides it is necessary to allow for the residue dependence of Δf. For double-stranded DNA it is necessary to allow for long-range interactions as given by the Stockmayer formula for σ.[42−45]

Acknowledgment. We would like to thank M. N. Tamashiro and P. Pincus as well as I. Rouzina and V.

A. Bloomfield for providing us with their results prior to publication. The work of A.B. was supported by the Marie Curie Fellowship HPMF-CT-1999-00328.

Appendix A: Probability of Sizes

The minimization of the free energy (eq 7) with respect to the probabilities $P_h(n)$ and $P_c(n)$ yields

$$y \ln P_h(n) = \mu_1^h - y + n\mu_2^h \quad \text{(A1)}$$

$$y \ln P_c(n) = \mu_1^c - y + n\mu_2^c \quad \text{(A2)}$$

These are solved by $P_h(n) = A_h B_h^n$ and $P_c(n) = A_c B_c^n$ where

$$A_{h(c)} = \exp\left(\frac{\mu_1^{h(c)} - y}{y}\right) \quad \text{(A3)}$$

$$B_{h(c)} = \exp\left(\frac{\mu_2^{h(c)}}{y} y\right) \quad \text{(A4)}$$

The parameters A_h, A_c, B_h, and B_c are determined using the constraints

$$1 = \sum_{n=1}^{\infty} P_h(n) = \frac{A_h B_h}{1 - B_h} \quad \text{(A5)}$$

$$\frac{\theta}{y} = \sum_{n=1}^{\infty} n P_h(n) = \frac{A_h B_h}{(1 - B_h)^2} \quad \text{(A6)}$$

$$1 = \sum_{n=1}^{\infty} P_c(n) = \frac{A_c B_c}{1 - B_c} \quad \text{(A7)}$$

$$\frac{1 - \theta}{y} = \sum_{n=1}^{\infty} n P_c(n) = \frac{A_c B_c}{(1 - B_c)^2} \quad \text{(A8)}$$

yielding $A_h = y/(\theta - y)$, $B_h = (\theta - y)/\theta$, $A_c = y/(1 - \theta - y)$, and $B_c = (1 - \theta - y)/(1 - \theta)$ thus leading to the expressions 8 for $P_h(n)$ and 9 for $P_c(n)$.

Appendix B: Fractions of Helices and of Helical Monomers

The equilibrium conditions $\partial F_0/\partial\theta = \partial F_0/\partial y = 0$ with F_0 given by eq 10 lead to

$$(1 - \theta)(\theta - y) = s\theta(1 - \theta - y) \quad \text{(B1)}$$

$$y^2 = \sigma(\theta - y)(1 - \theta - y) \quad \text{(B2)}$$

From eq B2, we obtain

$$y = \frac{1}{2\alpha}[\sqrt{1 + 4\alpha\theta(1 - \theta)} - 1] \quad \text{(B3)}$$

or

$$y = \frac{2\theta(1 - \theta)}{\sqrt{1 + 4\alpha\theta(1 - \theta)} + 1} \quad \text{(B4)}$$

where $\alpha = (1 - \sigma)/\sigma$. The second condition (eq B1) yields

$$y = \frac{\theta(1 - \theta)(s - 1)}{\theta(s + 1) - 1} \quad \text{(B5)}$$

Combining these two equations allows us to eliminate y and obtain an equation for θ as a function of s and σ

$$\sqrt{1 + 4\alpha\theta(1 - \theta)} + 1 = \frac{2\theta(s + 1) - 2}{s - 1} \quad \text{(B6)}$$

leading to

$$\theta = \frac{1}{2} + \frac{1}{2}\frac{s - 1}{\sqrt{(s - 1)^2 + 4\sigma s}} \quad \text{(B7)}$$

$$y = \frac{2\sigma s((s - 1)^2 + 4\sigma s)^{-1/2}}{s + 1 + \sqrt{(s - 1)^2 + 4\sigma s}} \quad \text{(B8)}$$

Appendix C: The Freely Jointed Chain Model

The freely jointed chain model may be considered as the macromolecular analogue of the Langevin theory of paramagnetism.[46] A flexible chain comprising N monomers is the counterpart of N non interacting magnetic dipoles. The monomer length a is the analogue of the magnetic moment μ, the applied tension $\boldsymbol{f}$ plays the role of the magnetic field $\boldsymbol{B}$ and the induced magnetic moment, M is the analogue of the end to end distance R. Each monomer is assigned an orientational energy $-fa\cos\phi$, where ϕ is the angle between the force and the segment. The reduced end-to-end distance R/Na of the polymer, at a temperature T, is the mean value of $\cos\phi$

$$\frac{R}{Na} = \langle\cos\phi\rangle_T = \frac{\int_0^{2\pi} \cos\phi e^{x\cos\phi} \sin\phi d\phi}{\int_0^{2\pi} e^{x\cos\phi} \sin\phi d\phi} = \mathcal{L}(x) \quad \text{(C1)}$$

where $x = fa/kT$ is the reduced force and $\mathcal{L}(x) = \coth x - 1/x$ is the Langevin function. The corresponding elastic free energy in the fixed f ensemble, as given by $F_{el}(f) = -\int_0^f R\, df'$, leads to the following free energy per monomer

$$\frac{F_{el}(f)}{NkT} = -\mathcal{L}_{int}(x) \quad \text{(C2)}$$

where $\mathcal{L}_{int}(x) = \ln[\sinh x/x]$, is the integrated Langevin function.

The elastic free energy in the fixed R ensemble $F_{el}(R) = \int_0^R f\, dR' = Rf + F_{el}(f)$ is the Legendre transform of $F_{el}(f)$ where $r = R/Na$ and x are related by $r = \mathcal{L}(x)$. For weak extensions $F_{el}(R)$ reduces to the familiar Gaussian form $F_{el}(R) \simeq 3kTR^2/2Na^2$ while for strong extensions it diverges as $F_{el}(R) \sim \ln f \sim -\ln(1 - r)$.

Appendix D: The Wormlike Chain Model

Within the wormlike chain model (WLC), the polymer is viewed as a bendable rod.[47] In the following we consider the case of a chain of constant contour length L and a single bending modulus $E = kTP$ where P is the persistence length of the unperturbed chain. The chain trajectory is described by the position of a point on the chain, $\boldsymbol{R}(s)$, at the contour length s. The unit tangent vector $\boldsymbol{u}(s) = \partial\boldsymbol{R}/\partial s$ specifies the local orientation of the chain. A straight rod corresponds to $\boldsymbol{u}(s) = \text{const}'$ or $\partial\boldsymbol{u}/\partial s = 0$. Accordingly, the bending energy is a quadratic function of $\partial\boldsymbol{u}/\partial s$. Since $\boldsymbol{u}\cdot\partial\boldsymbol{u}/\partial s = 0$, the only quadratic form is

$$U_{bend} = \frac{E}{2}\int_0^L ds \left(\frac{\partial\boldsymbol{u}}{\partial s}\right)^2 \quad \text{(D1)}$$

and the conformational distribution of the chain, its partition function, up to a contour length s, is

$$\Psi \propto \exp\left[-\frac{P}{2}\int_0^s \mathrm{d}s' \left(\frac{\partial \boldsymbol{u}}{\partial s'}\right)^2\right] \quad \text{(D2)}$$

where Ψ is a function of $\boldsymbol{u}(0) = \boldsymbol{u_0}$ and $\boldsymbol{u}(s)$, i.e., $\Psi = \Psi(\boldsymbol{u_0},\boldsymbol{u}(s))$. This describes a Gaussian process with the constraint $\boldsymbol{u}^2 = 1$. As such it is analogous to rotational diffusion and obeys $\partial\Psi/\partial s = (1/2P)\hat{\mathscr{R}}^2\Psi$ where $\hat{\mathscr{R}} \equiv \boldsymbol{u} \times \partial/\partial \boldsymbol{u}$ is the rotational operator that is related to the angular momentum operator in quantum mechanics, $\hat{L}$, as $-i\hat{\mathscr{R}} = \hat{L}$. The correlations along the unperturbed chain decay as $\langle \boldsymbol{u}(s) \cdot \boldsymbol{u}(0)\rangle = \exp(-s/P)$. When the chain is subjected to an external field $U_{\mathrm{ext}}(\boldsymbol{u})$ the expression for Ψ becomes

$$\Psi \propto \exp\left[-\frac{P}{2}\int_0^s \mathrm{d}s' \left(\frac{\partial \boldsymbol{u}}{\partial s'}\right)^2 - \frac{1}{kT}\int_0^s \mathrm{d}s'\, U_{\mathrm{ext}}(\boldsymbol{u})\right] \quad \text{(D3)}$$

The effect of tension $\boldsymbol{f}$ is specified by $U_{\mathrm{ext}}(\boldsymbol{u}) = -\boldsymbol{f}\cdot\boldsymbol{u} = -f\cos\phi$ thus leading to

$$\frac{\partial \Psi}{\partial s} = \left(\frac{1}{2P}\hat{\mathscr{R}}^2 + \frac{f\cos\phi}{kT}\right)\Psi \quad \text{(D4)}$$

It is possible to find an eigenstate satisfying $\partial\psi/\partial s = -g\psi$. For a long chain, the free energy is extensive and the eigenfunction expansion of $\Psi(\boldsymbol{u_0},\boldsymbol{u}(L))$ involves terms of the form $\exp(-gL)\psi(\boldsymbol{u_0})\psi(\boldsymbol{u}(L))$ where gkT is the free energy per unit length due to the bending fluctuations.

An analytical solution of eq D4 is not available. Marko and Siggia[35] argued that the WLC free energy is determined by the lowest g in the spectrum and utilized a variational approach in order to determine g. The force law is then recovered from $R/L = -kT\mathrm{d}g/\mathrm{d}f$. To this end they chose the trial function $\psi(\boldsymbol{u}) \propto \exp[\alpha\cos(\phi)/2]$ where α is a variational parameter and ϕ is the angle between $\boldsymbol{u}$ and $\boldsymbol{f}$. As we shall see later, $\alpha = 2f\lambda/kT$ where $\lambda = \lambda(f)$ is a force-dependent decay length that replaces P in characterizing the decay of $\langle \boldsymbol{u}(s)\cdot\boldsymbol{u}(0)\rangle$. Assuming $\int \mathrm{d}\boldsymbol{u}^2\, \psi^2(\boldsymbol{u}) = 1$, this leads to[48]

$$ga = \min_{\alpha} Ga = \min_{\alpha}\left(\frac{\alpha a}{4P} - x\right)\mathscr{L}(\alpha) \quad \text{(D5)}$$

where $\mathscr{L}(x) = \coth x - 1/x$ is the Langevin function and $x = fa/kT$ is the reduced force. The argument $\alpha = \alpha(n_{\mathrm{h}}\gamma x)$ that minimizes g is specified by $\partial G/\partial\alpha = 0$ which leads to

$$\mathscr{L}(\alpha) = [4n_{\mathrm{h}}\gamma x - \alpha]\mathscr{L}'(\alpha) \quad \text{(D6)}$$

where $\mathscr{L}'(x) = d\mathscr{L}(x)/dx$. Here $n_{\mathrm{h}}\gamma x = Pf/kT$. The end-to-end distance is simply deduced from

$$R = -L\frac{dga}{dx} = -L\frac{\partial Ga}{\partial x} - L\frac{\partial Ga\,\partial\alpha}{\partial\alpha\ \partial x} = -L\frac{\partial Ga}{\partial x} = L\mathscr{L}(\alpha) \quad \text{(D7)}$$

This end-to-end distance is identical to the one obtained from the FJC model if the reduced force $2n_{\mathrm{h}}\gamma x$ is replaced by $\alpha(n_{\mathrm{h}}\gamma x)$.[40] For small forces, when $\mathscr{L}(\alpha) \approx \alpha/3$ and $\mathscr{L}'(\alpha) \approx {}^1/_3$, eq D6 leads to $\alpha(n_{\mathrm{h}}\gamma x) \approx 2n_{\mathrm{h}}\gamma x$ or $\lambda \approx P$. For high tension, α increases more slowly. In this limit, $\mathscr{L}(\alpha) \approx 1 - 1/\alpha$ while $\mathscr{L}'(\alpha) \approx 1/\alpha^2$ and eq D6 yields $\alpha(n_{\mathrm{h}}\gamma x) \simeq (4n_{\mathrm{h}}\gamma x)^{1/2}$ or $\lambda \approx P(kT/fP)^{1/2}$. The difference between the variational approximation and the exact solution of the WLC model is less than 1.5% over the whole range of forces. The associated elastic free energy, $F_{\mathrm{el}}(f) = -\int_0^f R\ \mathrm{d}f'$ is

$$F_{\mathrm{el}}(f) = LkTg \quad \text{(D8)}$$

since $R\ \mathrm{d}f = -LkT\mathrm{d}g$ and where g corresponds to the explicit expression in eq D5 with $\alpha = \alpha(n_{\mathrm{h}}\gamma x)$ according to eq D6. In eq 49, this elastic free energy is used for a helical domain of length $L = n\gamma a$ with n monomers.

Appendix E: Rigid Helices Approximation

In this appendix we determine the probabilities $P_{\mathrm{c}}(n)$ and $P_{\mathrm{h}}(n)$ of a coil and a helical segment with n monomers as well as the fractions of helical monomers θ and of helices y.

The extremum condition of F_{chain} (eq 38) with respect to $P_{\mathrm{c}}(n)$ and $P_{\mathrm{h}}(n)$ yields

$$y \ln P_{\mathrm{c}}(n) = \mu_1^{\mathrm{c}} - y + n\mu_2^{\mathrm{c}} \quad \text{(E1)}$$

$$y \ln P_{\mathrm{h}}(n) = \mu_1^{\mathrm{h}} - y + n\mu_2^{\mathrm{h}} + y\,\mathscr{L}_{\mathrm{int}}(n\gamma x) \quad \text{(E2)}$$

This leads to

$$P_{\mathrm{c}}(n) = \exp\left(\frac{\mu_1^{\mathrm{c}} - y}{y} + n\frac{\mu_2^{\mathrm{c}}}{y}\right), \quad \text{(E3)}$$

$$P_{\mathrm{h}}(n) = \exp\left(\frac{\mu_1^{\mathrm{h}} - y}{y} + n\frac{\mu_2^{\mathrm{h}}}{y}\right)\frac{\sinh(n\gamma x)}{n\gamma x}. \quad \text{(E4)}$$

The probability $P_{\mathrm{c}}(n)$ is easily determined since it is not modified by the stretching. The normalization of the probability $P_{\mathrm{c}}(n)$ and the mean size $k_{\mathrm{c}} = (1-\theta)/y = \sum_n nP_{\mathrm{c}}(n)$ lead to (Appendix A)

$$\exp\left(\frac{\mu_1^{\mathrm{c}} - y}{y}\right) = \frac{y}{1-\theta-y}, \quad \text{(E5)}$$

$$\exp\left(\frac{\mu_2^{\mathrm{c}}}{y}\right) = \frac{1-\theta-y}{1-\theta} \quad \text{(E6)}$$

and

$$P_{\mathrm{c}}(n) = \frac{y}{1-\theta-y}\left(\frac{1-\theta-y}{1-\theta}\right)^n \quad \text{(E7)}$$

The probability $P_{\mathrm{h}}(n)$ is modified by the stretching and is more difficult to obtain. To do so, we consider the two remaining equilibrium conditions $\partial F_{\mathrm{chain}}/\partial\theta = 0$ and $\partial F_{\mathrm{chain}}/\partial y = 0$ after substitution of the equilibrium form of $P_{\mathrm{c}}(n)$ as given by eq E7

$$\ln s = -\ln\left(\frac{1-\theta-y}{1-\theta}\right) + \frac{\mu_2^{\mathrm{h}}}{y} + \mathscr{L}_{\mathrm{int}}(x) \quad \text{(E8)}$$

$$\ln\sigma = \ln\left(\frac{y}{1-\theta-y}\right) + \sum_{n=1}^{\infty} P_{\mathrm{h}}(n)\ln P_{\mathrm{h}}(n) - \frac{\mu_2^{\mathrm{h}}\theta}{y^2} - \sum_{n=1}^{\infty} P_{\mathrm{h}}(n)\mathscr{L}_{\mathrm{int}}(n\gamma x) \quad \text{(E9)}$$

Using eq E4 for $P_h(n)$ we may rewrite eq E9 as

$$\ln \sigma = \ln\left(\frac{y}{1-\theta-y}\right) + \frac{\mu_1^h}{y} - 1 \tag{E10}$$

From eqs E10 and E8, we obtain

$$\exp\left(\frac{\mu_1^h - y}{y}\right) = \frac{\sigma(1-\theta-y)}{y} \tag{E11}$$

$$\exp\left(\frac{\mu_2^h}{y}\right) = \left(\frac{1-\theta-y}{1-\theta}\right)\frac{sx}{\sinh x} \tag{E12}$$

thus leading to

$$P_h(n) = \frac{\sigma(1-\theta-y)^{n+1}}{y\,(1-\theta)^n}\left(\frac{sx}{\sinh x}\right)^n \frac{\sinh(n\gamma x)}{n\gamma x} \tag{E13}$$

θ and y are obtained using the normalization condition for $P_h(n)$ and the constraint on the average size of the helical domains $k_h = \theta/y$

$$\sum_{n=1}^{\infty}\frac{(1-\theta-y)^{n+1}}{(1-\theta)^n}\left(\frac{s\,x}{\sinh x}\right)^n \frac{\sinh(n\gamma x)}{n\gamma x} = \frac{y}{\sigma} \tag{E14}$$

$$\sum_{n=1}^{\infty}\frac{(1-\theta-y)^{n+1}}{(1-\theta)^n}\left(\frac{sx}{\sinh x}\right)^n \frac{\sinh(n\gamma x)}{\gamma x} = \frac{\theta}{\sigma} \tag{E15}$$

Using the definition $\sinh(x) = (e^x - e^{-x})/2$ as well as the relations

$$\sum_{n=1}^{\infty} X^n = \frac{X}{1-X} \tag{E16}$$

$$\sum_{n=1}^{\infty} \frac{X^n}{n} = -\ln(1-X) \tag{E17}$$

we obtain

$$\frac{y}{\sigma} = \frac{1-\theta-y}{2\gamma x}\ln\left(\frac{1-sAe^{-\gamma x}}{1-sAe^{\gamma x}}\right) \tag{E18}$$

$$\frac{\theta}{\sigma} = \frac{(1-\theta-y)\,s\,A\sinh(\gamma x)}{\gamma x\,(1-sAe^{\gamma x})(1-sAe^{-\gamma x})} \tag{E19}$$

where

$$A(x,\theta,y) = \frac{(1-\theta-y)x}{(1-\theta-y)\sinh(x)} \tag{E20}$$

These two equations for θ and y can be easily solved numerically. To this end we introduce a new variable $\epsilon = (1-\theta-y)/(1-\theta)$. Equation E18 may then be rewritten as

$$1-\epsilon = \frac{\sigma\epsilon}{2\gamma x}\ln\left[\frac{1-sA(x,\epsilon)e^{-\gamma x}}{1-sA(x,\epsilon)e^{\gamma x}}\right] \tag{E21}$$

with $A(x,\epsilon) = \epsilon x/\sinh(x)$. This equation allows to determine ϵ as a function of s, σ, γ, and x. θ is then obtained from eq E19 in the form

$$\frac{\theta}{1-\theta} = \frac{\sigma\epsilon sA(x,\epsilon)\sinh(\gamma x)}{\gamma x[1-sA(x,\epsilon)e^{-\gamma x}][1-sA(x,\epsilon)e^{\gamma x}]} \tag{E22}$$

Knowing ϵ and θ, y is simply determined from $y = (1-\epsilon)(1-\theta)$.

To obtain the end-to-end distance R as a function of the tension f, we use $R = -\partial F_{chain}/\partial f$, which leads to

$$r = (1-\theta)\mathcal{L}(x) + y\sum_{n=1}^{\infty} P_h(n)\, n\gamma\, \mathcal{L}(n\gamma x) \tag{E23}$$

The second term may be tackled exactly using the definition $\mathcal{L}(x) = \coth x - 1/x$

$$r = (1-\theta)\mathcal{L}(x) + y\sum_{n=1}^{\infty} P_h(n)\, n\gamma \frac{\cosh(n\gamma x)}{\sinh(n\gamma x)} - \frac{y}{x}$$

$$= (1-\theta)\mathcal{L}(x) + \frac{\sigma}{2x}(1-\theta-y)\,(\Lambda_+ + \Lambda_-) - \frac{y}{x} \tag{E24}$$

with

$$\Lambda_\pm = \frac{sA(x,\theta,y)e^{\pm\gamma x}}{1-sA(x,\theta,y)e^{\pm\gamma x}} \tag{E25}$$

The expression for r may be simplified using eq E19

$$\Lambda_+ + \Lambda_- = \frac{2sA(x,\theta,y)\cosh(\gamma x) - 2s^2A^2(x,\theta,y)}{[1-sA(x,\theta,y)\,e^{\gamma x}][1-sA(x,\theta,y)\,e^{-\gamma x}]}$$

$$= \frac{2\theta\gamma x[\cosh(\gamma x) - sA(x,\theta,y)]}{\sigma(1-\theta-y)\sinh(\gamma x)} \tag{E26}$$

which yields

$$r = (1-\theta)\mathcal{L}(x) + \theta\gamma\left[\coth(\gamma x) - \frac{sA(x,\theta,y)}{\sinh(\gamma x)}\right] - \frac{y}{x} \tag{E27}$$

References and Notes

(1) Mehta, A. D.; Rief, M.; Spudich, J. A.; Smith, D. A.; Simmons, R. M. *Science* **1999**, *283*, 1689−1695.
(2) Smith, S. B.; Finzi, L.; Bustamante, C. *Science* **1992**, *258*, 1122−1126.
(3) Bensimon, A.; Simon, A.; Chiffaudel, A.; Croquette, V.; Heslot, F.; Bensimon, D. *Science* **1994**, *265*, 2096−2098.
(4) Austin, R. H.; Brody, J. P.; Cox, E. C.; Duke, T.; Volkmuth, W. *Phys. Today* **1997**, *50*, 32−38.
(5) Maier, B.; Bensimon, D.; Croquette, V. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12002−12007.
(6) Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, H.; Gaub, H. E. *Science* **1997**, *276*, 1109−1112.
(7) Kellermayer, M. S. Z.; Smith, S. B.; Granzier, H. L.; Bustamante, C. *Science* **1997**, *276*, 1112−1116.
(8) Tskhovrebova, L.; Trinick, J.; Sleep, J. A.; Simmons, R. M. *Nature* **1997**, *387*, 308−312.
(9) Oberhausser, A. H.; Marszalek, P. E.; Erickson, H. P.; Fernandez, J. M. *Nature* **1998**, *393*, 181−185.
(10) Rief, M.; Oesterhelt, F.; Heymann, B.; Gaub, H. E. *Science* **1997**, *275*, 1295−1297.
(11) Li, H.; Rief, M.; Oesterhelt, F.; Gaub, H. E. *Adv. Mater.* **1998**, *3*, 316−319.
(12) Oesterhelt, F.; Rief, M.; Gaub, H. E. *New J. Phys.* **1999**, *1*, 6.1−6.11.
(13) Flory, P. J. *Science* **1956**, *124*, 53−60.
(14) Buhot, A.; Halperin, A. *Phys. Rev. Lett.* **2000**, *84*, 2160−2163.
(15) Tamashiro, M. N.; Pincus, P. *Phys. Rev. E* **2001**, *63*, 021909.

(16) Birshtein, T. M.; Ptitsyn, O. B. *Conformations of Macromolecules*; John Wiley: New York, 1966.
(17) Poland, D.; Sheraga, H. *Theory of Helix-Coil Transitions in Biopolymers*; Academic Press: New York, 1970.
(18) Grosberg, A. Yu.; Khokhlov, A. R. *Statistical Physics of Macromolecules*, AIP Press: New York, 1994.
(19) A transfer matrix version of the $S_{mix} = 0$ approximation for this problem is describe in ref 16. The analysis is concerned only with the case of chains just above the melting temperature of the helix, modeling the helix as a rigid rod. This approach yields the general features of the force curve as depicted in Figure 5. However, this analysis fails to identify the relevant force and length scales because the elasticity of the blocks is not specified.
(20) Deming, T. J. *Nature* **1997**, *390*, 386−389.
(21) Buhot, A.; Halperin, A. *Europhys. Lett.* **2000**, *50*, 756−761.
(22) Yu, Y.-C.; Berndt, P.; Tirrell, M.; Fields. G. B. *J. Am. Chem. Soc.* **1996**, *118*, 12515−12520.
(23) Fields, G. B.; Lauer, J. L.; Dori, Y.; Forns, P.; Yu, Y.-C.; Tirrell, M. *Biopolymers* **1998**, *47*, 143−151.
(24) James, K.; Levene, H.; Parsons, J. R.; Kohn, J. *Biomaterials* **1999**, *20*, 2203−2212.
(25) Tjia, J. S.; Aneskievich, B. J.; Moghe, P. V. *Biomaterials* **1999**, *20*, 2223−2233.
(26) To be precise, this approach is equivalent to a transfer matrix description utilizing a second-order transfer matrix. In turn, this description is sufficient for the analysis of the helix−coil transition in an infinite chain.
(27) Zimm, B. H.; Bragg, I. K. *J. Chem. Phys.* **1959**, *31*, 526−535.
(28) It is customary to estimate Δh by the energy of the H bond, E_H. The values of E_H vary in the range 10−20 kJ/mol. For $E_H = -10$ kJ/mol, we find $\sigma \approx 3.3 \times 10^{-4}$ at room temperature.[18]
(29) The constraint that helical domains incorporate at least three monomers can be easily introduced by changing the lower boundary of sums involving $P_h(n)$ from $n = 1$ to $n = 3$.
(30) The final expression for the mixing entropy for the free chain can be calculated directly.[14] $S_{mix} = k \ln W$ where W is the product of two factors: $W_{th} = \binom{N\theta}{Ny}$ the number of possible placements of Ny boundaries to separate $N\theta$ helical bonds into Ny helical domains and $W_{tc} = \binom{N(1-\theta)}{Ny}$ the number of ways to place Ny boundaries to separate $N(1-\theta)$ coil bonds into Ny coil domains. Using the Sterling formula $N! \sim N \ln N - N$, we obtain $S_{mix}(\theta, y, N) = k \ln W \simeq Nk[\theta s(y/\theta) + (1-\theta)s(y/(1-\theta))]$ where $s(x)$ stands for $-x\ln x - (1-x)\ln(1-x)$.
(31) Landau, L. D.; Lifshitz, E. M. *Statistical Physics*, Pergamon: Oxford, England, 1980.
(32) Borisov, O. V.; Halperin, A. *Macromolecules* **1997**, *30*, 4432−4444.
(33) In the fixed f ensemble, the free energy is $F_{chain}/NkT = -\theta \ln s - (1-\theta)\mathcal{L}_{int}(x) - \gamma\theta x$ where $\mathcal{L}_{int}(x) = \ln[\sinh(x)/x]$ (see Appendix C) and the boundaries of the plateau are specified by the equilibrium condition $\partial F_{chain}/\partial\theta = 0 = -\ln s + \mathcal{L}_{int}(x_\pm) - \gamma x_\pm$ while the end-to-end distance is $R = -\partial F_{chain}/\partial f = Na[(1-\theta)\mathcal{L}'(x) + \theta\gamma]$.
(34) In this case, the $S_{mix} = 0$ approximation leads to (see Appendix D and section IVB) $F_{chain}/NkT = -\theta \ln s - (1-\theta)\mathcal{L}_{int}(x) - (\theta/4n_h)[4n_h\gamma x - \alpha]\mathcal{L}'(\alpha)$. The domain boundaries are specified by $0 = -\ln s + \mathcal{L}_{int}(x_\pm) - [4n_h\gamma x_\pm - \alpha_\pm]\mathcal{L}'(\alpha_\pm)/4n_h$ with $\alpha_\pm \equiv \alpha(n_h\gamma x_\pm)$ and the end-to-end distance is $R = Na[(1-\theta)\mathcal{L}'(x) + \theta\gamma\mathcal{L}'(\alpha)]$.
(35) Marko, J. F.; Siggia, E. D. *Macromolecules* **1995**, *28*, 8759−8770.
(36) A careful study of the behavior at this apparent jump shows that there is no first-order transition.
(37) Even though the force law displays a jump, the phase transition is a second-order transition since the order parameter θ changes continuously and attains unity.
(38) Léger, J. F.; Romano, G.; Sarkar, A.; Robert, J.; Bourdieu, L.; Chatenay, D.; Marko, J. F. *Phys. Rev. Lett.* **1999**, *83*, 1066−1069.
(39) Kreuzer, H. G.; Wang, R. L. C.; Grunze, M. *New J. Phys.* **1999**, *1*, 21.1−21.16.
(40) The factor 2 comes from the difference between the persistence length $P = n_h\gamma a$ and the Kuhn length $2n_h\gamma a$.
(41) To characterize the slope, we expand $\tilde{s}(x)$ around $x_\pm$ where $\tilde{s}(x_\pm) = 1$. In the immediate vicinity of $x_\pm$ we have $\tilde{s}(x) \simeq 1 + \delta(x - x_\pm)$ where δ is a constant. Accordingly, $\theta \simeq {}^1/_2(1 + ((\tilde{s} - 1)/\sqrt{4\sigma\tilde{s}}) \simeq {}^1/_2[1 + \delta(x-x_\pm)/2\sigma^{1/2}]$ and $r = r_\pm + \delta(x-x_\pm)/2\sigma^{1/2}\{\mathcal{L}'(x_\pm) - \gamma\mathcal{L}'[\alpha(x_\pm)]\}$ leading to $x - x_\pm \sim \sigma^{1/2}(r - r_\pm)$.
(42) Rouzina, I.; Bloomfield, V. A. *Biophys. J.* **2001**, *80*, 882−893.
(43) Rouzina, I.; Bloomfield, V. A. *Biophys. J.* **2001**, *80*, 894−900.
(44) Williams, M. C.; Wenner, J. R.; Rouzina, I.; Bloomfield, V. A. *Biophys. J.* **2001**, *80*, 874−881.
(45) Williams, M. C.; Wenner, J. R.; Rouzina, I.; Bloomfield, V. A. *Biophys. J.* **2001**, *80*, 1932−1939.
(46) Hill, T. L. *An Introduction to Statistical Thermodynamics*; Addison-Wesley: Reading, England, 1960.
(47) Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*, Clarendon Press: Oxford, England, 1986.
(48) The variational eigenstate $\psi = M \exp[\alpha \cos(\phi)/2]$, where M is the normalization factor. With $d^2u = (1/4\pi)\, d\omega \sin\phi\, d\phi$, the normalization condition $\int d^2u\, \psi^2 = 1$ yields $M^2 = \alpha/\sinh(\alpha)$. Since there is no ω dependence, $\hat{\mathcal{R}} = i\hat{\varphi}\partial/\partial\theta$ where $\hat{\varphi}$ is the appropriate unit vector. Consequently eq D4 leads to $g = \int d^2u\, \{(1/2P)(\hat{\mathcal{R}}\psi)^2 - (f/kT)\psi^2 \cos\phi\} = (M^2/2)\int_{-1}^{1}\{(\alpha^2/8P)(1-y^2)\exp(\alpha y) - (f/kT)y\exp(\alpha y)\}\, dy$.

MA011631W