

6-1998

# Sequence-Disorder Effects on DNA Entropic Elasticity

Philip C. Nelson
*University of Pennsylvania*, nelson@physics.upenn.edu



Recommended Citation
Nelson, P. C. (1998). Sequence-Disorder Effects on DNA Entropic Elasticity. *Physical Review Letters, 80* 5810-5812.
http://dx.doi.org/10.1103/PhysRevLett.80.5810

# Sequence-Disorder Effects on DNA Entropic Elasticity

**Abstract**
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.

**Disciplines**
Physical Sciences and Mathematics | Physics

UPR-781T

# Sequence Effects on DNA Entropic Elasticity

Philip Nelson

*Department of Physics and Astronomy, University of Pennsylvania*

*Philadelphia, PA 19104 USA*

(25 November 1997)

## Abstract

DNA stretching experiments are usually interpreted using the worm-like chain model; the persistence length $A$ appearing in the model is then interpreted as the elastic stiffness of the double helix. In fact the persistence length 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* persistence length deduced from torsionally-constrained stretching experiments suffers no such correction. Our calculation is very simple and analytic; it applies to DNA and other polymers with weak intrinsic disorder.

*Introduction and Summary:* 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 bonding 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.* found that a very simple model, the "worm-like chain", fit the force-extension data over four orders of magnitude in $f$ [5]. The model attributes to DNA just one parameter, the bend persistence length $A_{\rm eff}$; subsequent experiments have refined its value to [6] $A_{\rm eff} = 40\,{\rm nm}$ [7]. In a refinement of the technique, Strick *et al.* devised a *torsionally-constrained* stretching experiment [8,9]; analyses of the corresponding directed walk problem led to values of the twist persistence length $C_{\rm eff}$ between 75 and 120 nm [10–13]; in each case $C_{\rm eff}/A_{\rm eff}$ was found to exceed unity.

The purpose of this note is to show that the value $A_{\rm eff}$ 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 $A_{\rm eff}$ *underestimates* the true elastic stiffness $A$, while $C_{\rm eff}$ accurately reflects the true $C$, as announced in [12]. Thus the large observed value of $C_{\rm eff}/A_{\rm eff}$ is perhaps not as mysterious as it at first seems.

Recently Bensimon *et al.* have independently studied these and other issues [14]. Using a different model from ours, they found analytical formulæ 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 $A_{\rm eff}$ differs from [14], as described below. We will also retain the torsional degree of freedom needed to study the twist stiffness.

The result of this note 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 unit. The sequence of natural DNA has a small component with period equal to the helix repeat [15], but mainly the sequence imparts random natural bends to the rod [16]. Trifonov *et al.* noted that even in the absence of any thermal fluctuations a randomly-kinked rod would follow a random walk of some persistence length $P$, which they called the "static persistence length." They argued that the effective persistence length of such a coil at nonzero temperature would be [17]

$$A_{\rm eff} = A/(1+\lambda) \ , \tag{1}$$

where $A \cdot k_{\rm B}T$ is the true elastic stiffness of the rod and $\lambda \equiv A/P$, and they verified formula (1) with Monte Carlo simulations [18]. Trifonov *et al.* computed the numerical value $P = 216\,{\rm nm}$ 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 $A = 78\,{\rm nm}$, $A_{\rm eff} = 45\,{\rm nm}$, and hence $\lambda = 0.4$ [19,20] [21].

One might imagine that under extensional force the kinked rod would simply follow the usual worm-like chain result with $A$ replaced by $A_{\rm eff}$ from (1). Indeed *this is correct for weak disorder* (small $\lambda$). In contrast, Bensimon *et al.* found that at weak disorder $A_{\rm eff} = A(1 - \frac{1}{2}\sqrt{\lambda})$ [14], while Marko and Siggia argued that at high force disorder is immaterial: $A_{\rm eff} = A$ [22].

For a rod under torsional stress similarly define $C_{\rm eff}$ by the torque $\tau$ needed to change the linking number Lk to a value different from its relaxed value $\mathrm{Lk}_0$:

$$\tau = C_{\rm eff} \cdot k_{\rm B}T \cdot \omega_0 \frac{\mathrm{Lk} - \mathrm{Lk}_0}{\mathrm{Lk}_0} \ . \tag{2}$$

Here $\omega_0 = 1.85/\mathrm{nm}$ is the rotation per unit length of relaxed DNA. In contrast to (1) we have

$$C_{\rm eff} = C + \cdots \tag{3}$$

where the ellipsis denotes terms vanishing at high force or greater than first order in $\lambda$.

*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 [23].

To describe the rod conformations, let $\hat{E}_a(s)$ be 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 an orthonormal matrix $\mathsf{E}(s)$. Let $\mathbf{\Omega} \equiv \mathsf{E}^{-1}\dot{\mathsf{E}} \equiv \sum_i \Omega_i \mathsf{T}_i$, where the dot denotes $\mathrm{d}/\mathrm{d}s$. $\mathsf{T}_i$ are the three antisymmetric $3 \times 3$ matrices generating rotations, *e.g.* $[T_1]_{23} = +1$. The elastic energy of a conformation is then:

$$E_{\rm elas}/k_{\rm B}T = \frac{1}{2}\int_0^L \mathrm{d}s \, \left[A(\Omega_1 - \zeta_1)^2 + A(\Omega_2 - \zeta_2)^2 + C(\Omega_3 - \zeta_3)^2\right] \ . \tag{4}$$

To this energy we now add a term describing the work done by the external force,

$$-\frac{f}{k_{\rm B}T}\int \mathrm{d}s \, E_{33} \ . \tag{5}$$

The functions $\zeta_i(s)$ appearing in (4) specify the random kinks [24]. We give them an isotropic, Gaussian distribution:

$$[\![\zeta_i(s)]\!] = 0 \ ; \quad [\![\zeta_i(s)\zeta_j(s')]\!] = \frac{\lambda}{A}\delta(s-s') \begin{bmatrix} 1 & & \\ & 1 & \\ & & g \end{bmatrix}_{ij} \ . \tag{6}$$

Here the double brackets signify an average over an ensemble of many possible sequences [25]. 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 $[\![\hat{E}_3(0)\cdot\hat{E}_3(s)]\!] = 1-\frac{s}{P}+\mathcal{O}(s^2)$. The constant $g$ in eqn. (6) will drop out of our answers.

Thus even neglecting thermal undulations altogether, straightening the rod requires some extensional force to overcome the *elastic* energy (4). We must now introduce *entropic* effects as well, and compute the full extension

$$Z/L = [\![\langle E_{33}(0)\rangle]\!] \ , \tag{7}$$

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

$$\mathsf{E} = \mathrm{e}^{-\phi \mathsf{T}_3}\mathrm{e}^{-\theta \mathsf{T}_2}\mathrm{e}^{-\psi \mathsf{T}_3} \ . \tag{8}$$

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

$$E_{\mathrm{elas}}/k_{\mathrm{B}}T = \int \mathrm{d}s \left[ A(|\mathcal{W}|^2 - \mathcal{W}\mathcal{Z}^* - \mathcal{W}^*\mathcal{Z}) + \tfrac{1}{2}C(\dot{\psi} + \dot{\phi}\cos\theta + \zeta_3)^2 - \frac{f}{k_{\mathrm{B}}T}\cos\theta \right] \ . \tag{9}$$

We have dropped the divergent constant $\int |\mathcal{Z}|^2$ from (9) 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 last term of (9) by shifting the definition of $\psi$. Since $\psi$ does not enter the first term, while the next two terms already contain the disorder field $\mathcal{Z}$, this shift eliminates $\zeta_3$ altogether 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 reference frame rotated at $s$ by an angle $\int^s \mathrm{d}s'\,\zeta_3(s')$ relative to the old one.

What makes our problem interesting is that the disorder $\mathcal{Z}$ can *not* be so trivially eliminated, due to a clash between the $A$ terms and the $f$ term. To leading nontrivial order in the disorder strength $\lambda$ the $\mathcal{Z}$-terms of (9) contribute

$$1 + A^2 \int \mathrm{d}s\mathrm{d}s' \, \mathcal{W}(s)\mathcal{Z}^*(s)\mathcal{W}^*(s')\mathcal{Z}(s') \ , \tag{10}$$

to the Boltzmann weight $\mathrm{e}^{-E_{\mathrm{elas}}/k_{\mathrm{B}}T}$. Performing the average over the $\mathcal{Z}$ fields eliminates one of the integrations over $s$, so that the correction factor is the leading term of $\mathrm{e}^{A\lambda \int \mathrm{d}s \, |\mathcal{W}|^2}$. Comparing to (9), we see that to $\mathcal{O}(\lambda)$ the effect of disorder is simply to replace $A$ by $A_{\mathrm{eff}} = A(1-\lambda)$, leaving $C$ unchanged. This proves (1,3) 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 [26].

We can easily incorporate an external torque $\tau$ applied at the ends of the rod: $\tau$ couples to the change in Link density, which in our variables is simply $\dot{\psi} + \dot{\phi} + \zeta_3$. We added $\zeta_3$ to the formula of [27] in order to measure the *change* in Link from the unstressed value; the same shift in the definition of $\psi$ used earlier thus eliminates $\zeta_3$ here as well.

Thus within our approximations *the only effect of sequence on entropic elasticity is to reduce the effective bend persistence length*, as claimed in eqns. (1,3). The first correction to eqn. (3) in powers of $1/\sqrt{f}$ is also simple to obtain by substituting eqn. (1) into the formula for the effective stiffness given in [12,13] [28].

*Discussion:* The model investigated above may seem highly reductionist, neglecting as it does all 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 (6), 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. It also predicts that single-molecule stretching ex-

periments on long, intrinsically-straight DNA would show the same increase in effective persistence length seen at zero force, for example in [19,20]. 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 worm-like chain model to stretching experiments.

## ACKNOWLEDGMENTS

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 DMR95–07366.

# REFERENCES


[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 **389**, 251 (1997).

[4] R. H. Austin *et al.*, Physics 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 $A_{\text{eff}} \approx 47\,\text{nm}$ 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, preprint (1997).

[10] J. F. Marko and A. Vologodskii, Biophys. J. **73**, 123 (1997).

[11] C. Bouchiat and M. Mézard, preprint `cond-mat/9706050` (1997).

[12] J. D. Moroz and P. Nelson, Proc. Natl. Acad. Sci. USA **in press**, (1997).

[13] J. D. Moroz and P. Nelson, preprint (1997).

[14] D. Bensimon, D. Dohmi, and M. Mézard, preprint `cond-mat/9711051` (1997).

[15] E. Trifonov and J. Sussman, Proc. Natl. Acad. Sci. USA **77**, 3816 (1980).

[16] J. Widom, J. Mol. Biol. **259**, 579 (1996).

[17] 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.

[18] Later analytical calculations also supported the formula (J. A. Schellman and S. C. Harvey, Biophys. Chem. **55**, 95 (1995)), though Bensimon *et al.* have criticized Schellman and Harvey's treatment of the disorder [14].

[19] J. Bednar *et al.*, J. Mol. Biol. **254**, 579 (1995).

[20] P. Furrer *et al.*, J. Mol. Biol. **266**, 711 (1997).

[21] 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)).

[22] J. Marko and E. Siggia, Macromolecules **28**, 8759 (1995).

[23] 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 [13]. We also neglect any random variation in the elastic constants themselves (see *e.g.* B. S. Fujimoto and J. M. Schurr, Nature **344**, 175 (1990)); sequence effects enter only in random preferred angles from each segment to the next. We neglect self-avoidance effects, a good approximation at moderately high stretching force [12,13]. Finally we will restrict to forces $f \ll 1000\,\mathrm{pN}$ so that the extensibility of the DNA duplex itself is negligible [7].

[24] Eqn. (4) 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 [14], 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 [18]. Marko and Siggia considered a mechanical model of a chain of semicircles of radius $P$; this model does not have random curvature at all [22].

[25] For intensive quantities such as $Z/L$ our averaging procedure is equivalent to the actual

case of a single, very long, sequence.

[26] T. C. Lubensky, in *Ill-condensed matter*, edited by R. Balian, R. Maynard, and G. Toulouse (Elsevier, New York, 1979).

[27] B. Fain, J. Rudnick, and S. Östlund, Phys. Rev. **E55**, 7364 (1996).

[28] In fact it is simple to rederive the formulas for effective twist stiffness and Link-dependence of the extension to leading nontrivial order in $1/\sqrt{f}$, using (9) with $\mathcal{Z} = \zeta_3 = 0$. At higher orders, however, this method becomes cumbersome and the approach of [12,13] is much easier.