Stretching DNA

John F. Marko*

Center for Studies in Physics and Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021-6399

Eric D. Siggia

Laboratory of Atomic and Solid State Physics, Clark Hall, Cornell University, Ithaca, New York 14853-2501

Received May 9, 1995; Revised Manuscript Received September 18, 1995[⊗]

ABSTRACT: A statistical mechanical treatment of the wormlike chain model (WLC) is used to analyze experiments in which double-stranded DNA, tethered at one end, is stretched by a force applied directly to the free end, by an electric field, or by hydrodynamic flow. All experiments display a strong-stretching regime where the end-to-end distance approaches the DNA contour length as $1/(\text{force})^{1/2}$, which is a clear signature of WLC elasticity. The elastic properties of DNA become scale dependent in the presence of electrostatic interactions; the effective electric charge and the intrinsic bending elastic constant are determined from experiments at low salt concentration. We also consider the effects of spontaneous bends and the distortion of the double helix by strong forces.

I. Introduction

Perhaps the most elementary notion in polymer statistical mechanics is that to extend the ends of a long, linear flexible polymer, a force must be applied. The work done by this stretching goes into reduction of the conformational entropy of the chain. Thanks to huge technical advances in manipulation of the structure of double-helix DNA, it has become feasible to measure the force vs extension of single 10–100 μm long DNAs. In recent experiments by Smith et al.,[1] one end of a DNA was attached to a surface, while the other end was attached to a 3-μm-diameter magnetic bead which was then used to put the polymer under uniform tension (Figure 1a). Rather different experiments of Schurr et al.[2] and Perkins et al.[3] anchored one end of a DNA and then stretched the polymer using either an electric field or the drag force exerted by hydrodynamic flow past the coil (Figure 1b).

In this paper, we discuss these kinds of stretching experiments from the point of view of equilibrium statistical mechanics. The experimental results are rather rich in details that can be understood quantitatively by simple analytic calculations, thanks to some special features of DNA. First, double-helix DNA, or B-DNA, is very stiff[4–6] so that the energy associated with conformational fluctuations may be modeled using merely linear elasticity of a thin, uniform rod: i.e., using the "wormlike chain" (WLC).[7] Second, self-interactions or excluded volume effects are negligible under almost all of the experimental conditions[8] (see section III.C). Lastly, single double helices of up to 100 μm long[9] may be readily obtained and manipulated: one can easily obtain the long-chain limit favored by theorists.

Previous theoretical models for single-DNA stretching[10] have ignored the thin-rod elasticity known to describe DNA bending, which is a serious omission. Over most of the experimental range (end-to-end extensions from 30 to 95% of the contour length), the difference between the end-to-end extension and the total molecule contour length goes to zero as $1/f^{1/2}$, where f is the applied force (Figure 2). The only regime where a generic polymer model (e.g., the Edwards model, or the freely-jointed chain) is appropriate to describe DNA is for very weak stretching. Section II will discuss the basic statistical mechanics of the WLC under tension. This problem was treated numerically by Fixman and Kovac,[11] and some analytical details were later discussed by Crabb and Kovac,[12] but a complete theoretical picture has been lacking and is now demanded by DNA stretching experiments.[1,3] Thanks to the high quality of experimental data and its generally good agreement with WLC elasticity, we can next shift our attention to the question of how and why the naive WLC model *fails* to describe stretching experiments.

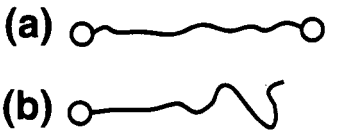

Figure 1. DNA stretching experiments are done with tethered chains either (a) by applying a force to an object ("bead") attached to the free end or (b) by applying a force to the DNA itself along its length with, e.g., electric or hydrodynamic flow fields. In (a) the tension along the DNA is uniform; in (b) the tension in the chain increases, and fluctuation decreases, as one moves from the free end to the tethered end. Tethering of the beads on the left might be accomplished by localizing them mechanically, or in an optical trap.

Section III treats electrostatic effects, important since DNA is charged; at low ionic strengths DNA is stiffened by Coulomb self-repulsion. Odijk, Skolnick, and Fixman[13] and, more recently, Barrat and Joanny[14] have shown that DNA elasticity at low ionic strength should be *scale dependent*: thus the effective persistence length should go down as the force stretching the WLC goes up. This occurs due to the reduction of the WLC fluctuation correlation length to less than the Debye screening length at high forces and turns out to be clearly observable at low ionic strengths in the experiments of Smith et al.[1] Vologodskii[15] recently used Monte Carlo simulations to show that the WLC with Debye–Hückel interactions could capture the experimentally observed trends; our analytical treatment gives further understanding of electrostatic stiffening effects. Experimental data at low ionic strength clearly indicate scale-dependent elasticity in accord with theory. Considered across a large range of ionic strengths, the data

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1995.

0024-9297/95/2228-8759$09.00/0

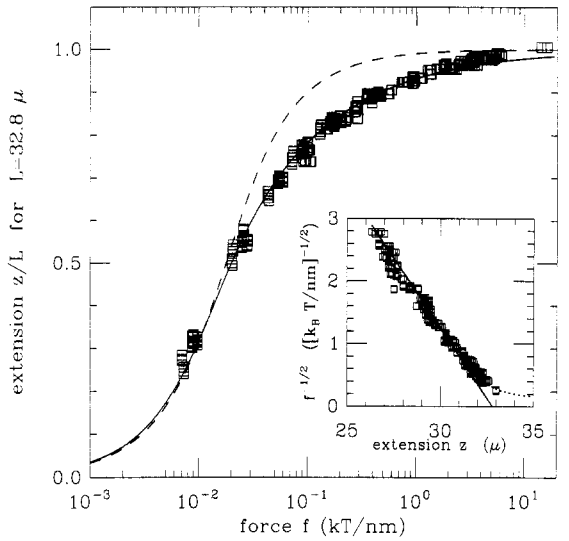

Figure 2. Fit of numerical exact solution of WLC force–extension curve to experimental data of Smith et al.[1] (97004 bp DNA, 10 mM Na^+). The best parameters for a global least-squares fit are $L = 32.8$ μm and $A = 53$ nm. The FJC result for $b = 2A = 100$ nm (dashed curve) approximates the data well in the linear low-f regime but scales incorrectly at large f and provides a qualitatively poorer fit. Inset: $f^{-1/2}$ vs z for the highest forces; the exact WLC result (solid line) is in this plot a straight line extrapolating to $L = 32.8$ μm from which the experimental points begin to diverge above $z \approx 31$ μm; including intrinsic elasticity (eq 19 with $\gamma = 500$ k_BT/nm, dotted curve) improves the fit.

also make plausible a crossover from an entropic elasticity regime to an intrinsic stretching elasticity regime (where the DNA contour length slightly increases), recently suggested by Odijk.[16] In the same section, we describe why one can largely ignore effects of excluded volume and spontaneous bends that may occur along DNA because of its heterogeneous base-pair sequence.

Section IV discusses experiments that stretch tethered DNAs with one free end (Figure 1b) with an electric field (again relying on the polyelectrolyte character of DNA) or with hydrodynamic flow.[3] Because of the complexities of dealing with a nonuniform and self-consistently determined tension, these kinds of experiments furnish less stringent tests of elastic theory but are closer to the kinds of ways DNAs and other polymers get stretched in the natural world. Finally, section V discusses recent experiments[17] showing that strong forces cause the double-helical "secondary structure" of B-DNA to abruptly lengthen by a factor of about 1.85. Although the precise nature of the new DNA state is at this time unclear (perhaps it is an extended flat ribbon or separated random-coil-like single strands), the geometry of the lengthening is consistent with straightening of the double helix, and the force scale is consistent with what is necessary to overcome the cohesive free energy binding the DNA strands together.

II. Entropic Elasticity of the Wormlike Chain

Double-helical B-DNA is a stiff-rod polymer. At length scales comparable to the double-helix repeat of 3.5 nm or the diameter of 2.1 nm, the pairing and stacking enthalpy of the bases makes the polymer very rigid, with a well-defined contour length that may be measured either in nanometers or in base pairs (1 bp = 0.34 nm).[4] DNA conformations may therefore be described by a space curve $\mathbf{r}(s)$ of fixed total length L, where s is arc length and where the tangent vector $\hat{\mathbf{t}} = \partial_s\mathbf{r}$ is a unit vector.[18,19]

A long enough linear DNA is a flexible polymer with random-walk statistics with end-to-end mean-squared distance $R_0 = (bL)^{1/2}$, where b is the Kuhn statistical monomer size (excluded volume effects can be ignored in most of the experimental data considered in this paper;[8] see section III.C). The bending costs an energy per length of $k_BTA\kappa^2/2$, where $\kappa = |\partial_s^2\mathbf{r}|$ is the curvature (the reciprocal of the bending radius) and where A is the characteristic length over which a bend can be made with energy cost k_BT. This inextensible polymer model is variously called the wormlike chain (WLC), the Kratky–Porod model, and the persistent chain model. For the WLC, $R_0^2 = 2AL$, and thus $b = 2A$.[7]

Since A is also the characteristic distance along the WLC over which the tangent vector correlations die off,[7] it is called the *persistence length*. For DNA in vivo (where there is about 150 mM Na^+ plus other ions), one should keep in mind a value $A \approx 50$ nm or 150 bp,[5] although at low ionic strengths electrostatic stiffening can cause A to appear as large as 350 nm. Throughout this paper, $L \gg A$ is always assumed.

Like any flexible polymer, separation of the ends of a DNA by an amount $z \ll L$ costs free energy $F = 3k_BTz^2/(2R_0^2)$ and therefore requires a force $f = \partial F/\partial z = 3k_BTz/(2AL)$. Below the characteristic force of k_BT/A, the extension z is small compared to L and this linear force law is valid. Since 1 k_BT/nm = 4.1 piconewtons (pN), for $A = 50$ nm, $k_BT/A = 0.08$ pN: the forces needed to extend DNAs are very small compared to the piconewtons needed to fully extend conventional polymers (e.g., polystyrene) with Kuhn length $b < 1$ nm.

For forces beyond k_BT/A, the nonlinear entropic elasticity[11,12] of the WLC model with fixed total contour length determines the force–distance behavior. Only for forces of order the base-stacking/pairing energies/length $\approx 10k_BT$/nm $= 500k_BT/A$ will the constraint of fixed arc length cease to be a good approximation, thus rendering inapplicable the WLC model (see section V).

The effective energy of a stretched WLC is[11,20]

$$\frac{E}{k_BT} = \int_0^L \frac{A\kappa^2}{2}\,ds - fz \qquad (1)$$

where the force f appears as a Lagrange multiplier to fix the end-to-end extension $z \equiv \hat{\mathbf{z}}\cdot[\mathbf{r}(L) - \mathbf{r}(0)]$. Below we will compute the equilibrium extension using the Boltzmann distribution e^{-E/k_BT}. In the remainder of this paper, forces and extensions are taken to be along the z axis, and when forces appear with inverse-length dimensions, a factor of k_BT has been suppressed.

A. Simple Calculation of WLC Strong-Stretching Behavior. When large forces are applied to a WLC, the extension approaches the total length L, and the tangent vector fluctuates only slightly around $\hat{\mathbf{z}}$.[12] From the constraint $|\hat{\mathbf{t}}| = 1$, we see that if t_x and t_y are taken as independent components, the t_z fluctuations are quadratic in the two-vector $\mathbf{t}_\perp \equiv [t_x,t_y]$, namely, $t_z = 1 - \mathbf{t}_\perp^2/2 + \mathcal{O}(\mathbf{t}_\perp^4)$. To quadratic order, $\kappa^2 = (\partial_s\mathbf{t}_\perp)^2$, and we obtain the Gaussian approximation to (1):[12]

$$\frac{E}{k_BT} = \frac{1}{2}\int_0^L ds\,[A(\partial_s\mathbf{t}_\perp)^2 + f\mathbf{t}_\perp^2] - fL \qquad (2)$$

where we have expressed the extension in (1) as $z = \int ds\, t_z$ and where terms of higher than quadratic order in $\mathbf{t}_\perp$ have been dropped.

Fourier transforms ($\tilde{\mathbf{t}}_\perp(q) \equiv \int ds\, e^{iqs}\mathbf{t}_\perp(s)$) decouple the energy into normal modes:

$$\frac{E}{k_BT} = \frac{1}{2}\int\frac{dq}{2\pi}[Aq^2 + f]|\tilde{\mathbf{t}}_\perp|^2 - fL \qquad (3)$$

The average of $\mathbf{t}_\perp^2$ at any point s is simply given by equipartition:

$$\langle \mathbf{t}_\perp^2\rangle = \int\frac{dq}{2\pi}\langle|\tilde{\mathbf{t}}_\perp(q)|^2\rangle = 2\int\frac{dq}{2\pi}\frac{1}{Aq^2+f} = 1/(fA)^{1/2} \qquad (4)$$

The leading factor of 2 in (4) counts the two components of $\mathbf{t}_\perp$. The extension is

$$\frac{z}{L} = \hat{\mathbf{t}}\cdot\hat{\mathbf{z}} = 1 - \langle \mathbf{t}_\perp^2\rangle/2 = 1 - 1/(4fA)^{1/2} \qquad (5)$$

and we see that for large forces, z approaches L with a distinctive $1/f^{1/2}$ behavior.[12] This is to be contrasted with, for example, the freely-jointed chain (FJC—a random walk of independent, fixed-length steps) model in which $L - z \sim 1/f$ for large forces.[1,11]

The large-force limit is characterized by a decreasing correlation length for the tangent vector fluctuations $\mathbf{t}_\perp$. From (3), we can just read off this correlation length as $\xi = (A/f)^{1/2}$. For low forces $f \approx k_BT/A$, $\xi \approx A$ as we would expect. For smaller forces still, the fluctuation of $\hat{\mathbf{t}}$ away from $\hat{\mathbf{z}}$ is large, and the Gaussian approximation is inapplicable.

We can expect the WLC model to be inapplicable when the force is so large that the correlation length ξ is reduced to of order the helix repeat length ≈ 3.5 nm. Using $f \approx (k_BT/A)(A/\xi)^2$, we see that for $A = 50$ nm the WLC should apply for forces up to $300k_BT/A$, or up to $z/L \approx 0.97$. This force scale is comparable to that mentioned above for the applied force to overcome base pairing and stacking interactions, $\approx 500k_BT/A$.

The side-to-side excursions of the chain over arc length s, $R_\perp^2 = \langle[\mathbf{r}_\perp(s) - \mathbf{r}_\perp(0)]^2\rangle$, can also be calculated exactly in the Gaussian limit using $\mathbf{t}_\perp \equiv \partial_s\mathbf{r}_\perp$:

$$R_\perp^2 = \frac{2}{f}\left(s - \frac{1 - \exp[-s(f/A)^{1/2}]}{(f/A)^{1/2}}\right) \qquad (6)$$

Note that if the $s \to 0$ limit of (6) is divided by s^2, (4) is rederived. Equation 6 works for forces down to $f \approx k_BT/A$, where R becomes the typical unstretched DNA coil size $R_\perp \approx (As)^{1/2}$; for much larger forces, $R_\perp \approx (As)^{1/2}/(fA)^{1/2}$.

A useful summary of these results is an approximate interpolation formula for the WLC force versus extension:

$$\frac{fA}{k_BT} = \frac{z}{L} + \frac{1}{4(1 - z/L)^2} - \frac{1}{4} \qquad (7)$$

This is asymptotically exact in the large- and small-force limits and has the scaling property that fA/k_BT is a function only of z/L (manifest already in (1) for $L \gg A$). However, the quality of the experimental data of Smith et al. requires that one fit the exact force–extension relation.[21]

B. Exact Force versus Extension of the WLC. The WLC energy (1) in terms of $\hat{\mathbf{t}}$ is

$$\frac{E}{k_BT} = \int_{s_0}^{s_1} ds\left[\frac{A}{2}(\partial_s\hat{\mathbf{t}})^2 - f\hat{\mathbf{z}}\cdot\hat{\mathbf{t}}\right] \qquad (8)$$

and is analogous to the action of a mechanical rotator $\hat{\mathbf{t}}$ of moment A for times s between s_0 and s_1 (generalized from 0 and L), subject to a polarizing field f acting in the z-direction.[11,20] Given values for $\hat{\mathbf{t}}$ at s_0 and s_1, the partition function for the WLC is

$$Z(s_1,\hat{\mathbf{t}}_1;s_0,\hat{\mathbf{t}}_0) = \int D\hat{\mathbf{t}}\,\exp[-E(s_1,s_0)/k_BT] \qquad (9)$$

where the path integral includes a constant to set the normalization $\int d^2t_1\, Z(1;0) = 1$ for all s_1, s_0, and $\hat{\mathbf{t}}_0$.

Given the probability distribution ψ for the tangent vector at s', we may compute it at s using (9):

$$\psi(\hat{\mathbf{t}},s) = \int d^2t'\, Z(s,\hat{\mathbf{t}};s',\hat{\mathbf{t}}')\psi(\hat{\mathbf{t}}',s') \qquad (10)$$

Examination of (10) in the limit $s - s' \to 0$ shows that ψ satisfies a linear Schrodinger-like equation:[11,20]

$$\partial_s\psi = \left[\frac{\mathbf{L}^2}{2A} + f\cos\theta\right]\psi \qquad (11)$$

where θ is the angle between $\hat{\mathbf{z}}$ and $\hat{\mathbf{t}}$ and where $\mathbf{L} = \hat{\mathbf{t}} \times \nabla_{\hat{\mathbf{t}}}$ is the "angular momentum" operator.

Translational invariance of (11) in s indicates that we can find eigenstates of (11):

$$\partial_s\psi = -g\psi \qquad (12)$$

For these states, $\psi(L) = e^{-gL}\psi(0)$, and (9) tells us that $\log Z(L, 0) = -gL$. The WLC free energy[22] is thus simply related to the smallest g in the spectrum of (11). The extension is computed from g via (9):

$$\frac{z}{L} = -\frac{\partial g}{\partial f} \qquad (13)$$

The right-hand side of (11) is easily diagonalized numerically (Appendix), but since we seek the minimum g, a variational approach can profitably be used. The variational wavefunction $\psi(\hat{\mathbf{t}}) \propto \exp[a\cos\theta]$ (where $\theta = \hat{\mathbf{z}}\cdot\hat{\mathbf{t}}$) reproduces the exact large- and small-f results obtained previously. For small variational parameter a, ψ is linear in $\cos\theta$ as is $f(z)$, while for large a, ψ is Gaussian in θ, which is the ground state for the oscillator described by (2).

Assuming the normalization $\int d^2t\ \psi^2(\hat{\mathbf{t}}) = 1$, the variational free energy is

$$g = \min_a\left\{\int d^2t\left[\frac{(\mathbf{L}\psi)^2}{2A} - f\cos\theta\psi^2\right]\right\}$$
$$= \min_a\left\{\left(\frac{a}{2A} - f\right)\left(\coth 2a - \frac{1}{2a}\right)\right\} \qquad (14)$$

For large- and small-force limits, the limiting behavior of (14) may be analytically obtained. For $fA \ll 1$, $a = fA$ is also small, $g = -Af^2/3$, and thus $z/L = 2fA/3$. For $fA \gg 1$, $a = (fA)^{1/2}$ diverges, $g = -f + (f/A)^{1/2}$, and therefore $z/L = 1 - (4fA)^{-1/2}$. In between, (14) may be solved numerically. The force versus extension curves for the interpolation formula (7) (which differs from the exact result in extension by about 7% for $fA \approx 1$), the solution of (14) (accurate to about 1.5%) and the numerical 'exact' result (see Appendix) are all shown in Figure 3 for comparison.

C. Fit of Pure WLC Elasticity to Experimental Data at 10 mM Salt. Adjustment of the constants L and A to obtain a least-squares fit of experimental data of Smith et al. (97004 bp DNA in 10 mM Na^+) to the exact WLC result[21] gives $L = 32.8$ μm and $A = 53$ nm. These values give a very good global fit to this data set (Figure 2) and the WLC provides a qualitatively better description of the data compared to the FJC (dashed line). In addition, the persistence length extracted this

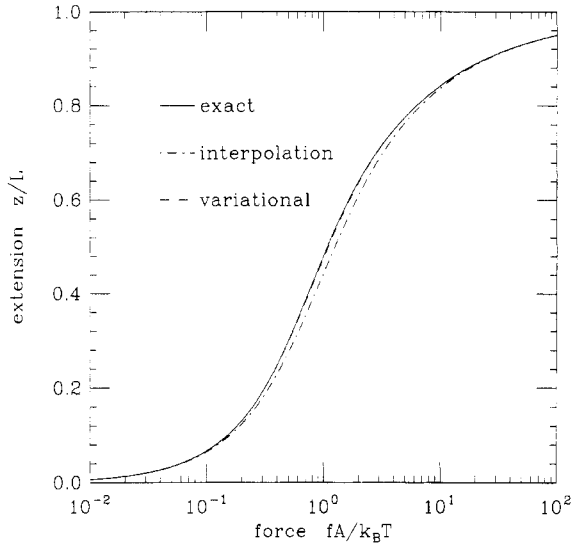

Figure 3. Comparison of three different calculations of the WLC extension z/L as a function of fA/k_BT. The solid line shows the numerical exact result (Appendix); the dashed line shows the variational solution (13–14); the dot-dashed line shows the interpolation formula (7). All three expressions are asymptotically equal for large and small fA/k_BT.

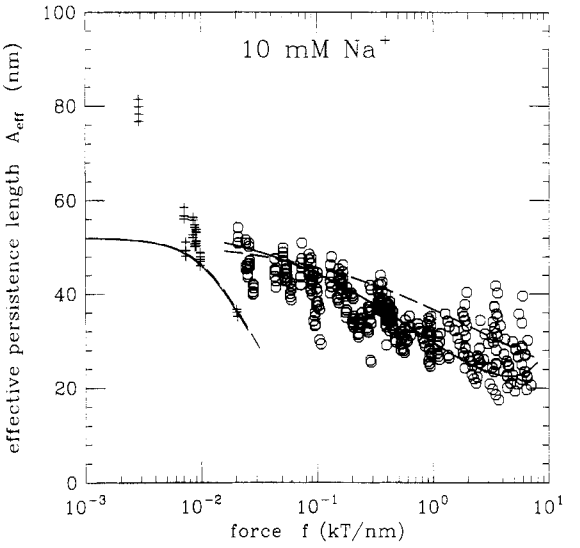

Figure 4. Apparent persistence lengths as defined in (15) for high and low forces, extracted from the 10 mM experimental data of Smith et al.[1] in Figure 2 using $L_{eff} = 33.7$ μm. The curves are the two theoretical fits of sections III (solid curves show $L = 32.8$ μm, $A_0 = 50.8$ nm, $A = 53$ nm, and $\gamma = 500$ k_BT/nm; dashed curves show $L = 33.7$ μm, $A_0 = 15$ nm, $A = 52$ nm, and γ infinite) for z vs f transformed in an identical fashion.

way matches the result of a battery of other, less direct measurements.[5,6]

The inset of Figure 2 shows that $f^{-1/2}$ goes to zero approximately linearly in z for $z \to L$ (the solid line shows the asymptotic WLC result $z/L = 1 - (k_BT/[4Af])^{1/2}$ for $A = 53$ nm and $L = 32.8$ μm). However, for the highest forces, the measured extensions exceed $L = 32.8$ μm, and $f^{-1/2}$ is not quite linear in z (note that this effect is very hard to see in the main part of Figure 2 where z/L is plotted versus f). These effects point to a failure of the pure WLC model at high forces. A possible explanation is that the B-DNA is slightly stretched by the highest tensions (e.g., consider that a slight untwisting of the double helix will slightly lengthen it; the linear twisting elasticity of B-DNA is well characterized[5,6]). Odijk has recently noted that relaxing the constraint of fixed length in the WLC model via the addition of linear stretching elasticity yields the asymptotic stretching[16] $z/L = 1 - (k_BT/[4Af])^{1/2} + f/\gamma$, where L is the total contour length of the unstressed chain and where γ is an elastic constant with dimensions of force (when $\gamma^{-1} \to 0$, the inextensible WLC result is obtained). The estimate $\gamma = 16k_BT\,A/D^2$ can be made by supposing DNA to be a homogeneous elastic rod of radius $D = 1.2$ nm;[16] taking $A = 53$ nm, we guess $\gamma \approx 600$ k_BT/nm. Turning to the data shown in the inset of Figure 2 and keeping the previously fit $L = 32.8$ μm and $A = 53$ nm, we find that $\gamma = 500$ k_BT/nm can account for the slight deviation from WLC behavior at high forces (inset of Figure 2, dotted line).

A more revealing way to plot the data and directly test the WLC model is to define an effective persistence length, A_{eff}, computed separately in the high- and low-force limits from the simple exact analytic formulas relating z to f:

$$A_{eff} = \begin{cases} 3z/(2fL_{eff}) & fA \le 1 \\ 1/[4f(1 - z/L_{eff})^2] & fA \ge 1 \end{cases} \qquad (15)$$

Figure 4 shows the 10 mM data of Figure 2 transformed using $L_{eff} = 33.7$ μm in (15). This value of L_{eff} was chosen for consistency with the low-salt data discussed in the following section. Given only the data in Figure 2 however, $L_{eff} = 33.4$ μm would make A_{eff} in Figure 4 nearly f-independent, thereby establishing that a single persistence length can describe both limits. However, this would still leave the deviations shown in the inset of Figure 2 pointing to a failure of the naive WLC model. The scatter in A_{eff} for intermediate forces is in any event independent of the choice of L_{eff} within the range $L_{eff} = 33.4$–33.7 μm. The next section shows how the gentle decrease of A_{eff} for 10 mM Na^+ (Figure 4) may be explained by electrostatic self-repulsion, either with or without intrinsic stretching elasticity.

III. Electrostatic Stiffening

Figures 5a and 5b show A_{eff}, (15), for the experimental data of Smith et al.[1] for 97 004 bp DNAs with $L_{eff} = 33.7$ μm for 1 and 0.1 mM Na^+ (Figures 6a–c show extension z versus force f for 10, 1, and 0.1 mM Na^+). The limiting low-force persistence length in these cases greatly exceeds its high-force limit, with most of the variation occurring for $z/L > 0.5$. There is comparatively little variation in A_{eff} measured for high salt (10 mM, Figure 4).

For the low ionic strengths of 0.1 and 1 mM, the electrostatic screening length λ_D falls within the range of the WLC elastic correlation length, $\xi \approx (A/f)^{1/2}$. (For the Na_2HPO_4 solution used,[1] we expect complete 2:1 Na–HPO_4 dissociation at 300 K,[23] which gives $\lambda_D \approx 0.25M^{-1/2}$ nm,[24] where M is the Na^+ molarity; M is used to label the data sets). At low forces, electrostatic self-repulsion increases the effective persistence length; for high enough forces that $\xi \ll \lambda_D$, only the "intrinsic" elastic persistence length is observed. Although the electrostatic contribution to the low-force persistence length was discussed long ago by Odijk, Skolnick, and Fixman[13] (OSF), Barrat and Joanny[14] (BJ) only recently emphasized that the OSF results indicate *scale dependence* of the persistence length. We now incorporate these electrostatic effects into the force–extension calculations of the previous section to explain the low ionic strength results of Smith et al.[1] and also examine how the resulting fits change when intrinsic stretching elasticity (see section II.C) is included.

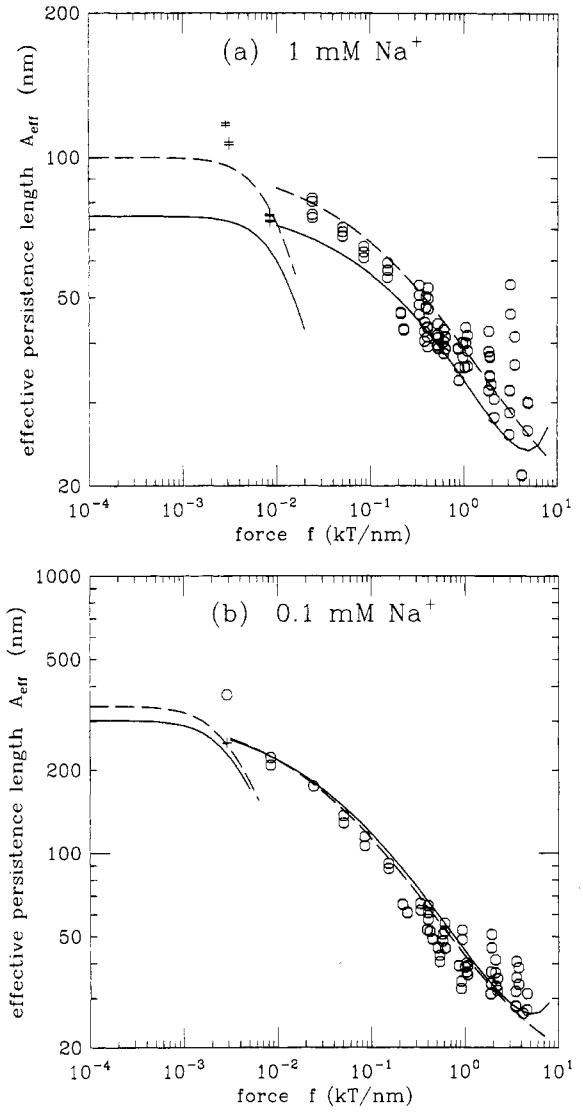

Figure 5. Apparent persistence lengths as defined in (15) for the data of Smith et al.[1] with $L_{eff} = 33.7$ μm, for (a) 1 mM and (b) 0.1 mM Na^+ solution. The high- and low-force persistence lengths are substantially different, indicating a failure of the simple WLC model. The two theoretical fits of section III transformed according to (15) are shown for $L = 32.8$ μm, $A_0 = 50.8$ nm, and $\gamma = 500$ k_BT/nm (solid lines) and $L = 33.7$ μm, $A_0 = 15$ nm, and γ infinite (dashed lines).

A. Adding Long-Ranged Interactions to the WLC. Following OSF and BJ, we introduce Debye–Hückel interactions:[13,14]

$$\frac{U}{k_BT} = \frac{l_B\nu^2}{2}\int_0^L ds \int_0^L ds' \, [v(|\mathbf{r}(s) - \mathbf{r}(s')|) - v(|s - s'|)] \tag{16}$$

The Bjerrum length $l_B = e^2/(\epsilon k_BT) = 0.7$ nm in water at 300 K with $\epsilon = 80$, and $v(r) = r^{-1}\exp[-r/\lambda_D]$. ν is the number of effective electron charges per unit length along a DNA (ν is the long-distance effective charge and should be considered as a free parameter since there is no *ab initio* theory for it, but it should be on the order of the value l_B^{-1} as suggested by the Manning condensation model[25,26]).

In the strong-stretching limit of section II.A, the Gaussian approximation to (16) is

$$\frac{U}{k_BT} = \frac{l_{OSF}}{2}\int \frac{dq}{2\pi} q^2 K(q) |\tilde{\mathbf{t}}_\perp(q)|^2 \tag{17}$$

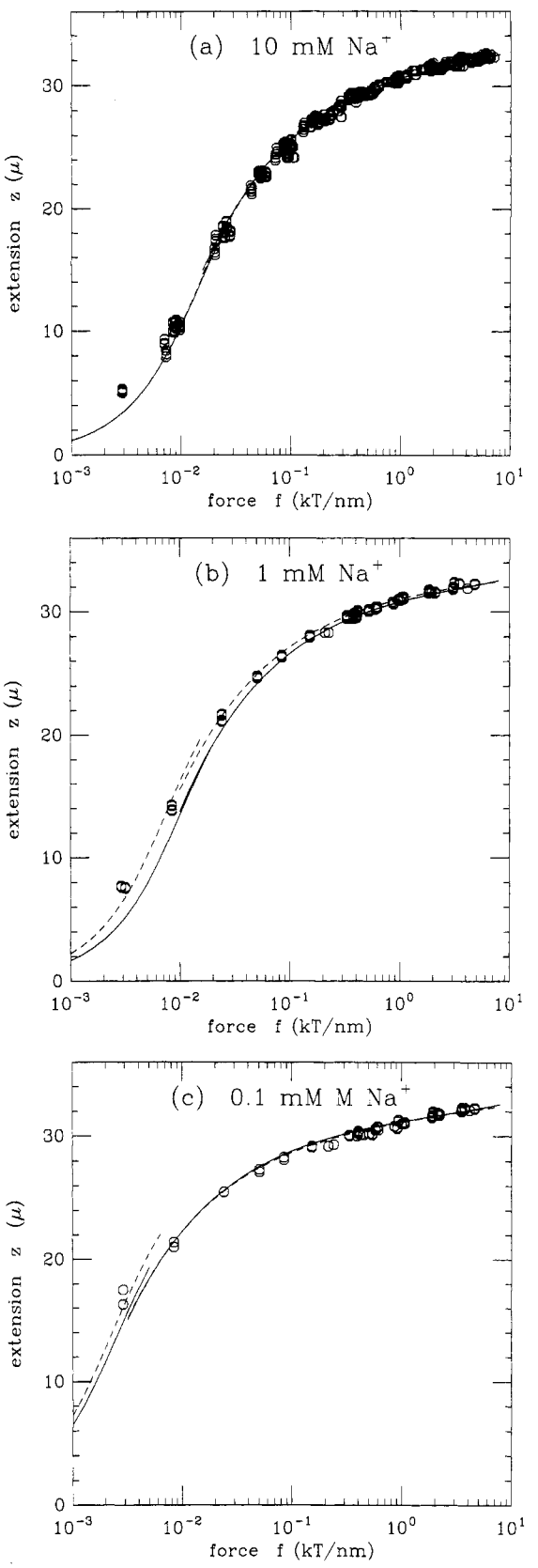

Figure 6. Extension versus force data of Smith et al.[1] compared with the two theoretical fits of section III: (a) 10 mM Na^+; (b) 1 mM Na^+; (c) 0.1 mM Na^+. The high-force fits are done with the Gaussian theory of (19) using $L = 32.8$ μm, $\gamma = 500$ k_BT/nm, and $A_0 = 50.8$ nm (solid lines) and $L = 33.7$ μm, γ infinite, and $A_0 = 15$ nm (dashed lines).

$$K(q) = \frac{2}{(\lambda_D q)^2}\left(\left[1 + \frac{1}{(\lambda_D q)^2}\right]\log[1 + (\lambda_D q)^2] - 1\right)$$

where $l_{OSF} = l_B(\lambda_D\nu)^2/4$. The Debye–Hückel kernel

behaves like $K(q=0) = 1$, $K(q\to\infty) \approx 4(\lambda_D q)^{-2} \log q$, and $K(q=2.029/\lambda_D) = 1/2$. The length l_{OSF} is the only object containing the unknown effective charge ν and will be treated as a phenomenological parameter.

The electrostatic energy (17) should be added to the force and intrinsic elasticity contributions of (1) to yield the effective energy for small $\mathbf{t}_\perp$:

$$\frac{E}{k_B T} = \frac{1}{2}\int \frac{dq}{2\pi}([A_0 + l_{OSF}K(q)]q^2 + f)|\tilde{\mathbf{t}}_\perp(q)|^2 - fL \quad (18)$$

where now A_0 denotes the "elastic persistence length". The coefficient of $q^2|\tilde{\mathbf{t}}_\perp|^2/2$ is the total, "scale-dependent persistence length"[14] $A_0 + l_{OSF}K(q)$, which varies from $A \equiv A_0 + l_{OSF}$ for $q \ll \lambda_D^{-1}$ to A_0 for $q \gg \lambda_D^{-1}$.

Repeating the steps from (3) to (5), we find that for strong forces (i.e., $fA \gg 1$) the extension is

$$\frac{z}{L} = 1 - \int \frac{dq}{2\pi} \frac{1}{[A_0 + l_{OSF}K(q)]q^2 + f} + \frac{f}{\gamma} \quad (19)$$

The final f/γ term accounts for the intrinsic linear stretching elasticity introduced in section II.C, which allows z to exceed L; this term is simply added to the z/L, which follows from (18) since in this Gaussian limit, bending and stretching fluctuations are decoupled.[16]

For forces $f \gg A/\lambda_D^2$ and γ infinite, only the intrinsic persistence length A_0 contributes to (19). Therefore the experimental data at high forces for low salt (where λ_D is large) most directly determine L and A_0. Since we expect ν to be nearly salt independent,[26] and therefore that $l_{OSF} \approx \lambda_D^2$, the disparity between A and A_0 should be largest at low salt. By the same reasoning, γ is best determined from high-salt data, where Coulomb effects are least important.

The small-fluctuation result (19) is accurate for $z/L > 0.5$ (as determined empirically from Figures 6a–c) but the experimental data extend down to $z/L \approx 0.1$, demanding a better calculation. However, the long-range electrostatic interaction (16) makes improvement of our Gaussian approximation problematic. Fortunately, for ≥0.1 mM Na^+, the screening length $\lambda_D \le A_0$, and a lower bound on the crossover force, A_0/λ_D^2, is large enough (i.e., $>1/A$) so that a calculation based on the WLC model starting from low extensions and using A will cross over to its Gaussian limit before screening effects come into play. This suggests there will be a substantial region of overlap between a low-force fit of A using the WLC model and (19).

B. Connecting the Theory to Experimental Results. To limit the number of free parameters, we immediately assume that only l_{OSF} depends on the salt concentration and demand that a single choice of A_0, γ, and L work throughout the concentration range. We use the Debye–Hückel formula $\lambda_D = 0.25/M^{1/2}$ nm, assuming $Na_2HPO_4 \rightarrow 2Na^+ + HPO_4^{-2}$, with the molarities always defined with respect to Na^+. The data can be equally well fit in two extreme limits, each of which ignores one possible additional fitting parameter. In the first scenario, we retain γ but insist that l_{OSF} for 10 mM is 0.01 of its value for 0.1 mM, i.e., that ν in the two limits is the same, and all the salt dependence comes from λ_D (this amounts to placing total confidence in the Debye–Hückel theory). The 1 mM data is then a check on the quality of this assumption. In the second scenario, we suppose that the chain is truly inextensible (we set $\gamma^{-1} = 0$), and we fit l_{OSF} separately for each salt concentration. For all of our fits, our errors represent the limits within which a subjectively "good fit" is observed.

1. Fit Assuming Finite Intrinsic Elasticity. In our first scenario, L and γ are extracted from the 10 mM data from the fit in the inset of Figure 2 (section II.C): $\gamma = 500 \pm 100\ k_BT/\text{nm}^{-1}$ and $L = 32.8\ \mu$m. From the 10 mM WLC fit, we have $A = 53$ nm while from the 0.1 mM data, we find $A = 310 \pm 20$ nm (Figure 5b). The relation $A = A_0 + l_{OSF} = A_0 + l_B(\lambda_D\nu)^2/4$, assumed to hold for 0.1 and 10 mM, gives two equations that we solve to obtain an elastic contribution to the persistence length of $A_0 = 50.8$ nm and an effective charge $\nu = 1.54$ nm^{-1}. This value of ν is close to the value $l_B^{-1} = 1.4$ nm^{-1} predicted by Manning condensation theory. A fit to the 1 mM data in Figure 5a gives $A = 85 \pm 10$ nm, which implies $l_{OSF} = 34 \pm 5$ nm, in quite good agreement with the value of 26 nm inferred from the assumption of $l_{OSF} = l_B(\lambda_D\nu)^2/4$ for $\nu = 1.54$ nm^{-1}.

The solid curves in Figures 4–6 show the Gaussian approximation (19) for $z/L > 0.4$, while for $z/L < 0.5$ we plot the exact WLC result using the large-scale $A = A_0 + l_{OSF}$. These curves are plotted for $A_0 = 50.8$ nm and $\nu = 1.54$ nm^{-1}, with l_{OSF} therefore 2.6, 26, and 260 nm for 10, 1, and 0.1 mM, respectively. Note that in Figures 4 and 5 we have used A_{eff} with $L_{eff} = 33.7\ \mu$m merely as a device to amplify small differences between data and fits at high force levels, which would be invisible in a more conventional z vs f plot; the physical DNA length at low force levels is $L = 32.8\ \mu$m in the present case.

2. Fit Assuming Fixed Contour Length. In the second scenario, the intrinsic stretching elasticity is ignored (γ^{-1} is set to zero), and we fit the remaining parameters as follows. The contour length L (fixed at one value at all force levels) is determined first, by demanding that for high enough forces and low salt, A_{eff} crosses over and has a well-defined plateau (i.e., we want $A_0/f \ll \lambda_D^2$). Too small an L will make A_{eff} increase at high force, while too large an L will eliminate the plateau. Unfortunately, for the 0.1 mM data, theory (figure 5b, dashed line) implies that we have to get to $f > 10k_BT/\text{nm}$ to really see a plateau. Hence our fit of $L = 33.7\ \mu$m was chosen to give a hint of curvature to the data points in Figure 5b. A value of $L = 33.4\ \mu$m would create a noticeable plateau for $f > 1k_BT/\text{nm}$, while A_{eff} defined with $L = 34.0\ \mu$m would have no plateau (as in Figure 5a, dashed line). Given L, we can move the high-force A_{eff} up and down by adjusting A_0. We thus find $A_0 = 15 \pm 5$ nm, with error bars set to where a noticeably inferior fit was obtained for all three molarities.

A value of $A_0 \approx 15$ nm would imply $A \approx 25$ nm for physiological salt concentrations (≈0.1 M and above), contradicting most experiments which imply that A sticks at its 10 mM value as the salt concentration increases (however, some experiments suggest lower values for A; see p 279 in ref 5). One could fit a larger value of $A_0 \approx 35$ nm, which moves the curves up, if one simultaneously increases λ_D to $0.38/M^{1/2}$ nm, which moves the curves to the left. A larger value of A_0 could also be made to fit if L was allowed to depend upon the salt concentration. However, given the assumption that $\gamma^{-1} = 0$, a small value of A_0 is unavoidable.

The last step in this fit entails using the low-force data to adjust A separately for each solution and thus to determine $l_{OSF} = 37$, 85, and 325 nm for 10, 1, and 0.1 mM Na^+, respectively, with errors of order 10% (Figures 4–6, dashed lines). These results correspond to effective charges of $\nu = 5.7$, 2.8, and 1.7 nm^{-1}, respectively, tending toward the Manning value of one uncondensed ion per Bjerrum length, $1/l_B = 1.4$ nm^{-1}

at low ionic strength, but there is more variation in ν with molarity than one would expect. As we have seen above, if a larger value of A_0 could be justified, l_{OSF} would be reduced in percentage terms much more for high rather than low salt concentrations, thus leveling the ν values (e.g., with $A_0 = 35$ nm, $\nu = 3.9$, 2.4, and 1.7 nm^{-1}, respectively). Systematic errors in Debye–Hückel theory are expected in more concentrated solutions, which might explain the large ν values inferred by this fit.

3. Comparison of the Two Fits. Figures 4 and 5 show that the two scenarios (solid lines: WLC + electrostatics + extensibility; dashed lines: inextensible WLC + electrostatics) fit the data about equally well. However, the value of $\nu = 1.54$ nm^{-1} independent of ionic strength as suggested by Manning condensation, the use of an a priori reasonable value of $\gamma = 500\ k_BT$/nm, and the value of the intrinsic elastic persistence length $A_0 = 50.8$ nm make the first scenario physically more appealing. It is important that the scatter of experiment around theory visible in Figures 4, 5a, and 5b for A_{eff} all but disappears in the conventional extension vs force plots (Figures 6a–c). We note that the value of A determined from the two fitting procedures was the same for 10 and 1.0 mM salt but differed marginally (310 nm with γ finite vs 340 nm with γ infinite) for the lowest salt concentration.

C. Self-Avoidance. The lowest extensions studied experimentally ($z/L < 0.25$ in Figures 6a and 6b) are systematically larger than the theoretical z/L. This is emphasized in the low-force apparent persistence lengths (extreme left side of Figures 4 and 5a). This disparity might be attributed to systematic error in measuring the low-force extensions of a few microns (the 3 μm size of the bead and the slow fluctuations of the DNA + bead both serve to make the low-force regime difficult to study). On the other hand, could this effect be attributed to swelling of the unstretched polymer by excluded-volume interactions? The fundamentally nonlocal self-avoidance effect is not included in the above model, which was designed only to account for the effects of electrostatic interactions *locally* along a DNA.

Here we quickly review how simple Flory theory provides a quantitative estimate of when self-avoidance between distant points along the chain is important.[8] From the excluded phase space volume $4A^2D$ of a rod of diameter D and length $2A$ (the WLC Kuhn length), we conclude that self-avoidance effects become important for an unstretched random coil (using end-to-end radius $R_0 = (2AL)^{1/2}$) when $LD^2/(8A^3) > 1$.[8] For $L = 33$ μm, $D = 2\lambda_D = 0.5M^{-1/2}$ nm, and low-force persistence lengths $A = 50$, 100, and 350 nm for Na^+ molarities $M = 10^{-2}$, 10^{-3}, and 10^{-4}, respectively, we find $LD^2/(8A^3) \approx 0.83$, 1.03, and 0.24.

Therefore, at zero extension we can expect some swelling via self-avoidance as is suggested by the low-force data in Figures 4, 5a, and 6a–c. However, for even small extensions to a few times the random-coil size $R_0 = (2AL)^{1/2}$, self-avoidance effects will be suppressed by the consequent reduction in DNA concentration. A simple Flory estimate indicates that for extension to a length z, the above criterion for swelling due to self-avoidance is suppressed by a factor R_0/z. Thus for extensions $z/L \geq 0.25$, $R_0/z \leq 0.2$, and self-avoidance effects are unimportant. Of course, once $z/L > 0.4$, the tangent vector is predominantly along $\hat{\mathbf{z}}$, and self-avoidance effects disappear completely.

D. Intrinsic Bending. Biological DNAs (e.g., the λ-phage genome used in all experiments discussed in this paper) have heterogeneous base pair sequences. Consequently, biological DNAs have sequence-dependent permanent bends along their length.[27] Although it is plausible that such intrinsic bending of DNA should be locally and uniquely determined by sequence,[28] it is not yet known how to reliably do so. A possible simplification follows from the observation that at large scales ($\geq$1 kb), protein-coding sequences appear to have Gaussian statistics, corresponding essentially to random walks in sequence:[29] thus large-scale intrinsic bends might be considered to be random. What can we say about the effects of large-scale random intrinsic bends on the WLC elasticity discussed in the previous section?

Using electrophoretic mobility data, Trifonov and coworkers have proposed a simple model which associates a small ($\approx$4°) sequence-dependent bend with nearest-neighbor base pairs.[30] The shape obtained from application of this model to the 48.5 kb λ-phage sequence corresponds to a *thermal* WLC configuration with a persistence length of $A_{struct} \approx 400$ nm.[31]

This suggests a simple statistical model for large-scale intrinsic DNA shape:[28] a particular configuration taken from the WLC ensemble with persistence length A_{struct}. If on top of these intrinsic random bends there are thermal bends with WLC bending persistence length A, Trifonov et al. noted[28] that the random-coil fluctuations of this model correspond to those of a WLC with an apparent persistence length $A_{app} = A/[1 + A/A_{struct}]$. The intrinsic bends are an additional source of disorder and shorten the apparent persistence length.

For $A_{struct} = 400$ nm and $A = 50$ nm, the intrinsic bends are gradual compared to those excited thermally. Therefore, we might expect a slight reduction of the unstretched coil size and a consequent slight decrease in the extension obtained for a given force in the linear Edwards regime to result from intrinsic bends. However, for $L \approx 30$ μm, the small effect suggested by $A_{struct} \approx 400$ nm will be more than offset by the swelling from self-avoidance (the very different z/L dependence of self-avoidance and disorder effects must be kept in mind; see section III.C).

For strong stretching, one might imagine the length stored in intrinsic bends could alter the large-f force law.[1] However, a simple calculation indicates that this is not the case. If we suppose the DNA is a series of half-circles of radius A_{struct} and if we suppose that the energy of deformation is $A(\kappa - A_{struct}^{-1})^2 ds/2$ (via generalization of (1) to spontaneous curvature A_{struct}^{-1}), then the extension z for large force varies as $1 - z/L \approx (A/A_{struct})^4 (fA)^{-2}$. The prefactor is small because we expect $A < A_{struct}$; more importantly, the $1/f^2$ power-law decay is much faster than the $1/f^{1/2}$ contribution of WLC thermal fluctuations. These primitive estimates suggest that for DNA stretching experiments, intrinsic bends characterized by $A_{struct} \gg A$ can be ignored, especially in the large-force regime. On the other hand, sharp or strongly correlated bends can be expected to produce interesting effects.

IV. Stretching Tethered DNAs with Electric and Flow Fields

The experiments discussed in the previous section[1] were well designed to extract the elastic properties of DNA because (a) one end of the molecule is anchored, and a known force is exerted on the other end, giving a uniform tension along the chain that can be exactly described thermodynamically; (b) no fluorescent dye (e.g., ethidium bromide, DAPI, or YOYO) which could modify the elastic properties of the DNA is needed since only the bead on the free end needs to be imaged; and

(c) high forces first lead to the Gaussian regime discussed in section II.A, which is very easy to treat, and eventually to a regime where the base pairs unstack (section V). Here the uniform tension allows one to probe coexistence of the B-form fixed-length phase with the unstacked phase.[17]

A second class of experiments use electric fields and hydrodynamic flow to stretch tethered DNAs *without* beads attached at their free ends. They are characterized by highly inhomogeneous stretching: near the free end, there is very little stretching, while near the tethered end, the tension is the sum of forces applied to the remainder of the chain and can be very large (Figure 1b). Such experiments are closer to the kind of stretching that might occur naturally during gel electrophoresis[2] or in shear flows.[3] In this section, we show how the thermodynamic expressions for the WLC derived in section II can be employed to treat these inhomogeneous situations.

A. Uniform Electric Field. We begin with the problem of a fixed electric field $\mathbf{E} = E\hat{\mathbf{z}}$ acting on a DNA molecule which has one end fixed in space and the other free. In practice, a uniform field can be maintained in an ionic solution only by driving a net current, and this can be stably done in standard electrophoresis setups.[2,32]

In the interesting limit of the screening length much less than the largest dimension of the body, the forces produced by electrophoresis are of hydrodynamic origin and result from a slight imbalance between the electric field acting directly on the surface charge and the shear stress of the flow in the screening layer.[32] How this applies to a deformable object like a polymer is still an open question, and in this section, we examine the consequences of a purely phenomenological description where the tension at a point s on a chain anchored at $s = 0$ is $\nu E\hat{\mathbf{z}}\cdot[\mathbf{r}(L) - \mathbf{r}(s)]$. The phenomenological constant ν has the dimensions of charge/length but, when computed for a solid rod, involves also the diameter and screening length. For a random coil configuration the tension at $s = 0$ scales as the radius of gyration in accordance with arguments of Ajdari et al.[32] We can express our ansatz for the tension f at a general point s in a form suitable for further analysis as

$$f(s) = \nu E\hat{\mathbf{z}}\cdot\int_s^L \mathrm{d}s\ \hat{\mathbf{t}}(s) \tag{20}$$

The tension at s decreases as one approaches the free end and is proportional to $\hat{\mathbf{z}}\cdot[\mathbf{r}(L) - \mathbf{r}(s)]$, the distance along z to the end of the chain.

We can close (20) in the limit when the time-averaged tension $f(s)$ (and therefore $\hat{\mathbf{z}}\cdot\hat{\mathbf{t}}(s)$) varies with s on a scale much larger than the correlation length for thermal fluctuations, $\xi(s) \approx A/[1 + fA]^{1/2}$ (this formula for ξ interpolates between the large- and low-tension limits discussed in section I.A; for $f \geq 1/A$, our closure criterion $\xi|\mathrm{d}\log f/\mathrm{d}s| \ll 1$ reduces to $\mathrm{d}\xi/\mathrm{d}s \ll 1$). Under these conditions, the averaged left-handed side of (20) may be related as a function of s to the local tangent vector orientation $y(s) \equiv \langle\hat{\mathbf{z}}\cdot\hat{\mathbf{t}}\rangle$ by using just the thermodynamic relation for the force in terms of the averaged tangent vector, i.e., z/L of section II.

In view of other approximations to follow, it suffices to use the approximate WLC force–extension relation (7). Thus (20) becomes explicitly

$$y + \frac{1}{4(1-y)^2} - \frac{1}{4} = \epsilon\int_s^L \frac{\mathrm{d}s}{A} y \tag{21}$$

where we have introduced the dimensionless electric field $\epsilon \equiv \nu EA^2/k_BT$ (the force applied to a persistence length of DNA in units of k_BT/A). For $E = 1$ V/cm, $A = 50$ nm, and $\nu < 2e/\mathrm{nm}$, we have $\epsilon < 0.02$. It is this small parameter which will lead to the slow variation of f and y on the scale ξ, thereby justifying our WKB-like approximation.

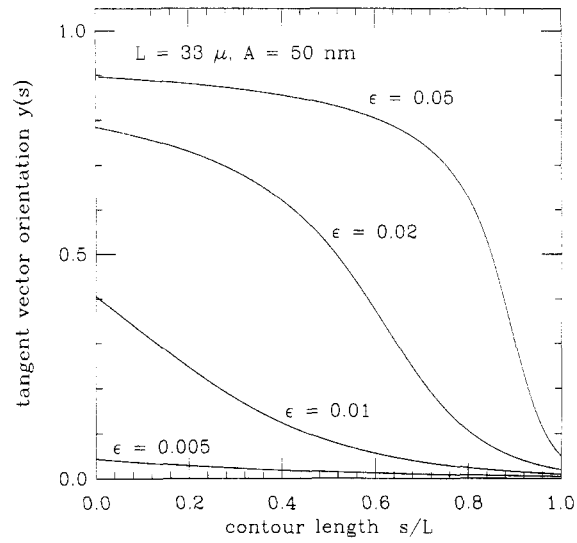

Figure 7. Tangent vector orientation profiles $y(s) = \langle\hat{\mathbf{t}}(s)\cdot\hat{\mathbf{z}}\rangle$ versus contour length s for a tethered $L = 33$ μm DNA with $A = 50$ nm stretched by an electric field. Results for dimensionless electric fields $\epsilon \equiv \nu EA^2/k_BT = 0.005$, 0.01, 0.02, and 0.05 are shown. When $\epsilon L/A \gg 1$, the stretching is appreciable: $y > 0.3$ along more than half the chain.

Differentiation of (21) yields a local equation for $y(s)$:

$$\left[1 + \frac{1}{2(1-y)^3}\right]\frac{\mathrm{d}y}{\mathrm{d}s} = -\epsilon\frac{y}{A} \tag{22}$$

This equation indicates that $\xi|\mathrm{d}\log y/\mathrm{d}s| \leq \epsilon$ everywhere, and therefore for $\epsilon \ll 1$ our WKB approximation is justified. Note that we must not use the boundary condition $y(L) = 0$ since integration of (22) from the free end would then lead to $y(s) \equiv 0$. This is caused by the fact that our thermodynamic relation assumed a long chain and neglected boundary conditions.

To derive a reasonable boundary condition for (22), we imagine the last persistence length to be a freely reorienting rod in an external field. This suggests $y(L) \approx \epsilon \ll 1$ as a boundary condition for (22), which may then be solved in closed form:

$$\frac{\epsilon L}{A}(1 - s/L) = \frac{1}{2}\log\frac{y^3/(1-y)}{\epsilon^3/(1-\epsilon)} + \frac{1}{4(1-y)^2} + \frac{1}{2(1-y)} - \frac{1}{4(1-\epsilon)^2} - \frac{1}{2(1-\epsilon)} \tag{23}$$

We see that the right-hand side depends only weakly on the boundary condition $y(s{=}L)$, and thus any imprecision in its definition is immaterial. Figure 7 shows profiles for the tangent orientation, $y(s)$, for a 33 μm DNA, with $A = 50$ nm. The amount of DNA per length along $\hat{\mathbf{z}}$ such as one would measure from fluorescence intensity is just proportional to y^{-1}.

Equation 23 implies that the tethered end of the chain is well aligned with the field if $\epsilon L/A \geq \ln(\epsilon^{-1})$. The characteristic parameter, $\epsilon L/A$ in (23) follows by comparing the typical force necessary to align DNA, k_BT/A with the total electric force on the molecule computed in the fully extended limit, νEL. It is of course the

thermal disorientation of the chain that gives the logarithmic factor.

In the limit $\epsilon L/A \gg 1$, most of the chain is stretched and we can solve (23) by just keeping the term $\sim 1/(1-y)^2$ on the right-hand side. The asymptotic behavior of the total extension $z = \int_0^L ds\, y(s)$ is just

$$z/L \approx 1 - 1/(\epsilon L/A)^{1/2} \tag{24}$$

This equation is just the strong-stretching extension (5) with applied force $f = \nu LE/4$. The square root is a signature of WLC strong-stretching elasticity. Experiments in this limit can test the scaling properties of WLC elasticity (the power law) and measure the effective electrophoretic charge (the prefactor).

B. Hydrodynamic Flow. Perkins et al. have used hydrodynamic flow to extend DNA molecules from 16 to 83 μm in length by attaching a bead to one end and holding it in a steady flow in a laser trap ("optical tweezers").[3] The bead was a negligible perturbation on the flow. Here we will use the WLC model as the basis for a semiquantitative description of the elongation vs flow velocity.

The main challenge is to estimate the drag force on the DNA, since the drag depends on the shape and the molecule is simultaneously deformed. The particularities of low Reynolds number hydrodynamics make it fairly easy to make estimates to within a factor of ≈ 2, but difficult to do better. A simplifying feature is the hydrodynamic screening which makes the tangle of DNA appear to the flow like a solid body. For instance, in the simple case of a random coil, there will be a pressure difference $\approx \eta v/a$ across the equivalent sphere of radius a, where η is the viscosity and v the velocity. If d is the characteristic spacing of the DNA filaments inside, this pressure will only create a flow $\approx d^2 v/a^2 \ll v$ through the DNA tangle, which therefore appears solid. A second simplification is that the drag for small Reynolds number is set by the longest dimension of the body, with a coefficient that varies slowly with the aspect ratio. For instance, the drag on a sphere is $6\pi\eta a v$ and that on a prolate ellipsoid (major axis a, minor axis b) is $4\pi\eta a v/\log(a/b)$ for $a \gg b$.[33]

This minimal dependence of the drag on the body's shape allows a reasonable fit to the extension experiments using a single parameter to describe the stretching of the molecule, the characteristic tangent orientation $y = \langle \hat{\mathbf{t}} \cdot \hat{\mathbf{z}} \rangle$. We ignore the variation of tension with s along the molecule and thus the increasing transverse fluctuations as one proceeds toward the free end (a family of ellipses with common b and variable $a(s) \geq b$ would be a reasonable guess for the drag on the DNA between s and L).

We will use once again the simple interpolation formula (7) to relate y and the force. To close the problem we therefore must express the drag force interms of y. For strong stretching, we identify the minor semiaxis with the transverse wandering distance $R_\perp$ (6) and the major semiaxis with $\alpha Ly/2$, where α will be used as an $\mathcal{O}(1)$ fitting parameter. For weak stretching, neither a nor b should be allowed to drop below the unperturbed random-coil hydrodynamic radius $\approx 0.375(AL)^{1/2}$ [34] (the numerical prefactor makes eq 25 go to a standard result for the drag on a random coil in the $y \to 0$ limit).

Both these limits are included in the following interpolation formula for the drag force:

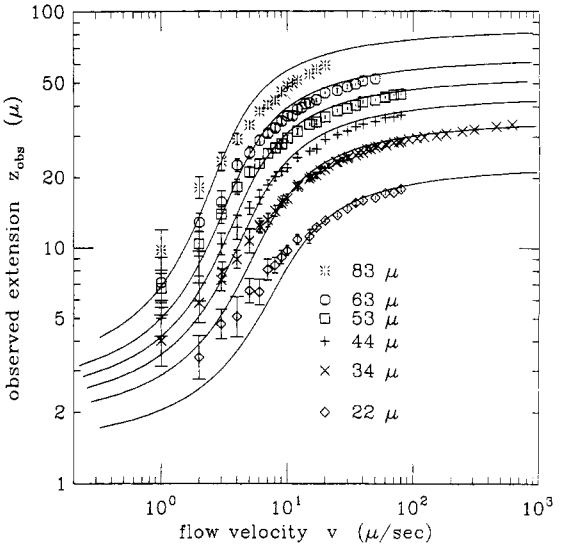

Figure 8. Stretching of a tethered DNA with hydrodynamic flow. Theoretical (curves, eq 25 with $\eta = 0.01$ P, $A = 50$ nm, and $\alpha = 0.25$) and experimental (symbols, data of Perkins et al.[3]) extensions z_{obs} as a function of flow velocity v are shown for $L = 22$, 35, 44, 53, 63, and 83 μm.

$$f_D = \frac{4\pi\eta v\alpha(Ly/2 + R_H)}{\log(1 + Ly/[2R_H]) + 2\alpha/3} \tag{25}$$

where $R_H = 0.375(AL)^{1/2}(1-y^2)$ describes the transverse (hydrodynamic) radius in both the weak- and strong-force limits. This drag force can then be equated to the WLC stretching force (7) to determine the stretching y.

These very simple ideas illustrate that high and low flows generate extensions that scale very differently with v and L. For weak stretching, the Stokes drag $\eta v(AL)^{1/2}$ balances the Edwards elasticity k_BTy/A (numerical factors are ignored). For strong stretching, the drag is roughtly $\eta vL/\log[L/b]$ (for the moment, we suppose that the ellipsoid minor semiaxis b is fixed) and balances the WLC elastic force $k_BT/[A(1-y)^2]$. In these two limits, we thus obtain the scaling behaviors

$$y \approx \frac{A^{3/2}\eta v L^{1/2}}{k_BT} \qquad y \ll 1 \tag{26}$$

$$1 - y \approx \left(\frac{k_BT}{A\eta vL/\log[L/b]}\right)^{1/2} \qquad y \to 1$$

Thus y is not a function of one combination of $\eta v/k_BT$, L, and A. Before we compare with details of the results of Perkins et al., we note that their measured extensions do not go to zero at zero flow (see Figure 2b of ref. 3). Their measurements of the longitudinal size of the "cloud" of fluorescently labeled DNA at low extensions are measures of the random-coil size. We may correct for this by adding a random-coil correction to our theoretical extension, contrived to disappear as $y \to 1$ in the same way as the transverse fluctuations:

$$z_{obs} = Ly + (2AL)^{1/2}(1-y^2) \tag{27}$$

In Figure 8 we compare experimental data of Perkins et al.[3] for extension as a function of flow, with the combination of (25), (27), and (7). The DNA lengths were $L = 22$, 35, 44, 53, 63, and 83 μm, and we used $\eta = 0.01$ P, $A = 50$ nm, and set $\alpha = 0.25$ to optimize the fit. For large extensions, the scaling behavior $1 - y \propto v^{-1/2}$ is plainly apparent (see inset of Figure 1c of

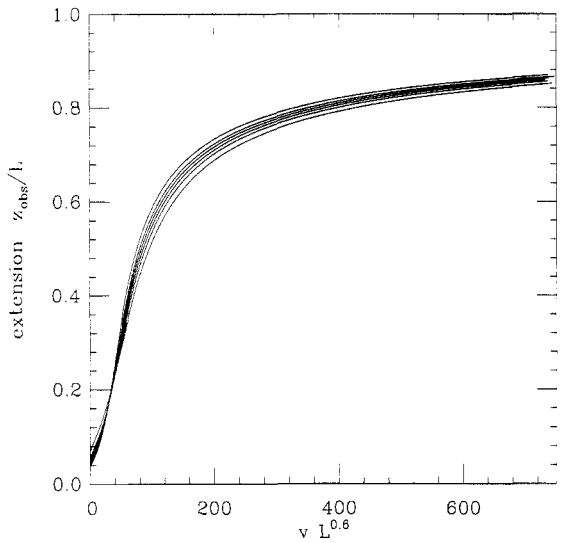

Figure 9. Theoretical curves for stretching of a chain by a flow from Figure 8 replotted as z_{obs}/L versus $vL^{0.6}$; a superimposition of curves very similar to that reported by Perkins et al.[3] is obtained.

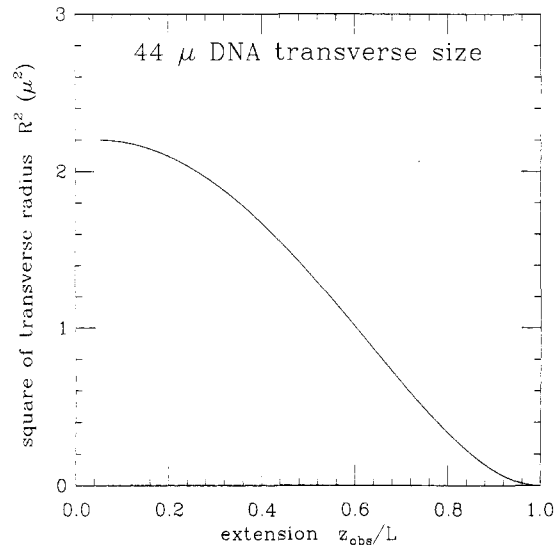

Figure 10. Square of transverse radius R^2 for a DNA stretched by flow versus extension z_{obs}/L for $L = 44$ μm. Between 0.2 and 0.8, there is a roughly linear dropoff of R^2 even though the asymptotic behavior for $z_{obs}/L \to 1$ is quadratic.

Perkins et al.), and once again the square root is a signature of WLC elasticity.

Over much of the experimental data range, the two limiting combinations of velocity and length which enter (26) ($vL^{1/2}$ for weak stretching and $vL/\log L$ for strong stretching) can be jointly fit as vL^x for some exponent x. For instance, in Figure 9, we show the theoretical curves from Figure 8 replotted as z_{obs}/L versus $vL^{0.6}$. The result is very similar to the collapse of the data shown in Figure 2b of Perkins et al.[3] where $vL^{0.54}$ was used as the abscissa.

Finally, in Figure 10 we show the square of the total transverse radius $R^2 = AL(1 - y^2)^2$ versus extension z_{obs}/L for $L = 44$ μm. D. Smith reported[35] to us that for extensions between 0.35 and 0.75, there is a roughly linear dropoff of the square of the transverse size. Very near $z_{obs}/L = 1$, R^2 has a quadratic "foot", but over most of the extension range (0.35–0.8) it indeed drops off nearly linearly. Further details of the conformation (e.g., the inevitable "trumpet" shape[10]) depend on the precise way that the tension builds up with distance from the free end, which in turn depends on the flow past the elongated coil.

V. Unstacking DNA

The stacking and base-pairing interactions provide a few k_BT per base pair of cohesive energy to the DNA double helix.[36] This energy density corresponds to a force on the order of $10k_BT$/nm $\approx$ 40 pN. This is a very large force compared to the $\approx k_BT/A \approx 0.08$ pN characteristic scale for WLC entropic elasticity. Therefore a model with fixed contour length, and an unperturbed ds-DNA helix should be very reasonable for the experiments discussed in sections II, III, and IV. On the other hand, if forces on the order of 40 pN are applied to a DNA, we should expect the double-helix structure to deform: we expect to observe DNA's intrinsic stretching elasticity. In sections II and III, we have already noted that data of Smith et al.[1] at forces $\approx$ 10 pN suggest the onset of linear stretching elasticity.[16] In this section, we discuss evidence for, and a simple model for, extremely nonlinear elastic response of DNA that results from an abrupt change in structure of the double helix under high tensions.

Recently[17] the measurements in Figure 2 have been extended to higher static forces. A transition was observed for a force $\approx$ 50 pN, where the B-DNA unsnaps and transforms to a state a factor $\alpha \approx 1.85$ times longer. Clear evidence was also provided for coexistence between the B-form and the unfolded phases since the force remained nearly at its critical value, while the extension of the molecule, z, varied in the range $1 < z/L < \alpha$, where L is the B-form contour length. This transition was plausibly attributed to unstacking and unwinding the base pairs. The stacking forces are short-ranged and could be expected to yield abruptly. The simple-minded model of straightened sugar–phosphate backbones (in B-DNA they are helically wrapped around a 2 nm diameter cylinder with a 3.5 nm helix repeat) gives an extension $\alpha \approx 2$.

In this section, we merely set up the thermodynamics to describe the two-phase coexistence, show how parameters measured for thermal denaturation constrain the unbinding force, and speculate about other features of the two-phase region. In addition to α, we will characterize the unstacked phase by a second parameter τ defining its free energy per unit length (of B-DNA). We neglect for the moment the very interesting but unexplored thermal fluctuations of the unstacked state, which due to its much reduced persistence length we expect to be more pronounced than for B-DNA.

The B-DNA regions are described by an averaged stretching $y \equiv \langle \hat{\mathbf{t}} \cdot \hat{\mathbf{z}} \rangle$, and a free energy per length (in the fixed-extension ensemble[22]) $\mathcal{F}(y) = y^2/(2A) + y^2/[4A(1 - y)]$ (note that $\mathcal{F}$ is the work done stretching the chain and may be derived from the interpolation formula (7) using the boundary condition $\mathcal{F}(y=0) = 0$, and is properly convex and exact in the limit $y \to 1$, in which we work). In analogy with liquid–gas coexistence at fixed volume, we write the total free energy per length:

$$\frac{F}{L} = (1 - \phi)\mathcal{F}(y) + \tau\phi \tag{28}$$

where ϕ is the fraction of unstacked DNA. We add the constraint of fixed extension:

$$\frac{z}{L} = (1 - \phi)y + \alpha\phi \tag{29}$$

which simply states that total length z is the sum of lengths of B-DNA and unstacked portions of the polymer.

For small extension, the tension in the DNA is low, $\phi = 0$, $z/L = y$. However, at some critical extension z^*, it becomes favorable to create unstacked DNA. This occurs when the free energy density of the $\phi = 0$ state, $\mathcal{F}(z/L)$, becomes equal to the line tension of the unstacked phase minus the change in free energy due to the elongation of the B-DNA under the imposed force:

$$\mathcal{F}(z^*/L) = \tau - \mathcal{F}'(z^*/L)(\alpha - z^*/L) \qquad (30)$$

Less physically, eq 30 is obtained by minimization of the free energy (28) subject to the constraint (29). The force at which the unstacking begins to occur is just $f^* = \mathcal{F}'(z^*/L) = [\tau - \mathcal{F}(z^*/L)]/[\alpha - z^*/L]$.

For extensions beyond $z > z^*$, the tension stays constant at f^*, and the unstacking fraction grows linearly with z:

$$\phi = \frac{z - z^*}{\alpha L - z^*} \qquad (31)$$

The constant tension in the chain is precisely what is required to do the work necessary to convert the WLC to unstacked DNA. Finally, when $z = \alpha L$, the DNA is fully unstacked ($\phi = 1$) and in this model, no further extension is possible. For real DNA, there will be additional length stored in the thermal fluctuations of the unstacked region and the resultant elasticity can be measured until either the chemical bonds in the backbone or the attachments to the bead break (chemical bonds will break at very roughly 5 eV/nm $\approx$ 600 pN; streptavidin-biotin links commonly used in stretching experiments break at about 150 pN[38]).

A value of τ can be inferred from the measured critical force $f^* = 50$ pN and the extension $\alpha = 1.85$, which when substituted into (30) imply $\tau = 10.9 k_B T$/nm and $z^*/L = 0.980$ (note τ is not simply $f^* = 50$ pN $= 12.2 k_B T$/nm). Since the sequence of the λ-phage genome used in the experiment is known, we can compute $\Delta G/L = 2.03$ kcal/mol $= 9.82 k_B T$/nm for the helix–coil transition using the measured nearest-neighbor stacking energies from ref 36 and our value of 0.347 nm/bp. To relate the free energy of two single-stranded random coils to τ (the reference free energy is unstretched B-DNA in all cases), we add to $\Delta G/L$ the free energy difference of unstacked ribbon relative to free coils and the elastic energy to extend the ribbon. Both contributions are positive: the former because DNA melts into free coils and not a ribbon. Although $\tau > \Delta G/L$, the difference is not large, so the unstacked state may effectively be two single strands.

In the absence of any variation in τ along the DNA, general statistical mechanical arguments imply that the two coexisting phases will mix.[39] The size l of the droplets of pure phase can be estimated by balancing the entropy of the "domain wall" separating the two phases against the energetic cost of creating one. The entropy is $\approx k_B T \ln(l/\xi)$, where $\xi \approx 1$ nm is the (zero-dimensional) domain wall thickness, while the energy is $\approx \tau\xi \approx 10 k_B T$, implying $l \approx \xi e^{\tau\xi/k_B T} \approx 10$ μm. These figures are largely guesses. To the extent, however, that $l > A$, the entropy associated with the mixing of domains of the two phases is much smaller than the free energy of the WLC chain elasticity, which would diverge as $x \to L$ were the unstacking transition not to intervene. In practice, long-range correlations in the G/C vs A/T composition will make it favorable to put the unstacked phase where the B-DNA interstrand binding is smallest. This heterogeneity is probably responsible for the $\approx$1 pN force needed to interconvert one phase into the other, rounding of the force plateau, and hysteresis observed during interconversion of the two phases. The unstacking transition, analyzed here, is also implicated in the experiments which stretch DNA with a receding miniscus, judging by the forces and overextensions involved.[40] Further support for the phase coexistence model outlined above is provided by Chatenay et al.[41] who very recently measured the force vs distance for overextension of B-DNA and observed a prominent region of constant force over a range of extensions from roughly 1 to 1.8 times the B-DNA length.

Other situations arise where two states of a polymer with different extensions are brought to pseudophase coexistence by stretching. Twisted DNA molecules form plectonemic supercoils, but since the ends of a plectoneme are coincident, any imposed extension will require some fraction of solenoidally supercoiled DNA.[42] There exist proteins which when bound to DNA reduce its length (e.g., histones, TATA-binding proteins), so one might stretch such a complex sufficiently to make it thermodynamically favorable for the proteins to release (based on the $\approx 30 k_B T$ binding energy of the 146 bp of DNA in a nucleosome,[43] we have estimated that a force of 2 pN should be sufficient to liberate histones from chromatin fiber). Finally, we note that similar ideas were applied to the coexistence of collapsed and extended-chain conformations encountered when a polymer "brush" is stretched in a poor solvent.[44]

VI. Conclusions

The experiments analyzed in this article are a compelling illustration that biomaterials—in this case DNA—can furnish a quantitative test of basic aspects of polymer physics that otherwise would remain somewhat academic. Three classes of experiments may be discerned depending on how the force is applied to the DNA (one end is always tethered), (a) static force is applied to the free end only; (b) an electrophoretic force is applied; (c) hydrodynamic drag is applied directly to the DNA. The theory is quantitative in all but the first case since otherwise hydrodynamics is involved and the shape of the DNA has to be determined self-consistently with the force.

The WLC model can be solved exactly for perfect screening in case a, and far more precise experiments are possible by extensions of the techniques pioneered in ref 1. Thus we anticipate a lively exchange between theory and experiment as a number of small effects emerge unambiguously from the data. Prominent among these will be an improved treatment of Coulomb effects. Experiments at $\sim$0.1 M salt will immediately decide whether the small value of $A_0 = 15$ nm determined with no intrinsic elasticity is viable or whether a finite γ is required to fit experiment. For salt concentrations near physiological, one should probably linearize around the correct Poisson–Boltzmann equation in (16), to obtain some corrected kernel in (17). Ultimately, the only approach may be numerical.[15] In low salt, it will no longer be adequate to treat all the crossover in the Gaussian limit and to fit the low-force data with the WLC model containing the long-distance correlation length A. Finally, some unique signature of chemical effects such as intrinsic bends may emerge.

One aspect of DNA physics that has not been addressed here are manifestations of the twist rigidity which leads to an internal twist-angle correlation length of order 75 nm at high salt.[5] Thus DNA supercoils writhe when twisted, like a rubber hose. All force measurements to date have anchored only one of the

DNA strands and thus allowed the molecule to come into twist equilibrium. However, with an imposed twist and reasonable values of force, the DNA is a mixture of plectonemic and solenoidal forms, and changes in extension occur by conversion of one form to the other (the plectoneme has zero extension but is the preferred state in the at zero force). The force–extension curves change dramatically; Gaussian calculations were reported in ref 42.

Acknowledgment. We gratefully acknowledge helpful discussions with A. Ajdari, S. Burley, C. Bustamante, D. Chatenay, P. Cluzel, Y. Cui, P. Furrer, F. MacKintosh, T. Perkins, D. Smith, S. Smith, A. Stasiak, J.-L. Viovy, and J. Widom. We especially thank T. Odijk for his helpful comments regarding the intrinsic stretching elasticity of DNA. J.M. gratefully acknowledges support at RU from the Meyer Foundation; this research was also supported by the National Science Foundation through NSF Grant No. DMR 9012974 and NSF–MRL Award No. DMR-9121654.

Appendix. Numerical Calculation of WLC Free Energy

We may find the spectrum of the right-hand side of (11) by just expanding ψ using spherical harmonics:

$$\psi(\hat{\mathbf{t}}) = \sum_l \psi_l Y_{l0}(\hat{\mathbf{t}}) \tag{32}$$

where ψ_l are the expansion coefficients and where we have anticipated that the ground state must have axial symmetry and thus no $m \neq 0$ components. Insertion of (32) into (12) yields the matrix equation

$$g\psi_l = \sum_{l'}\left[\frac{l(l+1)}{2A}\delta_{l,l'} - \frac{f}{((2l+1)(2l'+1))^{1/2}}(l\delta_{l',l-1} + l'\delta_{l,l'-1})\right] \tag{33}$$

which is easily diagonalized to obtain the minimum eigenvalue g.

An upper cutoff on l renders the problem finite: practically, one can just increase l until convergence to the precision desired over a certain force range is obtained. Figure 3 has a cutoff at $l = 100$, which provides eight-digit accuracy (defined by the change in results relative to calculations with maximum $l = 80$) over the force range ($fA < 100$) relevant to experiments on DNA.

References and Notes

(1) Smith, S. B.; Finzi, L.; Bustamante, C. *Science* **1992**, *258*, 1122.
(2) Schurr, J. M.; Smith, S. B. *Biopolymers* **1990**, *29*, 1161. Smith, S. B.; Bendich, A. J.; *Biopolymers* **1990**, *29*, 1167. Smith, S. B., private communication, 1995.
(3) Perkins, T. T.; Smith, D. E.; Larson, R. G.; Chu, S. *Science* **1995**, *268*, 83.
(4) Bates, A. D.; Maxwell, A. *DNA Topology*; IRL Press: Oxford, 1993; Section 1.
(5) Hagerman, P. J. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 265.
(6) Crothers, D. M.; Drak, J.; Kahn, J. D.; Levene, S. D. *Methods Enzymol.* **1992**, *212*, 3.
(7) Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Cambridge University Press: 1986; Section 8.8.
(8) Odijk, T. *Biopolymers* **1979**, *18*, 3111. Manning, G. S. *Biopolymers* **1981**, *20*, 1751. Post, C. B. *Biopolymers* **1983**, *22*, 1087.
(9) The DNAs in human chromosomes are single molecules ≈ 4 cm long; those in the newt *Triturus cristatus carnifex* measure about 1 m.
(10) Brochard-Wyart, F., preprint, 1995. Brochard-Wyart, F. *Europhys. Lett.* **1993**, *23*, 105. Brochard-Wyart, F.; Hervet, H.; Pincus, P. *Europhys. Lett.* **1994**, *26*, 511.
(11) Fixman, M.; Kovac, J. *J. Chem. Phys.* **1973**, *58*, 1564.
(12) Kovac, J.; Crabb, C. C. *Macromolecules* **1982**, *15*, 537.
(13) Odijk, T. *J. Polym. Sci.* **1977**, *15*, 477. Skolnick, J.; Fixman, M. *Macromolecules* **1977**, *10*, 944.
(14) Barrat, J. L.; Joanny, J.-F. *Europhys. Lett.* **1993**, *24*, 333.
(15) Vologodskii, A. *Macromolecules* **1994**, *27*, 5623.
(16) Odijk, T. *Macromolecules*, in press; private communication.
(17) Smith, S. B.; Cui, Y.; Hausrath, A. C.; Bustamante, C. *Biophys. J.* **1995**, *68*, A250.
(18) Fuller, F. B. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 815.
(19) Tanaka, F.; Takahashi, H. *J. Chem. Phys.* **1985**, *83*, 6017.
(20) Yamakawa, H. *Pure Appl. Chem.* **1976**, *46*, 135.
(21) Bustamante, C.; Smith, S.; Marko, J. F.; Siggia, E. D. *Science* **1994**, *265*, 1599.
(22) The number g is the fluctuating-z free energy per length (in k_BT units) and is analogous to the thermodynamic potential in the theory of gases. The work done stretching the WLC is the free energy in the fixed-z ensemble, $gL + fz$; it is analogous to the Helmholtz free energy.
(23) Lehninger, A. L.; Nelson, D. L.; Cox, M. M. *Principles of Biochemistry*; Worth: New York, 1993; Chapter 4.
(24) Israelachvili, J. N. *Intermolecular and Surface Forces*; Academic: New York, 1985; Section 12.14.
(25) Schellman, J. A.; Stigter, D. *Biopolymers* **1977**, *16*, 1415. Stigter, D. *Biopolymers* **1977**, *16*, 1435.
(26) Oosawa, F. *Polyelectrolytes*; Marcel Dekker: New York, 1971; Section 2.
(27) Crothers, D. M.; Haran, T. E.; Nadeau, J. G. *J. Biol. Chem.* **1990**, *265*, 7093.
(28) Trifonov, E. N.; Tan, R. K.-Z.; Harvey, S. C. *Structure and Expression*: Eds.; *DNA Bending and Curvature*; Olson, W. K., Sarma, R. H., Sundaralingam, M., Adenine: New York, 1987; p 243.
(29) Arneodo, A.; Bacry, E.; Graves, P. V.; Muzy, J. F. *Phys. Rev. Lett.* **1995**, *74*, 3293.
(30) Bolshoy, A.; McNamara, P.; Harrington, R. E.; Trifonov, E. N. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2312.
(31) P. Furrer (unpublished, 1995) computed the tangent vector correlations of the λ-phage sequence for this DNA bending model and noted that the decay of correlations corresponds to that expected in the WLC ($\langle\hat{\mathbf{t}}(s)\cdot\hat{\mathbf{t}}(0)\rangle \sim \exp[-s/A]$) for $A \approx$ 400 nm.
(32) Ajdari et al. have pointed out that there will be a hydrodynamic interaction of distant points along the molecule: the force on any part of the molecule will depend on the entire chain conformation: Long, D.; Viovy, J.-L.; Ajdari, A., preprint, 1995. Our simple locally forced model does not explicitly include this effect.
(33) Berg, H. C. *Random Walks in Biology*; Princeton University Press: Princeton, NJ, 1993; Chapter 6.
(34) The prefactor of 0.375 is that obtained from a microscopic calculation for the Zimm diffusion constant, with the end-to-end distance $bN^{1/2}$ replaced with the WLC value of $(2AL)^{1/2}$. See: Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Cambridge University Press: Cambridge, 1986; Section 4.2.
(35) Smith, D. E., private communication, 1994.
(36) Breslauer, K. J.; Frank, R.; Blocker, H.; Marky, L. A. *Proc. Natl. Sci. U.S.A.* **1986**, *83*, 3746.
(37) An analogous state of affairs exists for (liquid) lipid bilayer membranes, which have been shown to have a rather weak entropic elasticity; area that is stored in thermal undulations of a membrane can be removed by putting it under tension. Under strong tension, however, a membrane can be microscopically expanded: the flexible hydrocarbon chains on the lipids can gradually rearrange, giving a membrane some intrinsic stretching elasticity; see: Evans, E.; Rawicz, W. *Phys. Rev. Lett.* **1990**, *64*, 2094.
(38) Florin, E.-L.; Moy, V. T.; Gaub, H. E. *Science* **1994**, *264*, 415.
(39) Landau, L.; Lifshitz, I. M. *Statistical Mechanics*; Pergamon Press: New York, 1968; Section 152.
(40) Bensimon, D.; Simon, A. J.; Croquette, V.; Bensimon, A. *Phys. Rev. Lett.* **1995**, *74*, 4754.
(41) Cluzel, P.; Lebrun, A.; Heller, C.; Lavery, R.; Viovy, J.-L.; Chatenay, D.; Caron, F., preprint, 1995.
(42) Marko, J. F.; Siggia, E. D. *Science* **1994**, *265*, 506; *Phys. Rev. E* **1995**, *52*, 2912.
(43) Widom, J., private communication, 1995.
(44) Halperin, A. *Macromolecules* **1991**, *24*, 5393.

MA9506086