

Theory of Biopolymer Stretching at High Forces

Ngo Minh Toan^{*,†} and D. Thirumalai^{*,†,‡}

[†]*Biophysics Program, Institute for Physical Science and Technology and* [‡]*Department of Chemistry and Biochemistry, University of Maryland at College Park, College Park, Maryland 20742*

Received September 8, 2009; Revised Manuscript Received March 26, 2010

ABSTRACT: We provide a unified theory for the high force entropic elasticity of biopolymers solely in terms of the persistence length, ξ_p , and the monomer spacing, a . When the force $f > \mathcal{F}_h \sim k_B T \xi_p / a^2$ the biopolymers behave as freely jointed chains (FJC) while in the range $\mathcal{F}_l \sim k_B T / \xi_p < f < \mathcal{F}_h$ the worm-like chain (WLC) is a better model. We show that ξ_p can be estimated from the force extension curve (FEC) at the extension $x \approx 1/2$ (normalized by the contour length of the biopolymer). After validating the theory using simulations, we provide a quantitative analysis of the FECs for a diverse set of biopolymers (dsDNA, ssRNA, ssDNA, polysaccharides, and unstructured PEVK domain of titin) for $x \geq 1/2$. The success of a specific polymer model (FJC or WLC) to describe the FEC of a given biopolymer is naturally explained by the theory. Only by probing the response of biopolymers over a wide range of forces can the f -dependent elasticity be fully described.

1. Introduction

The elasticity of biopolymers, probed using single molecule methods, is often analyzed using the worm-like chain (WLC)^{1–3} model or the freely jointed chain (FJC) model. For example, the force–extension curves (FECs) of double stranded (ds) DNA (at high monovalent salt concentration) are accurately reproduced using the WLC model with essentially the persistence length (ξ_p) as the only adjustable parameter while the FJC model does not fit the data.⁴ The WLC model is also successful in describing the FEC of ssRNA⁵ and proteins.^{6,7} In contrast, the FEC of ssDNA is better explained using the FJC model than the WLC model.⁸ Given that the values of ξ_p of ssDNA and ssRNA are similar,^{9,10} it is hard to understand the source of the discrepancy. In addition, for polysaccharides^{11–14} the FJC is used to extract the parameters from the FEC.

At low temperatures and in the absence of force, a typical biopolymer (RNA, proteins and even DNA) adopts regular structures that are enthalpically stabilized (Figure 1). For example, RNA forms hairpins and can adopt compact structures. The manifestation of such structures can be gleaned from FEC. In some cases the presence of hairpins or other preferred structures appear as plateaus in the FEC.^{15–17} When the stretching force, f , is large enough to disrupt these structures the FEC is determined to a great extent by entropic forces (Figure 1). Here, we use general theoretical arguments to determine the polymer model that best describes the measured FEC. In particular, we derive a complete picture of the stretching physics of polymers at relative high forces, solely in terms of ξ_p and the tensile screening length $\xi_t \equiv k_B T / f$,¹⁸ where $k_B T$ is the thermal energy. In the process, the FECs for the FJC and WLC models naturally appear as limiting cases within a unified theory for the stretching of polymer models, regardless of the nature of interactions between the monomers. Because the theory makes use of only general geometrical and scaling arguments that describe the response of a polymer under tension in terms of ξ_p and ξ_t , we believe that our approach can be used to analyze the FECs of virtually any model polymer, and more importantly, many classes of biopolymers.

We propose that the FEC for any polymer should exhibit a near universal behavior as long as the enthalpic effects are small.

When enthalpic effects are relevant, low force elasticity exhibits complex dependence on temperature as illustrated in experiments¹⁹ and lattice simulations.^{20–22} Such nonuniversal effects are not considered here. For a generic polymer, at low forces we expect the extension to scale as $x \sim f$.^{23–25} As the force increases the FEC could exhibit a plateau region (poor solvent)²⁶ or a universal behavior (good solvent) where the extension scales as $x \sim f^{2/3}$ ¹⁸ (Pincus regime). This behavior cannot extend beyond the region $x \sim 1/2$, and in this regime entropic forces determine FEC. In the force range $\mathcal{F}_l \sim k_B T / \xi_p < f < \mathcal{F}_h \sim k_B T / \xi_p / a^2$ with a being approximately the monomer size, the FEC should coincide with the prediction of the WLC model. For $f > \mathcal{F}_h$ the polymer is better described using the FJC model. Thus, as long as $f > \mathcal{F}_h$ the FEC can be fit using the analytical results from WLC or FJC model regardless of the interactions that stabilize the ordered structures of the biopolymer at low forces ($f < \mathcal{F}_h$). The model that provides the most accurate fit of the measured FEC depends on the force range of the FEC as well as the specific biopolymer, which is characterized by ξ_p and a . In other words, the elasticity of the biopolymer (or any polymer for that matter) depends on the force.⁸

2. Theory

For a chain subject to a force f along the z -axis, the tensile energy is $\mathcal{H}_S = -f a \sum_{i=1}^N \cos \theta_{f,i}$, where $\theta_{f,i}$ is the angle between bond i and the force vector. In the limit the tensile screening length $\xi_t \equiv k_B T / f$ ^{18,24} is less than ξ_p , the excluded volume or other solvent-mediated intrachain interactions are irrelevant.^{27,28} We make the physically reasonable assumption that, at high values of f , the chain Hamiltonian has the global minimum when all the bond angles are zero, which corresponds to the rod state. For chains with special bond constraints such as the freely rotating chain, we shall assume that we can define virtual bonds and virtual bond angles that satisfy this assumption.

2.1. The FJC Regime Is Universal for All Discrete Chains at High Force. In the following we denote

$$\langle X(\theta) \rangle_{\mathcal{H}_S} \equiv \int_0^\pi X(\theta) \exp\left(-\frac{\mathcal{H}_S}{k_B T}\right) \sin \theta \, d\theta$$

as the average with respect to \mathcal{H}_S of X as an arbitrary function of some angle θ . For any two bonds i and j with

*Corresponding authors.

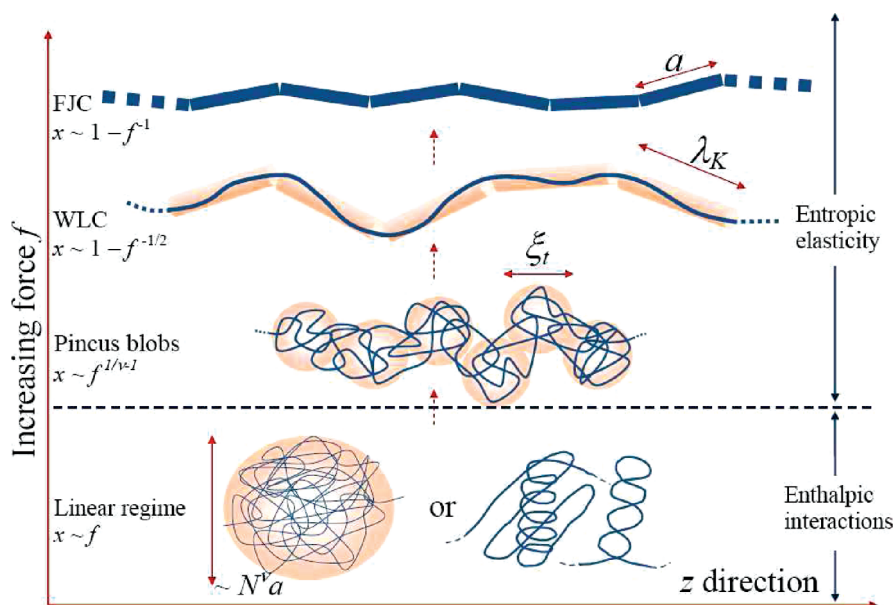


Figure 1. Summary of the various stretching regimes in polymers or biopolymers. In the smallest force regime, the extension increases linearly with the force. In the case of biopolymers (RNA and protein) the structure in the collapsed state is ordered. Polymers in poor solvents are collapsed. The size of the polymer or a globular protein is about $N^v a$, where N is the number of monomers, a is the monomer spacing and v is the scaling exponent, which is $\approx 1/3$ in poor solvent, $1/2$ in Θ -solvent or for ideal chains, and $3/5$ in good solvent.²⁴ In good or θ -solvent, as the force increases the chain extends into a series of Pincus blobs with typical size $\xi_t = k_B T / f$. In this regime, the extension scales as $x \propto f^{1/(v+1)}$.^{18,24} When the force is comparable to $\mathcal{F}_l = k_B T / \xi_p$ ($\xi_t \approx \xi_p$), each Pincus “blob” contains just a single segment on the order of a persistence length. In this force range, the Pincus scaling law breaks down and the elasticity is described by the WLC model. The chain can be viewed as an effective freely jointed chain of segments of length

$$\lambda_K \sim \sqrt{\xi_t \xi_p}$$

Intrachain interactions may still be relevant within the length scale λ_K but are negligible on larger length scales. The extension goes as $1 - x \propto f^{-1/2}$. Finally, when the force reaches $\mathcal{F}_h \sim k_B T \xi_p / a^2$, all chain bonds are strongly aligned with the force and the intrinsic chain interactions are irrelevant. Thus, the chain behaves as a freely jointed chain (segment length a) with the extension scaling as $1 - x \propto f^{-1}$. In some cases of biopolymers discussed in the main text, \mathcal{F}_h can be so high such that bond overstretching can occur before the force reaches \mathcal{F}_h and the theoretical FJC regime may not be actually observed. The universal aspects discussed here are not applicable when enthalpic interactions are relevant.

bond angle θ_{ij} , because of the triangle inequality $0 \leq \theta_{ij} \leq \theta_{fi} + \theta_{fj}$, the average of $\cos(\theta_{ij})$ with respect to *only* \mathcal{H}_S satisfies:

$$\begin{aligned} \langle \cos \theta_{ij} \rangle_{\mathcal{H}_S} &\geq \langle \cos \theta_{f,i} \cos \theta_{f,j} \rangle_{\mathcal{H}_S} - \langle \sin \theta_{f,i} \sin \theta_{f,j} \rangle_{\mathcal{H}_S} \\ &= \langle \cos \theta_{f,i} \rangle_{\mathcal{H}_S} \langle \cos \theta_{f,j} \rangle_{\mathcal{H}_S} - \langle \sin \theta_{f,i} \rangle_{\mathcal{H}_S} \langle \sin \theta_{f,j} \rangle_{\mathcal{H}_S} \\ &= \left(\coth(a/\xi_t) - \frac{\xi_t}{a} \right)^2 - \left(\frac{\pi}{2} \frac{I_1(a/\xi_t)}{\sinh(a/\xi_t)} \right)^2 \end{aligned} \quad (1)$$

and becomes independent of i and j . As the force becomes so large such that $\xi_t \ll a$, eq 1 for *any* two bonds reduces to

$$\langle \cos \theta \rangle_{\mathcal{H}_S} \gtrsim \exp\left(-\frac{(\pi+4)\xi_t}{2a}\right) \quad (2)$$

On the other hand, the contributions due only to the intrachain interactions lead to

$$\langle \cos \theta \rangle_{\mathcal{H}_c}(s) = \exp\left(-\frac{s}{\xi_p}\right) \quad (3)$$

where \mathcal{H}_c is the chain Hamiltonian of *any form* and $s = |i - j|a$ is the arc-length. The expression in eq 3, which defines the persistence length, is assumed to be true for many types of uncharged biopolymers and for polyelectrolytes at sufficiently high counterion concentrations. At high forces, the contribution from the tensile energy to the average in eq 2 becomes comparable to the contribution from the chain interaction. Thus, the arc-length s at which the two contributions

equal decreases as the force increases. Eventually, as the force approaches \mathcal{F}_h such that $s = a$, we obtain

$$\exp\left(-\frac{(\pi+4)\xi_t}{2a}\right) \lesssim \langle \cos \theta \rangle_{\mathcal{H}_S}(\mathcal{F}_h) = \exp\left(-\frac{a}{\xi_p}\right) \quad (4)$$

which leads to an estimate of \mathcal{F}_h

$$\mathcal{F}_h = c \frac{k_B T \xi_p}{a^2} \quad (5)$$

with the constant c in the range from $(4 + \pi)/2$ to 4 (see below). For $s > a$, the average cosine is determined by the tensile energy \mathcal{H}_S whose effects on the chain are (much) larger than intrachain interactions. At such high forces, the bonds align themselves along the force with small fluctuations. At forces above \mathcal{F}_h , the stretching equation of a discrete semiflexible chain satisfying the above bond angle assumption becomes that of the FJC, i.e.

$$x = 1 - \frac{\xi_t}{a} \quad (6)$$

Thus, \mathcal{F}_h should be viewed as a crossover force because above \mathcal{F}_h all biopolymers should behave as FJCs. Note that, in order for \mathcal{F}_h to be finite, a in eq 5 must not vanish, i.e., the polymer chain should be viewed as a series of connected discrete links.

A similar form and notion of the crossover force were already pointed out elsewhere^{29,30} for two specific models, WLC and freely rotating chain (FRC). However, either the *explicit* form of the Hamiltonian (WLC) was required or

simulations for FRC were needed to infer an empirical value of \mathcal{F}_h . Here, general arguments based on the interplay between the force and intrinsic chain stiffness alone leads to an estimate of \mathcal{F}_h (eq 5). More importantly, \mathcal{F}_h can be obtained without assuming any specific energy function for the polymer and hence should be considered as a universal quantity for any discrete semiflexible chain.

The critical value \mathcal{F}_h defined in eq 5 could, in principle, be somewhat more complicated especially if the bonds connecting the monomers are extensible. In this case, the stretching equation at high f for the FJC (eq 6) has the form^{3,5,31,32}

$$x = 1 - \frac{\xi_t}{a} + \frac{f}{K} \quad (7)$$

where K is the stretch modulus that has the dimension of force.

2.2. Force-Dependent Kuhn Segments and the Universal Effective FJC Regime. Consider the regime

$$\frac{a^2}{c\xi_p} < \xi_t < \xi_p$$

(the lower bound is when $f \approx \mathcal{F}_h$). Similar to the Pincus argument at low force ($\xi_t > l_K$, the Kuhn length) in which the chain can be viewed as a series of blobs, the polymer breaks up into a series of segments whose length is *on the order of* the Kuhn length and behaves *almost* similarly to a FJC of Kuhn segments (see Figure 1). The relative extension in this force regime is, at most, $x \sim 1 - \xi_t/l_K$. However, we are not in the true FJC regime discussed above. The intrinsic intrachain interactions, which tend to prevent the chain from bending, are relevant. The relative extension x must be larger than that of the FJC with bond length a , namely $1 - \xi_t/a$. Thus, we have the bound

$$1 - \frac{\xi_t}{a} \leq x \leq 1 - \frac{\xi_t}{l_K}$$

In analogy with eq 6, we can formally write

$$x = 1 - \frac{\xi_t}{\lambda(f)} \quad (8)$$

where $\lambda(f)$ is the segment length of the effective FJC, which should be regarded as an effective force-dependent Kuhn length.

If the chain is viewed as a sequence of connected segments with effective length $\lambda(f)$ that resembles the FJC, the average cosine of the angle formed by any two of these λ -segments is given by eq 2 where a is replaced by λ :

$$\langle \cos \theta \rangle_{H_s}(\lambda) \gtrsim \exp\left(-c \frac{\xi_t}{\lambda}\right) \quad (9)$$

The assumption that $\xi_t \ll \lambda$ has to be justified *a posteriori*. The intrinsic tangent–tangent correlation of any two i and j effective monomer-like λ -segments is

$$\langle \cos \theta \rangle_{H_b}(s) = \exp\left(-\frac{|i-j|\lambda}{\xi_p}\right) \quad (10)$$

By equating the two eqs 9 and 10 (with $|i-j| = 1$) we can relate λ , f and ξ_p using

$$\lambda \approx \lambda_K \equiv \sqrt{c\xi_t\xi_p} \quad (11)$$

It follows from eq 11 that the force-dependent Kuhn segment λ_K is determined not only by the intrinsic stiffness of the chain through ξ_p but also by the tension through ξ_t .

The tensile screening length, which satisfies $\xi_t(\propto f^{-1}) \ll \lambda(f)$ ($\propto f^{-1/2}$) is similar to the *deflection length* proposed by Odijk³¹ using the explicit quadratic form of the Hamiltonian. The FEC in this regime satisfies,

$$x \approx 1 - \sqrt{\frac{\xi_t}{c\xi_p}} \quad (12)$$

which coincides with the results for the WLC when $c = 4$.^{1,29,33} Thus, by merely using the analogy to a semiflexible chain, which behaves as a FJC at high force and considering the effects of the intrinsic segment correlation and the effects of tension, we have obtained the characteristic stretching equation for *any* semiflexible chain. Such a regime exists in all semiflexible chains that satisfy the bond angle assumption. We shall refer to this as the effective FJC or WLC regime, generalizing the traditional name for any semiflexible chain. In practical applications the terms effective FJC and WLC can be used interchangeably.

2.3. The Unified Stretching Equation for both the FJC and WLC. For the two regimes FJC and WLC discussed above, we assume that the generalized equation for λ takes the simple form

$$\lambda(f, \alpha) \approx (\lambda_K^\alpha + a^\alpha)^{1/\alpha} \quad (13)$$

where $\alpha > 1$ is a model-dependent parameter. This equation reflects the observation that when $f < \mathcal{F}_h$, $\lambda(f)$ is almost λ_K ($\gg a$) and the chain behaves as an effective FJC; and when $f > \mathcal{F}_h$, $\lambda(f)$ tends to a ($\gg \lambda_K$) and the chain behaves as a true FJC. Thus, we have the equation for the two strong stretching regimes

$$x \approx 1 - \frac{\xi_t}{(\lambda_K^\alpha + a^\alpha)^{1/\alpha}} \quad (14)$$

which reduces exactly to the form for the discrete WLC²⁹ when $\alpha = 2$.

Because of the upper limit of λ , which is $l_K = 2\xi_p$, the lower limit on the force for observing the effective FJC regime (assuming $c \leq 4$) is

$$\lambda = l_K \Rightarrow \xi_t \approx \xi_p \Rightarrow \mathcal{F}_l \approx \frac{k_B T}{\xi_p} \quad (15)$$

Here, \mathcal{F}_l can be called the crossover force to the effective FJC regime from a low force regime (see Figure 1).

The physics of the stretching of all biopolymer molecules using discrete semiflexible chains is summarized in Figure 1. At very low forces, the extension increases linearly with the force.²⁵ For larger forces, but not too large, such that the tensile screening length is still significantly larger than the Kuhn length, one reaches the Pincus regime of blobs where the extension scales either linearly with the force (ideal chains) or as $f^{2/3}$ (self-avoiding chains).^{18,24} As the force increases until it is larger than \mathcal{F}_l , ξ_t becomes smaller than the Kuhn length, the chain behaves as an effective FJC with the law $1 - x \propto f^{-1/2}$. Eventually, when ξ_t becomes smaller than a^2/ξ_p ($f > \mathcal{F}_h$), the chain behaves as a true FJC under high tension with the FEC law $1 - x \propto f^{-1}$.

The theory provided here differs from previous studies in several aspects. (i) We obtain the universal behavior of response of any discrete semiflexible chain, regardless of the specific Hamiltonian for the polymer. In this sense, our theory is model-free, which allows us to analyze a broad range of experiments. (ii) The unified theory embodied in eq 14 automatically leads to the $1/2$ -rule from which ξ_p can be

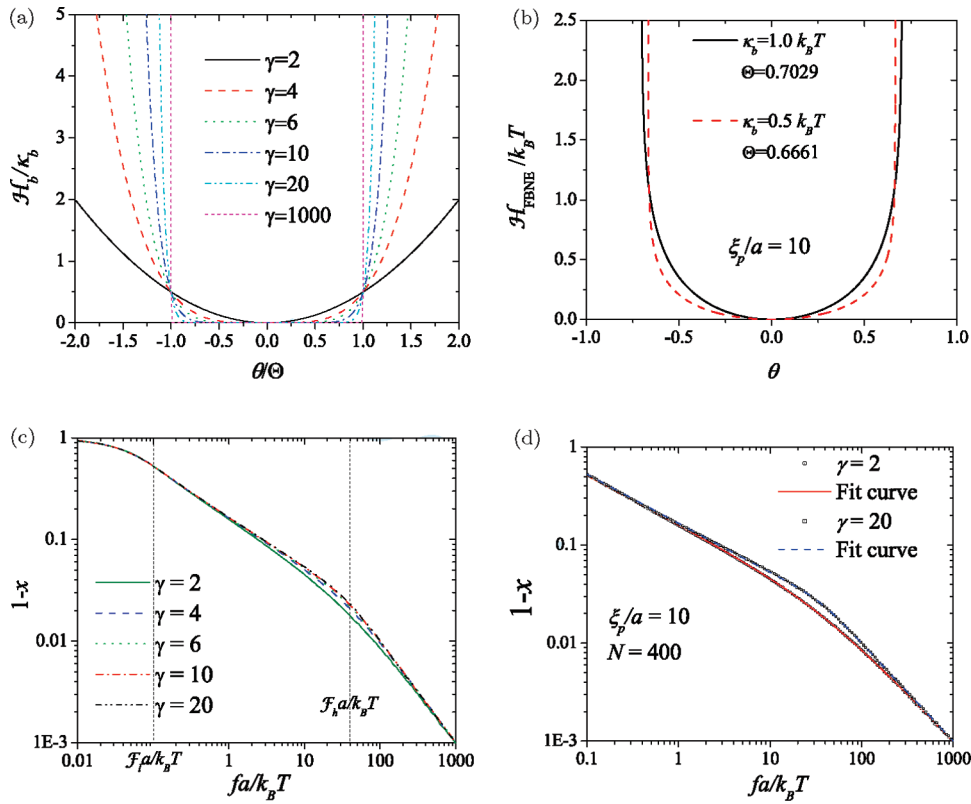


Figure 2. Simulation of model Hamiltonians and fits to the simulation data. (a) Plot of \mathcal{H}_b/k_b as a function of θ/Θ at various values of γ (eq 16). As γ increases from 2, which is approximately the value for the WLC model, the potential mimics the excluded volume interaction between consecutive monomers. (b) Plot of $\mathcal{H}_{FBNE}(\theta)/k_B T$ in eq 17 at two sets of values of κ_b and Θ with $\xi_p/a = 10$. (c) Log–log plot of $1 - x$ as a function of $fa/k_B T$ obtained from Monte Carlo simulations for the polynomial model with $\xi_p/a = 10$ and $N = 400$ and various values of γ . The scaling laws $1 - x \propto f^{-1/2}$ and $1 - x \propto f^{-1}$, as well as the two crossover forces \mathcal{F}_l and \mathcal{F}_h are clearly visible. (d) The fits of eq 14 to the simulation data for $\gamma = 2$ and 20.

analyzed. Indeed, this method of analysis allows us to estimate the dependence of persistence length on ionic strength for ssDNA (see below).

3. Simulations Confirm the Theoretical Predictions

To test our theory, we performed simulations using a number of models.

“Polynomial” Models (“Generalized” WLC). The Hamiltonian of the chain with N bonds is

$$\mathcal{H}_b = \frac{\kappa_b}{2} \sum_{i=1}^N \left(\frac{\theta}{\Theta} \right)^\gamma \quad (16)$$

where θ , the angle between two successive bonds, varies from 0 to π . When $\gamma = 2$ and $\Theta = 1$, the Hamiltonian approximately reduces to the WLC energy function and when γ is large, we recover the local approximation Thick Chain model.^{34,35} Figure 2(a) shows the dependence of \mathcal{H}_b on θ/Θ for several values of γ .

The FBNE (Finite Bendable Nonlinear Elastic) Chain. To model the finite bendability of a chain, we use the FBNE potential

$$\mathcal{H}_{FBNE} = \begin{cases} -\frac{\kappa_b}{2} \sum_{i=1}^N \log \left(1 - \left(\frac{\theta}{\Theta} \right)^2 \right), & \theta < \Theta; \\ \infty, & \theta \geq \Theta \end{cases} \quad (17)$$

which resembles the FENE potential^{36,37} for chain bonds.

To compare all simulation data, we chose parameters in the models such that the persistence length is fixed at $\xi_p/a = 10$.

The ratio $\mathcal{F}_h/\mathcal{F}_l \approx c(\xi_p/a)^2 = 100c$ is large enough to clearly observe the effective FJC regime. From eqs 12 and 6, it can be seen that plotting $1 - x$ against $f \times a/k_B T$ in double logarithmic scale would give us the characteristic slopes of $-1/2$ and -1 for the effective FJC and FJC regimes, respectively. Indeed, parts C and D of Figure 2 show that there are these characteristic slopes for the “polynomial” model at all γ considered. The same result is obtained for the FBNE and *self-avoiding* TC models^{25,34,35} (data not shown). Moreover, the fits using the generalized eq 14 with *two* free parameters c and α are excellent and yield, in accord with theoretical arguments, c in the range $(4 + \pi)/2$ to 4, and α in the range 2 to 7 and positively correlated with γ . Moreover, it is easily seen that the effective FJC regime sets in as soon as

$$f \approx 0.1 \frac{k_B T}{a} = \frac{a}{\xi_p} \frac{k_B T}{a} \equiv \mathcal{F}_l$$

The crossover from effective FJC to FJC occurs at

$$f \approx 40 \frac{k_B T}{a} \approx \frac{c \xi_p}{a} \frac{k_B T}{a} \equiv \mathcal{F}_h$$

Thus, the simulations confirm the theoretical arguments.

4. Analysis of Experiments

4.1. Significance of the Regimes and Crossover Forces. Our analysis provides a theoretical basis for choosing the appropriate polymer model to analyze the FECs of biopolymers. In order to choose a specific model it is necessary to estimate \mathcal{F}_l and \mathcal{F}_h , which require the values of a and ξ_p (see eqs 15 and 5) that are intrinsic properties of the biopolymer. The value

of the monomer spacing a can be obtained by analyzing the FEC at high force (FJC regime). The effects of the self-avoidance of the chain can be assessed most accurately in the linear and Pincus regimes (Figure 1). Chain extensibility can make it difficult to estimate a making it necessary to include models that treat chain extensibility explicitly.^{3,31,32} The persistence length is easily extracted from the FECs in the effective FJC regime. It is the most accurate quantity that can be inferred from experiments, since it is the effective FJC regime that is commonly detectable in those experiments.^{2-7,38-40} In practice, we use the force at $x \approx 0.5$ (referred to as $1/2$ -rule) to estimate $\mathcal{F}_l = k_B T / \xi_p$, and hence ξ_p (see eq 12 with $c \approx 4$). We illustrate the utility of our theory that depends on the interplay of the two “crossover” forces \mathcal{F}_l and \mathcal{F}_h in determining the elasticity of biopolymers by analyzing measured FECs.

dsDNA. Using the well-known values $\xi_p \approx 50 \text{ nm}^{1-3,41}$ and $a = 0.34 \text{ nm}$ at $T = 300 \text{ K}$, we obtain $\mathcal{F}_l \approx 0.1 \text{ pN}$ and $\mathcal{F}_h \approx 6400 \text{ pN}$. Values of f that are comparable to \mathcal{F}_l are accessible only using magnetic beads or optical tweezers,^{2-4,38} whereas $f \sim \mathcal{F}_h$ can be realized in AFM experiments. Because dsDNA undergoes a phase transition to an overstretching form at $f \approx 65 \text{ pN}$,⁴²⁻⁴⁵ the FJC limit cannot be observed in dsDNA. It should be stressed that even if a is greatly increased, up to the conceivably maximum value of 3.4 nm or one double helix turn, \mathcal{F}_h still exceeds $f \approx 65 \text{ pN}$. These estimates show that the FEC of dsDNA (excluding the overstretching region) should be in only the effective FJC regime, which explains the success of the WLC in quantitatively describing the FEC of dsDNA.

ssRNA. From the typical values of $\xi_p \approx 1 \text{ nm}^{9,10}$ at sufficiently high counterion concentrations and $a \approx 0.55 \text{ nm}^{46}$ for both ssRNA and ssDNA we obtain $\mathcal{F}_l \approx 4 \text{ pN}$ and $\mathcal{F}_h \approx 50 \text{ pN}$. Thus, both \mathcal{F}_l and \mathcal{F}_h are small enough that they can be accessed in typical LOT experiments, making it practical to infer the f -dependent changes in the elasticity. The values of \mathcal{F}_l and \mathcal{F}_h readily explain why the measured FEC for poly-U using optical tweezers⁵ in the force range from about 0.5 pN to 50 pN , which is about \mathcal{F}_h , is best described using the WLC model. We predict that the FJC behavior should emerge for forces larger than 50 pN for poly-U. The same arguments also hold good for ssDNA data at high ionic concentrations in the following paper by Saleh et al.

ssDNA. In a recent experiment, Saleh et al. measured FEC of ssDNA at varying counterion concentrations in the force range $0.1 \leq f \leq 50 \text{ pN}$.²⁷ The data suggests that for $f < f_c$, a critical force that depends on the salt concentration, the FEC is in the Pincus regime (Figure 1) in which the nonlinear response to force is determined largely by the excluded-volume interactions. We adopt the $1/2$ -rule to calculate $\xi_p(\lambda_D)$ as a function of the Debye length λ_D for ssDNA and poly-U. The estimated $\xi_p(\lambda_D)$ for these two polyelectrolytes are identical (Figure 3a). Moreover, we can fit the curves with $\xi_p(\lambda_D) = \xi_p^0 + A\lambda_D$, where ξ_p^0 is the “bare” persistence length and A is a proportionality constant. The values of A and ξ_p^0 for poly-U and ssDNA are nearly identical ($\xi_p^{0,\text{poly-U}} = 0.67 \text{ nm}$ and $\xi_p^{0,\text{ssDNA}} = 0.63 \text{ nm}$). The linear dependence of ξ_p on λ_D is in agreement with previous works^{47,48} as well as Saleh’s findings.²⁷ Although the procedure used by Saleh et al. to extract $\xi_p(\lambda_D)$ is different from the WLC-based analysis presented here, our estimated $\xi_p^{0,\text{ssDNA}}$ is in excellent agreement with their estimate of 0.6 nm .²⁷

Comparison of our predictions with the FEC data for ssDNA reported by Rief et al.⁴⁹ in the force range from 10 pN to 200 pN and in Tris buffer further validates the present analysis. In a different paper,⁸ Dessinges et al. combined the data with their data on ssDNA in the force range from

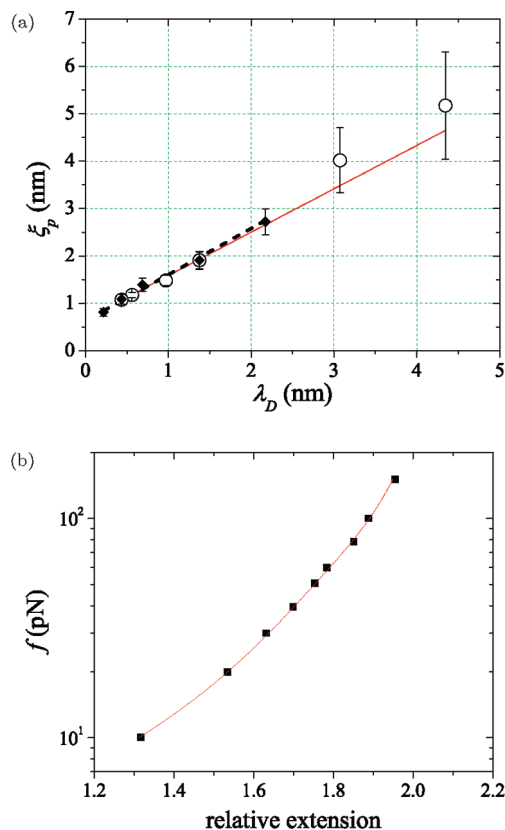


Figure 3. Analysis of the experimental results. (a) Estimates of $\xi_p(\lambda_D)$ for poly-U (circles) and ssDNA (filled symbols) using the $1/2$ -rule. Linear fits to the data (solid and dashed lines) yielded the bare persistence lengths $\xi_p^{0,\text{poly-U}} \approx 0.67 \text{ nm}$ and $\xi_p^{0,\text{ssDNA}} \approx 0.63$. (b) The fit of eq 14 (solid line) to the FEC of a ssDNA (squares) by Rief et al.⁴⁹ The extracted values for the parameters are: $\xi_p \approx 0.8 \text{ nm}$, $a \approx 0.5 \text{ nm}$ and $\alpha \approx 7$ leading to $\mathcal{F}_l \approx 5 \text{ pN}$ and $\mathcal{F}_h \approx 55 \text{ pN}$. The extracted values of ξ_p and a are consistent with those reported in refs 10 and 46, respectively. The value of $\alpha \approx 7$ implies that γ (see eq 16) is large. As a result, the bending energy increases more sharply than predicted by the WLC model (Figure 2a) as θ increases beyond a critical value.

0.05 pN to 20 pN and in a different buffer (see Figure 4 in ref 8). Using extensive data on ssDNA, they showed that at very low force, self-avoidance and electrostatic effects play dominant roles, and hence the elasticity of ssDNA is qualitatively different from the high force regimes. Indeed, by taking into account the electrostatic interactions in the so-called modified (extensible) FJC model, Dessinges et al. explained the low force part of the FEC. They also showed that *both* the WLC and the “bare” modified FJC, when applied independently, failed to fit the *entire* high force part ($x \geq 1/2$) of the FEC. Here we show that the FEC above $f = \mathcal{F}_l$ (data by Rief et al.⁴⁹) can be quantitatively analyzed using our theory. The equation for the high stretch regimes (eq 14), which has both the WLC and FJC high force behaviors, provides a very good fit to the data. We notice that the modified FJC model by Dessinges et al. with electrostatic interactions actually failed to fit the high force ($x > 1/2$) portion of the FEC.⁸ It could be argued that the good fit to the lower portion of the FEC was mainly because the electrostatic interactions, which play dominant roles, were well taken into account. The fact that whether the FJC or WLC is used for the low force regime should not considerably affect the outcome of the fit. But for the high force regime, it is essential to do a force range analysis prior to the actual fit, as it is done here. Our extracted values are $a \approx 0.5 \text{ nm}$, $\xi_p \approx 0.8 \text{ nm}$, and $\alpha \approx 7$ leading to $\mathcal{F}_l \approx 5 \text{ pN}$ and $\mathcal{F}_h \approx 55 \text{ pN}$ (see Figure 3b).

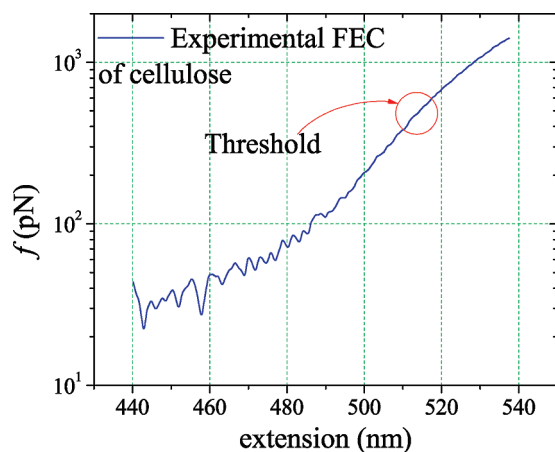


Figure 4. Illustration of overstretching in cellulose. A threshold is identified as a narrow region in the circle, where it shows a decrease in the slope of the FEC when plotted in log–linear scale.

Protein (PEVK Domain). PEVK (Pro-Glu-Val-Lys) domain is a proline-rich domain of the muscle protein titin and it has no definite folded structure. Using the accepted values $\xi_p \approx 0.5$ nm^{6,7,39,40} and $a = 0.38$ nm, $\mathcal{F}_l \approx 8.4$ pN, which is well within the covered range of almost all single molecule apparatuses, and $\mathcal{F}_h \approx 50$ pN. We expect that the WLC is more suitable for the PEVK domain than the FJC, and indeed FECs in the force range from 10 to 200 pN were analyzed using WLC.^{6,7,39,40} However, we can see the need for a discrete model rather than the continuum WLC to obtain a better fit.

Polysaccharides. Linear polysaccharides that occur naturally in a variety of cellular structures are subject to tensile stress. At high forces the sugar rings undergo a transition from chair to boat conformation, which clearly changes their elastic properties.^{12,13} Unlike other systems (dsDNA, for example) the change in elasticity is associated with enthalpic changes as a result of the conformational transition. Surprisingly, it is found that over a wide range of forces the FEC is best explained by the FJC model,¹⁴ which can be fully explained using our theory. For cellulose with the value $l_K \approx a \approx 0.54$ nm,^{11–14} we obtain $\mathcal{F}_l \approx \mathcal{F}_h \approx 15.3$ pN. This means that excluding the overstretch range, the entire range of FEC for cellulose reported in references by Marszalek et al.^{11–14} obtained using AFM with force range of 10 to 1000 pN is best described using the FJC. The same observation can be made for other polysaccharides, such as dextran ($l_K \approx a \approx 0.44$ nm in the chair conformation and $l_K \approx a \approx 0.57$ nm in the boat conformation) and amylose ($l_K \approx a \approx 0.45$ nm (chair) and $l_K \approx a \approx 0.54$ nm (boat)).¹² Thus, the intrinsic properties of polysaccharides show that this class of molecules behave as FJCs.

In all the above examples, it is necessary to assess the range of force in any particular FEC before using the polymer model that can best describe the data. In most cases, it is necessary to consider both the semiflexibility and discreteness of the biopolymer. Models such as the discrete WLC model²⁹ and the TC model^{34,35} are better for most fitting purposes. In analyzing most of the currently available data we find that these two models are adequate. However, as shown by Dessinges et al.⁸ the variations in the elasticity of ssDNA over a wide range of forces will require a combination of models. Our theory accounts for f -dependent variations in the elasticity in terms of naturally occurring length scales as long as $x \geq 1/2$. Below $x = 1/2$, additional forces may stabilize specific structures.

4.2. Overstretching in a FEC. For each biopolymer, there is a threshold force at which the transition to bond overstretching

occurs. Inextensible models can be used to fit only the experimental FECs below the threshold force. From the approximate eq 14 in the high stretch regime it can be shown that the slope of the FEC above \mathcal{F}_l in the log–linear scale is a strictly monotonically increasing quantity for a generic semiflexible chain. On the other hand, in the region where overstretching effect is significant, the extension can be approximated as³¹

$$x \approx 1 - \frac{C}{f^{1/m}} + \frac{f}{K}$$

where K , m , and C are positive constants. It can be shown that the rate of change of the slope in the region $x > 1$ (overstretched) is always negative. Then for a FEC with overstretching, the threshold force is the one ($> \mathcal{F}_l$) at which the slope of the curve $\ln f(x)$ vs x stops increasing. A slope that stays constant or decreases would signify the failure of the inextensible model.

The simple analysis, that accounts for extensibility through the stretch modulus K , allows us to gauge the threshold force from the FEC. If the data are precise, we can take the numerical derivative of $\ln f(x)$ in the region $f > \mathcal{F}_l$ and detect the force at which it reaches a maximum value or a plateau. Otherwise, the practical solution would be to plot $f(x)$ vs x in the log–linear plane then detect the change in the slope above \mathcal{F}_l by inspection. Although not rigorous, it turns out to be an effective way as demonstrated in Figure 4. For the poly-U data⁵ at various salt concentrations and at forces $f \leq 50$ pN, the practical technique reveals no decrease in the slope in the $x > 1/2$ region (data not shown). Therefore, we expect no overstretching in poly-U or ssDNA.

5. Conclusions

Our theory, which is valid as long as enthalpic interactions are negligible, shows that the WLC and FJC models at large forces are the two limiting universal behaviors for virtually all polymers and especially biopolymers. The crossover from WLC to FJC is determined by the persistence length ξ_p and the monomer spacing a . The validity of the theory is demonstrated through the excellent agreement with simulation results and the successful analysis of the force extension curves of a diverse set of biopolymers. The two crossover forces, \mathcal{F}_h and \mathcal{F}_l , which demarcate the boundary between the applicability of the WLC and FJC, provide a physical picture of the f -dependent elasticity of biopolymers. The derived $1/2$ -rule is a convenient way to obtain a good estimate of the persistence length of the biopolymer from the measured FEC, which in turn yields the model that is most appropriate for analyzing the FEC. The theory also shows that complete understanding of the elasticity of biopolymers requires measurement of FECs for a wide range of forces as obtained for ssDNA.⁸

Note Added in Proof: While this paper was in press, Dr. J. Kierfeld brought to our attention an article by J. Kierfeld et al. (Kierfeld, J.; Niamplöy, O.; Sa-yakanit, V.; Lipowsky, R. Stretching of semiflexible polymers with elastic bonds, *Eur. Phys. J. E* **2004**, *14*, 17–34) in which the semiflexible harmonic chain model with explicit quadratic form of the Hamiltonian was used to obtain expressions for crossover forces similar to those obtained here.

Acknowledgment. We are grateful to Changbong Hyeon, Greg Morrison, and Omar Saleh for useful discussions. This work was supported by a grant from the National Science Foundation (CHE 09-14033).

References and Notes

- (1) Marko, J. F.; Siggia, E. D. *Macromolecules* **1995**, *28*, 8759–8770.
- (2) Wang, M. D.; Yin, H.; Landick, R.; Gelles, J.; Block, S. M. *Biophys. J.* **1997**, *72*, 1335–1346.
- (3) Bouchiat, C.; Wang, M. D.; Allemand, J. F.; Strick, T.; Block, S. M.; Croquette, V. *Biophys. J.* **1999**, *76*, 409–413.
- (4) Smith, S. B.; Finzi, L.; Bustamante, C. *Science* **1992**, *258*, 1122–1126.
- (5) Seol, Y.; Skinner, G.; Visscher, K. *Phys. Rev. Lett.* **2004**, *93*, 118102.
- (6) Linke, W. A.; Kulke, M.; Li, H.; F.-Becker, S.; Neagoe, C.; Manstein, D. J.; Gautel, M.; Fernandez, J. M. *J. Struct. Biol.* **2002**, *137*, 194–205.
- (7) Li, H.; Oberhauser, A. F.; Redick, S. D.; C.-Vazquez, M.; Erickson, H. P.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10682–10686.
- (8) Dessinges, M.-N.; Maier, B.; Zhang, Y.; Peliti, M.; Bensimon, D.; Croquette, V. *Phys. Rev. Lett.* **2002**, *89*, 248102.
- (9) Caliskan, G.; Hyeon, C.; Perez-Salas, U.; Briber, R. M.; Woodson, S. A.; Thirumalai, D. *Phys. Rev. Lett.* **2005**, *95*, 268303.
- (10) Rivetti, C.; Walker, C.; Bustamante, C. *J. Mol. Biol.* **1998**, *280*, 41–59.
- (11) Marszalek, P. E.; Pang, Y.-P.; Li, H.; Yazal, J. E.; Oberhauser, A. F.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 7894–7898.
- (12) Marszalek, P. E.; Oberhauser, A. F.; Pang, Y.-P.; Fernandez, J. M. *Nature* **1998**, *396*, 661–664.
- (13) Marszalek, P. E.; Li, H.; Fernandez, J. M. *Nat. Biotechnol.* **2001**, *19*, 258–262.
- (14) Lee, G.; Nowak, W.; Jaroniec, J.; Zhang, Q.; Marszalek, P. *Biophys. J.* **2004**, *87*, 1456–1465.
- (15) Rief, M.; Fernandez, J. M.; Gaub, H. E. *Phys. Rev. Lett.* **1998**, *81*, 4764–4767.
- (16) Ke, C.; Humeniuk, M.; S-Gracz, H.; Marszalek, P. *Phys. Rev. Lett.* **2007**, *99*, 018302.
- (17) Seol, Y.; Skinner, G. M.; Visscher, K.; Buhot, A.; Halperin, A. *Phys. Rev. Lett.* **2007**, *98*, 158103.
- (18) Pincus, P. *Macromolecules* **1976**, *9*, 386–388.
- (19) Danilowicz, C.; Lee, C. H.; Coljee, V. W.; Prentiss, M. *Phys. Rev. E* **2007**, *75*, 030902.
- (20) Klimov, D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6166–6170.
- (21) Guttmann, A. J.; Jacobsen, J. L.; Jensen, I.; Kumar, S. *J. Math. Chem.* **2009**, *45*, 223–237.
- (22) Krawczyk, J.; Jensen, I.; Owczarek, A. L.; Kumar, S. *Phys. Rev. E* **2009**, *79*, 031912.
- (23) Flory, P. *Statistical mechanics of chain molecules*; Hanser Publisher: Munich, Germany, 1989.
- (24) de Gennes, P.-G. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca, NY, 1979.
- (25) Toan, N. M. Ph.D. Thesis, SISSA: Trieste, Italy, **2006**.
- (26) Morrison, G.; Hyeon, C.; Toan, N. M.; Ha, B. Y.; Thirumalai, D. *Macromolecules* **2007**, *40*, 7343–7353.
- (27) Saleh, O. A.; McIntosh, D.; Pincus, P.; Ribbeck, N. *Phys. Rev. Lett.* **2009**, *102*, 068301.
- (28) Netz, R. R. *Macromolecules* **2001**, *34*, 7522–7529.
- (29) Rosa, A.; Hoang, T. X.; Marenduzzo, D.; Maritan, A. *Macromolecules* **2003**, *36*, 10095–10102.
- (30) Livadaru, L.; Netz, R. R.; Kreuzer, H. J. *Macromolecules* **2003**, *36*, 3732–3744.
- (31) Odijk, T. *Macromolecules* **1995**, *28*, 7016–7018.
- (32) Storm, C.; Nelson, P. C. *Phys. Rev. E* **2003**, *67*, 051906.
- (33) Ha, B. Y.; Thirumalai, D. *J. Chem. Phys.* **1997**, *106*, 4243–4247.
- (34) Toan, N. M.; Marenduzzo, D.; Micheletti, C. *Biophys. J.* **2005**, *89*, 80–86.
- (35) Toan, N. M.; Micheletti, C. *J. Phys.: Condens. Matter* **2006**, *18*, S269–S281.
- (36) Kremer, K.; Grest, G. S. *J. Chem. Phys.* **1990**, *92*, 5057–5086.
- (37) Hyeon, C.; Dima, R. I.; Thirumalai, D. *Structure* **2006**, *14*, 1633–1645.
- (38) Strick, T. R.; Allemand, J.-F.; Bensimon, D.; Bensimon, A.; Croquette, V. *Science* **1996**, *271*, 1835–1837.
- (39) Li, H.; Linke, W.; Oberhauser, A.; Carrion-Vazquez, M.; Kerkvliet, J.; Lu, H.; Marszalek, P.; Fernandez, J. *Nature* **2002**, *418*, 998–1002.
- (40) Fisher, T. E.; Oberhauser, A. F.; Carrion-Vazquez, M.; Marszalek, P. E.; Fernandez, J. M. *Trends Biochem. Sci.* **1999**, *24*, 379–384.
- (41) Baumann, C. G.; Smith, S. B.; Bloomfield, V. A.; Bustamante, C. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6185–6190.
- (42) Smith, S. B.; Y., J. C.; Bustamante, C. *Science* **1996**, *271*, 795–799.
- (43) Cluzel, P.; Lebrun, A.; Heller, C.; Lavery, R.; Viovy, J.-L.; Chatenay, D.; Caron, F. *Science* **1996**, *271*, 792.
- (44) Wuite, G. J.; Smith, S. B.; Young, M.; Keller, D.; Bustamante, C. *Nature* **2000**, *404*, 103–106.
- (45) Bustamante, C.; Bryant, Z.; Smith, S. B. *Nature* **2003**, *421*, 423–427.
- (46) Hyeon, C.; Dima, R. I.; Thirumalai, D. *J. Chem. Phys.* **2006**, *125*, 194905.
- (47) Barrat, J.-L.; Joanny, J.-F. *Europhys. Lett.* **1993**, *24*, 333–338.
- (48) Dobrynin, A. V. *Macromolecules* **2005**, *38*, 9304–9314.
- (49) Rief, M.; Clausen-Schaumann, H.; Gaub, H. E. *Nature* **1999**, *6*, 346–349.