

## Sequence-Disorder Effects on DNA Entropic Elasticity

Philip Nelson

*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104*  
(Received 26 November 1997)

DNA stretching experiments are usually interpreted using the wormlike chain model; the parameter  $A$  appearing in the model is then interpreted as giving the elastic bend stiffness of the double helix. Actually, however, the value of  $A$  obtained by this method is a combination of bend stiffness and *intrinsic bend* effects reflecting sequence information, just as at zero stretching force. This observation resolves the discrepancy between the value of  $A$  measured in these experiments and the larger “dynamic persistence length” measured by other means. On the other hand, the *twist* stiffness deduced from torsionally constrained stretching experiments suffers no such correction. The calculation is very simple and analytic; it explains the success of the naive wormlike chain model over the entire force range of DNA stretching experiments. [S0031-9007(98)06015-3]

PACS numbers: 87.15.-v, 87.10.+e, 87.15.By

The DNA in living cells is often described as a passive database of pure information, the genome. In fact, however, the DNA molecule itself actively collaborates in its own packaging, transcription, regulation, and repair [1]. Unraveling the underlying mechanisms of these crucial processes requires an understanding of the basic mechanical properties of the DNA duplex. For example, the fundamental unit of DNA packaging, the nucleosome, is delicately balanced between elastic stresses and binding energies [2]; an accurate account of the former is clearly important for analyzing the stability of the whole complex. Since nucleosomal DNA is under torsional as well as bending stress [3], an accurate model incorporating both twist and bend is needed.

Recently a new class of experiments has permitted precise physical control over *single molecules* of DNA [4]. For example, a single molecule of known contour length  $L$  can be subjected to known stretching force  $f$  at its ends and the resulting extension (end-to-end length)  $Z$  measured. Simple arguments from polymer physics then predict that  $Z < L$  since thermal fluctuations keep a flexible rod from being perfectly straight;  $Z$  approaches  $L$  at large  $f$ . Remarkably, Bustamante *et al.* [5] found that a very simple model, the “wormlike chain,” fits the force-extension data over four orders of magnitude in  $f$ . The model attributes to DNA just one parameter, the elastic bend stiffness  $\kappa_{\text{eff}}$ , usually expressed in thermal units as  $A_{\text{eff}} = \kappa_{\text{eff}}/k_B T$ . Subsequent experiments have refined its value to [6]  $\kappa_{\text{eff}} = (40 \text{ nm})k_B T$  [7]. In an important extension of the technique, Strick *et al.* devised a *torsionally constrained* stretching experiment [8–10]. Analyses of the corresponding directed walk problem led to values of the twist stiffness  $\tilde{\kappa}_{\text{eff}}$  between  $(75 \text{ nm})k_B T$  and  $(110 \text{ nm})k_B T$  [11–14]; in each case,  $\tilde{\kappa}_{\text{eff}}/\kappa_{\text{eff}}$  was found to exceed unity. Large values for this ratio are not impossible, but they are disturbing in the light of classical beam theory: In continuum elasticity a large value of  $\tilde{\kappa}/\kappa$  implies a material with *negative Poisson ratio*, a situation not found in ordinary materials [15].

The purpose of this Letter is to show that the value  $\kappa_{\text{eff}}$  measured by stretching experiments does *not* directly reflect the bend stiffness of the DNA helix, but rather a certain combination of stiffness and *disorder* induced by the sequence of natural DNA. Since these two effects will enter in different combinations in other circumstances, for example, the nucleosome binding energy, it is important to disentangle them. In fact,  $\kappa_{\text{eff}}$  *underestimates* the true elastic bend stiffness  $\kappa$ , while  $\tilde{\kappa}_{\text{eff}}$  accurately reflects the true  $\tilde{\kappa}$ , as announced in [13]. Thus the large observed value of  $\tilde{\kappa}_{\text{eff}}/\kappa_{\text{eff}}$  is perhaps not as mysterious as it at first seems.

Recently Bensimon *et al.* [16] have independently studied these and other issues. Using a different model from ours, they found analytical formulas for low-force stretching and numerical results for all  $f$ , at both strong and weak disorder. Below we will restrict to the case of weak disorder, the case relevant for DNA. In this limit the calculation becomes very simple. The result obtained here for  $\kappa_{\text{eff}}$  differs from [16], as described below. We will also retain the torsional degree of freedom needed to study the twist stiffness.

The result of this Letter is perhaps not surprising in the light of extensive earlier work on DNA coils at *zero* applied tension. A uniform rigid stack of monomers must form some sort of helix, and in particular such a helix will have a straight axis in its undeformed state. DNA, however, is a stack of four different types of units. The sequence of natural DNA has a small component with period equal to the helix repeat [17], but mainly the sequence imparts random natural bends to the rod [18]. Trifonov *et al.* [19] noted that even in the absence of any thermal fluctuations a randomly kinked rod would follow a random walk of some persistence length  $P$ , the “static” (or “structural”) persistence length. They argued that the effective persistence length of such a coil at nonzero temperature would be  $A_{\text{eff}} = \kappa_{\text{eff}}/k_B T$ , where

$$\kappa_{\text{eff}} = \kappa/(1 + \lambda), \quad (1)$$

with  $\kappa$  the true elastic stiffness of the rod and  $\lambda \equiv \kappa/Pk_B T$ . They verified formula (1) with Monte Carlo simulations [20], then computed the numerical value  $P = 216$  nm and hence  $\lambda = 0.3$  starting from sequence information and estimates of the wedge angles. Later, Bednar *et al.* measured  $\lambda$  more directly by comparing random coils of natural DNA to synthetic constructs designed to be straight; they obtained  $\kappa = (78 \text{ nm})k_B T$ ,  $A_{\text{eff}} = (45 \text{ nm})k_B T$ , and hence  $\lambda = 0.4$  [21–23].

One might imagine that under extensional force the kinked rod would simply follow the usual wormlike chain result with  $\kappa$  replaced by  $\kappa_{\text{eff}}$  from Eq. (1). Indeed, we will show that *this is correct for weak disorder* (small  $\lambda$ ). In contrast, Bensimon *et al.* [16] found that at weak disorder  $\kappa_{\text{eff}} = \kappa(1 - \frac{1}{2}\sqrt{\lambda})$ , while Marko and Siggia [24] argued that at high force the disorder is immaterial:  $\kappa_{\text{eff}} = \kappa$ .

When one constrains the total linking number of DNA, as in the experiments of [8–10], the appropriate mathematical problem becomes the “wormlike ribbon chain” (or “torsional directed walk”) [11–14]. We will show that with weak disorder the kinked rod reproduces the results of the homogeneous rod with twist stiffness

$$\tilde{\kappa}_{\text{eff}} = \tilde{\kappa}. \quad (2)$$

*Calculation.*—We wish to evaluate the extension of a randomly kinked, flexible rod under an imposed tension  $f$ , and later an applied torque as well. We seek the leading term in an expansion in weak disorder; the extension to higher orders is straightforward [25].

To describe the rod conformations, let  $\hat{E}_a(s)$  be a “material frame,” i.e., an orthonormal triad describing the orientation of the rod segment at arclength  $s$  from the end, with  $\hat{E}_3$  the tangent to the rod axis. The spatial components  $E_{ia}$  of these three vectors thus form a  $3 \times 3$  orthogonal matrix  $\mathbf{E}(s)$ . Let  $\mathbf{\Omega} \equiv \mathbf{E}^{-1}\dot{\mathbf{E}} \equiv \sum_i \Omega_i \mathbf{T}_i$ , where the dot denotes  $d/ds$ .  $\mathbf{T}_i$  are the three antisymmetric  $3 \times 3$  matrices generating rotations, e.g.,  $[\mathbf{T}_1]_{23} = +1$ . The elastic energy of a conformation is then

$$E_{\text{elastic}} = \frac{1}{2} \int_0^L ds [\kappa(\Omega_1 - \zeta_1)^2 + \kappa(\Omega_2 - \zeta_2)^2 + \tilde{\kappa}(\Omega_3 - \zeta_3)^2]. \quad (3)$$

The functions  $\zeta_i(s)$  appearing in (3) specify the random kinks [26]. We give them an isotropic, Gaussian distribution,

$$\begin{aligned} \llbracket \zeta_i(s) \rrbracket &= 0, \\ \llbracket \zeta_i(s) \zeta_j(s') \rrbracket &= \frac{\lambda}{A} \delta(s - s') \begin{bmatrix} 1 & & \\ & 1 & \\ & & g \end{bmatrix}_{ij}. \end{aligned} \quad (4)$$

Here the double brackets signify an average over an ensemble of many possible sequences [27]. Considering the curve whose curvature is exactly  $\zeta_i(s)$ , one can see that  $P = A/\lambda$  is the structural persistence length mentioned above, by calculating  $\llbracket \hat{E}_3(0) \cdot \hat{E}_3(s) \rrbracket = 1 - \frac{s}{P} +$

$\mathcal{O}(s^2)$ . The constant  $g$  in Eq. (4) will drop out of our answers. Finally we add to the energy a term describing the work done by the external stretching force,

$$-f \int ds E_{33}. \quad (5)$$

Even neglecting thermal undulations altogether, straightening the rod requires some extensional force to overcome the *elastic* energy (3). We must now introduce *entropic* effects as well, and compute the full extension

$$Z/L = \llbracket \langle E_{33}(0) \rangle \rrbracket, \quad (6)$$

where the angle brackets are the usual thermal average.

To carry out the calculation, begin with the Euler angle representation of a rotation matrix, defining three fields  $\theta(s)$ ,  $\phi(s)$ , and  $\psi(s)$  by

$$\mathbf{E} = e^{-\phi \mathbf{T}_3} e^{-\theta \mathbf{T}_2} e^{-\psi \mathbf{T}_3}. \quad (7)$$

To exploit the assumed isotropy of the rod and its disorder, define the complex variable  $\mathcal{W} = (\Omega_1 + i\Omega_2)/\sqrt{2} = e^{-i\psi}(-i\dot{\theta} + \dot{\phi} \sin \theta)/\sqrt{2}$ . Similarly, let  $\mathcal{Z} = (\zeta_1 + i\zeta_2)/\sqrt{2}$ , which then obeys  $\llbracket \mathcal{Z}(s) \mathcal{Z}^*(s') \rrbracket = \frac{\lambda}{A} \delta(s - s')$  and  $\llbracket \mathcal{Z}(s) \mathcal{Z}(s') \rrbracket = 0$ . The energy then becomes

$$E_{\text{elastic}} = \int ds \left[ \kappa(|\mathcal{W}|^2 - \mathcal{W} \mathcal{Z}^* - \mathcal{W}^* \mathcal{Z}) + \frac{1}{2} \tilde{\kappa}(\dot{\psi} + \dot{\phi} \cos \theta + \zeta_3)^2 - f \cos \theta \right]. \quad (8)$$

We have dropped the divergent constant  $\int |\mathcal{Z}|^2$  from (8) because constants in the energy do not affect thermal averages. It is now clear that the disorder field  $\zeta_3(s)$  may be eliminated from the  $\tilde{\kappa}$  term of (8) by shifting the definition of  $\psi$  [28]. Since  $\psi$  does not enter the first term, while the next two terms already contain the disorder field  $\mathcal{Z}$ , this shift eliminates  $\zeta_3$  to leading order in the strength  $\lambda$ . The physical meaning of this shift is simple. Consider a straight, isotropic rod with a randomly rotating reference stripe painted on its surface. Nothing changes if we pass to a different material frame rotated at  $s$  by an angle  $\int^s ds' \zeta_3(s')$  relative to the old one.

What makes our problem interesting is that the disorder  $\mathcal{Z}$  *cannot* be so trivially eliminated, due to a clash between the  $\kappa$  terms and the  $f$  term. To leading nontrivial order in the disorder strength  $\lambda$  the  $\mathcal{Z}$  terms of (8) contribute

$$1 + \left( \frac{\kappa}{k_B T} \right)^2 \int ds ds' \mathcal{W}(s) \mathcal{Z}^*(s) \mathcal{W}^*(s') \mathcal{Z}(s'), \quad (9)$$

to the Boltzmann weight  $e^{-E_{\text{elastic}}/k_B T}$ . Performing the disorder average over the  $\mathcal{Z}$  fields eliminates one of the integrations over  $s$  [see Eq. (4)], so that the correction factor is the leading term of  $e^{\kappa \lambda \int ds |\mathcal{W}|^2}/k_B T$ . Comparing to (8), we see that to  $\mathcal{O}(\lambda)$  the effect of disorder is simply to replace  $\kappa$  by  $\kappa_{\text{eff}} = \kappa(1 - \lambda)$ , leaving  $\tilde{\kappa}$  unchanged. This proves (1) and (2) since we are working to first order in  $\lambda$ . To go beyond this order we must be careful to treat the disorder as quenched, for example, via the replica trick [29].

Thus, within our approximations, *the only effect of sequence on entropic elasticity is to reduce the apparent bend stiffness*, as claimed in Eqs. (1) and (2).

*Discussion.*—The model investigated above may seem highly reductionist, neglecting as it does all of the specific properties of DNA, e.g., the specific bends at particular base-pair junctions. Indeed, we have used a continuum model, where there are no base pairs at all. But it is precisely the existence of a good continuum limit, despite the very singular form of the assumed disorder (4), which gives the result universality. Like the phenomenon of entropic elasticity itself, random kinks affect the force-extension curve via fluctuations over length scales much longer than a base pair.

The analysis given here explains the qualitative success of models without disorder in fitting DNA stretching experiments: The force-extension curve is predicted to be of the same form as the naive wormlike chain, with only a renormalization of the bend stiffness. It also predicts that single-molecule stretching experiments on long, intrinsically straight DNA would show the same increase in effective persistence length as is seen at zero force, for example, in [21,22]. More importantly, it implies that the elastic stiffness relevant for deformation of a given segment of DNA on scales shorter than a micron is considerably greater than the value obtained by fitting the naive wormlike chain model to stretching experiments.

I would like to thank J. M. Schurr for a question which led to this work, and R. D. Kamien and T. C. Lubensky for valuable discussions. This work was supported in part by NSF Grant No. DMR95-07366.

- 
- [1] B. Alberts *et al.*, *Molecular Biology of the Cell* (Garland, New York, 1989).
  - [2] K. J. Polach and J. Widom, *J. Mol. Biol.* **254**, 130 (1995).
  - [3] K. Luger *et al.*, *Nature (London)* **389**, 251 (1997).
  - [4] R. H. Austin *et al.*, *Phys. Today* **50**, 32 (1997).
  - [5] C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith, *Science* **265**, 1599 (1994).
  - [6] The authors of [7] separated the true elastic contribution to  $A$  from the electrostatic contribution by extrapolating to high salt. They found the more usual value  $A_{\text{eff}} \approx 47$  nm in physiological salt concentrations.
  - [7] M. D. Wang *et al.*, *Biophys. J.* **72**, 1335 (1997).
  - [8] T. Strick *et al.*, *Science* **271**, 1835 (1996).
  - [9] T. R. Strick, J.-F. Allemand, D. Bensimon, and V. Croquette, *Biophys. J.* (to be published).
  - [10] J.-F. Allemand and V. Croquette (unpublished).
  - [11] J. F. Marko and A. Vologodskii, *Biophys. J.* **73**, 123 (1997).
  - [12] C. Bouchiat and M. Mézard, *Phys. Rev. Lett.* **80**, 1556 (1998).
  - [13] J. D. Moroz and P. Nelson, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14418 (1997).
  - [14] J. D. Moroz and P. Nelson, *cond-mat/9712004*, 1998.
  - [15] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon Press, Oxford, 1986), 3rd ed.

- [16] D. Bensimon, D. Dohmi, and M. Mézard, *Europhys. Lett.* (to be published).
- [17] E. Trifonov and J. Sussman, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3816 (1980).
- [18] J. Widom, *J. Mol. Biol.* **259**, 579 (1996).
- [19] E. N. Trifonov, R. K.-Z. Tan, and S. C. Harvey, in *DNA Bending and Curvature*, edited by W. K. Olson, M. H. Sarma, and M. Sundaralingam (Adenine Press, Schenectady, 1987), pp. 243–254.
- [20] Later analytical calculations also supported the formula [J. A. Schellman and S. C. Harvey, *Biophys. Chem.* **55**, 95 (1995)], though Bensimon *et al.* [16] have criticized Schellman and Harvey's treatment of the disorder.
- [21] J. Bednar *et al.*, *J. Mol. Biol.* **254**, 579 (1995).
- [22] P. Furrer *et al.*, *J. Mol. Biol.* **266**, 711 (1997).
- [23] Some dynamical studies have yielded even larger values for  $A$ , presumably because energy barriers make transitions to some conformations very slow, effectively freezing them out on short time scales [L. Song and J. M. Schurr, *Biopolymers* **30**, 229 (1990)].
- [24] J. Marko and E. Siggia, *Macromolecules* **28**, 8759 (1995).
- [25] The analysis neglects the helical character of DNA, taking the elasticity as well as the disorder to be isotropic about the rod axis. This is a good approximation since the important fluctuations are on length scales several times longer than the helical repeat; see the appendix to [14]. We also neglect any random variation in the elastic constants themselves [see, e.g., B. S. Fujimoto and J. M. Schurr, *Nature (London)* **344**, 175 (1990)]; sequence effects enter only via random preferred angles from each segment to the next. We neglect self-avoidance effects, a good approximation at moderately high stretching force [13,14]. Finally, we will restrict to forces  $f \ll 1000$  pN so that the extensibility of the DNA duplex itself is negligible [7].
- [26] Equation (3) is reminiscent of a model of membranes with random bends [D. C. Morse and T. C. Lubensky, *J. Phys. II (France)* **3**, 531 (1993)]. It differs from the approach of [16], where each link pivots freely in a cone whose preferred polar angle  $\theta_0(s)$  was random. As emphasized by Schellman and Harvey, this idealization can affect the calculation [20]. Marko and Siggia [24] considered a mechanical model of a chain of semicircles of radius  $P$ ; this model does not have random curvature at all.
- [27] For intensive quantities such as  $Z/L$  our averaging procedure is equivalent to the actual case of a single, very long, sequence.
- [28] This change of variables is also the appropriate one to use when an external torque is applied at the ends of the rod. The torque couples to the resulting change in linking number  $2\pi\Delta Lk$ . At high enough force the DNA will not double back on itself, and so the change in  $2\pi Lk$  may be written in the local form  $\int \dot{\psi} + \dot{\phi} + \dot{\zeta}_3$ . This formula is just the one given in B. Fain, J. Rudnick, and S. Östlund, *Phys. Rev. E* **55**, 7364 (1996), with  $\dot{\zeta}_3$  added in order to measure the *change* in Link from the unstressed value. The same shift in the definition of  $\psi$  just used in the text thus eliminates  $\dot{\zeta}_3$  here as well.
- [29] T. C. Lubensky, in *Ill-Condensed Matter*, edited by R. Balian, R. Maynard, and G. Toulouse (Elsevier, New York, 1979).