Statistical mechanics of double-helical polymers

Alvise De Col[1,2] and Tanniemola B. Liverpool[1,3]
[1]*Condensed Matter Theory Group, Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom*
[2]*Theoreticshe Physik, ETH Hönggerberg, CH-8093 Zürich, Switzerland*
[3]*Applied Mathematics, University of Leeds, Leeds LS2 9JT, United Kingdom*
(Received 30 October 2001; revised manuscript received 3 November 2003; published 2 June 2004)

We introduce a simple geometric model for a double-stranded and double-helical polymer. We study the statistical mechanics of such polymers using both analytical techniques and simulations. Our model has a single energy scale which determines both the bending and twisting rigidity of the polymer. The helix melts at a particular temperature T_c below which the chain has a helical structure and above which this structure is disordered. Under extension we find that for small forces, the behavior is very similar to wormlike chain behavior but becomes very different at higher forces.

DOI: 10.1103/PhysRevE.69.061907 PACS number(s): 87.15.By, 36.20.Ey, 61.25.Hq

Recent developments in single molecule manipulation techniques have led to the detailed study of the mechanical properties of DNA and other biomolecules as well as their response to applied fields. The model most used in the study of the large-scale properties of biopolymers is that of the wormlike chain [1] in which the polymer flexibility (structure) is determined by a single length, the persistence length L_p which measures the tangent–tangent correlations. For example, DNA has a persistence length $L_p \approx 50$ nm. Such coarse-grained models are needed to understand the *statistical* behavior of long chains with a *large* number of monomers. They are complementary to chemically specific models which describe accurately the *local* effects of external fields but generally cannot deal with long chains [2]. Generalizations of the wormlike chain to introduce twist degrees of freedom have also been proposed [3]. Whilst the wormlike chain model and its generalizations give a good account of DNA under small fields (perturbations), it fails when these perturbations become large, e.g., under tensional forces above 65 pN [4,5]. Such situations are biologically relevant in, for example, DNA replication and repair.

DNA is double stranded and helical. We introduce a geometric model which includes three important ingredients of DNA in B-form. First, the double-stranded nature of DNA given by the two phosphate backbones; second, the hydrogen bonds that keep the two strands together; and third, the interactions between the bases that bring about the twisted structure. Similar models could be used to study other double-helical polymers such as F-actin. The model is a generalization of the double-stranded semiflexible (ribbon) polymer introduced by Liverpool, Golestanian, and Kremer (LGK) [6] which takes into account the first two aspects but ignores the third. Other ribbon models have also been proposed by a number of authors [7]. It will be our conclusion that it is exactly this third property, the base-stacking interaction, which can account for many of the *elastic* properties of DNA. In our model, the *rigidity* of the individual strands is irrelevant for the effective persistence length, L_p, which is determined by the base-stacking interaction, unlike Z. Haijun *et al.* [8], who proposed another generalization of the LGK model with base-stacking interactions. In their model the bending rigidity of DNA is due to the elasticity of single strands. Our model is therefore consistent with the large difference between the persistence length of single- and double-stranded DNA measured experimentally. Our model can easily be extended to include excluded volume and electrostatic interactions which for the moment have been ignored [9]. We do not attempt to make quantitative comparison with experiments but point out qualitative differences between simpler models. Our aim is rather to suggest a minimal model required to understand experimental results.

The polymer is embedded in three-dimensional (3D) space. As for the ribbon, we see qualitatively different behavior for the high- and low-temperature regimes with a *melting*, which in this framework corresponds to an unwinding of the helix, at particular temperature $T_c \approx (4/9)^2[5+(2P/\pi a)^2/1+(2P/\pi a)^2]^{3/2}B/k_B a$, where P and a are, respectively, the pitch and diameter of the helix at $T=0$ and B the "stiffness" of the base-stacking interaction. Below T_c, there is a helical structure of the chain whilst above T_c, the helical structure vanishes. Unlike the ribbon, *both* the bend and twist rigidity of the double-helical polymer are determined by the base-stacking interaction and the stiffness of the individual strands is irrelevant. We also study the response of the double-helical polymer to an extensional force and find that for small forces, the mechanical response is very similar to that of a wormlike chain but becomes significantly different at higher forces.

The system is composed of two semiflexible chains modeling the sugar phosphate backbones, each with rigidity A, whose embeddings in three-dimensional space are defined by $\mathbf{R}_1(s)$ and $\mathbf{R}_2(s)$. Because of the covalent bonds ($E_{Cb} \sim 100 k_B T$ at room temperature) between the alternating phosphate and sugar molecules, the backbones are effectively inextensible at the forces and temperatures we consider. Similarly, the H bonds ($E_{Hb} \sim 10 k_B T$) between the two backbones are assumed to keep the distance between the chains constant.

The separation of the strands is fixed and given by a, and the H bonds give the constraint $\mathbf{R}_2(s)=\mathbf{R}_1(s)+a\mathbf{b}(s)$, where $\mathbf{b}(s)$ is a unit vector which we call the *bond-orientation vector* (see Fig. 1). The bond-orientation vector is perpendicular to both strands. The infinitesimal distance in three-

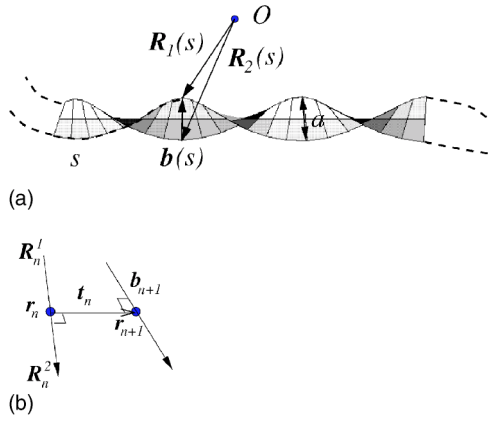

FIG. 1. (a) The schematic double-helical polymer with $\mathbf{R}_2(s) = \mathbf{R}_1(s) + a\mathbf{b}(s)$. The midsection (solid) is in the minimum energy configuration. (b) The discretization of the double-helix used in the simulation.

dimensional space between neighboring points on a single strand is a function of the local curvature, $\partial_s^2\mathbf{R}_i(s)$ of the strands, where $\partial_x A(x) \equiv \partial A/\partial x$. The base-stacking interaction can be modeled by a potential $V_i(\partial_s^2\mathbf{R}_i, \mathbf{b})$, where the subscript $i \in \{1,2\}$ refers to strand 1 or 2, which is a function of the curvatures of the chains and the bond-vector only, with a minimum given by a symmetric double helix. The Hamiltonian of the system is given by

$$\mathcal{H} = \int ds\{V_1(\partial_s^2\mathbf{R}_1, \mathbf{b}) + V_2(\partial_s^2\mathbf{R}_2, \mathbf{b}) + A(|\partial_s^2\mathbf{R}_1|^2 + |\partial_s^2\mathbf{R}_2|^2)\}. \tag{1}$$

The simplest example of such a potential which we will use for the remainder of this paper is the quadratic $V_1(\partial_s^2\mathbf{R}_1, \mathbf{b}) = B/2(\partial_s^2\mathbf{R}_1 - H\mathbf{b})^2$ and $V_2(\partial_s^2\mathbf{R}_2, \mathbf{b}) = B/2(\partial_s^2\mathbf{R}_2 + H\mathbf{b})^2$. This is a potential of *stiffness* $B/2$ with a minimum when equivalent points on the two strands have equal and opposite curvatures, $\pm H$. More complicated potentials with several minima are also possible (see later). We note that we have not included the asymmetry of the DNA helix but a similar but more complicated model could be defined for an asymmetric double helix [9]. We implement the model by introducing the "midcurve" $\mathbf{r}(s)$: $\mathbf{R}_1(s) = \mathbf{r}(s) + (a/2)\mathbf{b}, \mathbf{R}_2(s) = \mathbf{r}(s) - (a/2)\mathbf{b}$. In terms of the tangent to the midcurve $\mathbf{t}(s) = \partial_s\mathbf{r}$ and the bond director $\mathbf{b}$ the Hamiltonian of the system can now be written as

$$\mathcal{H}[\mathbf{t},\mathbf{b}] = \frac{B}{2}\int ds \quad \left[\left(\partial_s\mathbf{t}(s) + \frac{a}{2}\partial_s^2\mathbf{b}(s) - H\mathbf{b}\right)^2 + \left(\partial_s\mathbf{t}(s) - \frac{a}{2}\partial_s^2\mathbf{b}(s) + H\mathbf{b}\right)^2\right], \tag{2}$$

subject to the exact (local) constraints

$$\left(\mathbf{t} \pm \frac{a}{2}\partial_s\mathbf{b}\right)^2 = 1, \quad \mathbf{b}^2 = 1, \quad \left(\mathbf{t} \pm \frac{a}{2}\partial_s\mathbf{b}\right)\cdot\mathbf{b} = 0 \tag{3}$$

corresponding to the chain inextensibility, fixed distance between strands and definition of the bond vector, respectively. We make the assumption that $B \gg A$ since the bending rigidity of single-stranded DNA is very small, $(A/k_BT < 10\text{Å})$, compared to the double-stranded form. We can define a typical length $\ell = B/k_BT$, which together with the strand separation a and equilibrium radius of curvature H^{-1} form the three relevant lengths of the problem. This completes the formulation of the model.

Ground state behavior. We can calculate the ground state of this system from the condition $\mathcal{H} = 0$ which gives, $\mathbf{R}_1(s) = (a/2)\cos(\Theta s)\hat{e}_1 + a/2\ \sin(\Theta\ s)\hat{e}_2 + s\sqrt{1-(a^2\Theta^2/4)}\ \hat{e}_3$ and $\mathbf{R}_2(s) = -(a/2)\cos(\Theta\ s)\hat{e}_1 - (a/2)\sin(\Theta\ s)\hat{e}_2 + s\sqrt{1-(a^2\Theta^2/4)}\ \hat{e}_3$, where $\Theta = \sqrt{2H/a}$. This is a double helix with a pitch: $P = 2\pi/\Theta\sqrt{1-(a^2\Theta^2/4)}$ and the bond-director field, $\mathbf{b}(s) = -\cos(\Theta\ s)\hat{e}_1 - \sin(\Theta\ s)\hat{e}_2$, where $\{\hat{e}_i\}$ are an orthonormal set of unit vectors.

Finite T behavior. Inspite of the simplification of the model from using constraints, an exact analytic expression for the partition function of the model is still not available. For our analytic calculations, we use a very useful approximation [10] which describes correctly the equilibrium behavior of correlation functions and even distribution functions, if the end effects are taken into account carefully [11]. This is a mean field approach that relaxes the local constraints to global ones [10,11]. It may be thought of as a self-consistent theory and corresponds to the saddle-point evaluation of path integrals over the lagrange multipliers [12]. This approach fails for some dynamical properties [13] which are outside the scope of this paper. To implement this, we add a variational term to our Hamiltonian

$$\mathcal{H}_m/k_BT = \int ds((b/\ell)[\mathbf{t} - (a/2)\partial_s\mathbf{b}]^2 + (b/\ell)[\mathbf{t} + (a/2)\partial_s\mathbf{b}]^2 + (ca^2/4\ell^3)\mathbf{b}^2 + e/\ell[\mathbf{t} - (a/2)\partial_s\mathbf{b}]\cdot\mathbf{b} + e/\ell[\mathbf{t} + (a/2)\partial_s\mathbf{b}]\cdot\mathbf{b}),$$

where b, c, and e are dimensionless constant Lagrange multipliers. The details of the calculations are similar to those in LGK [6].

Two useful dimensionless geometric parameters $u = Ha/2$ and $v = 4(\ell/a)^2$ can be used to characterize the double-helical polymer. Note that v is proportional to T^{-2} and can be viewed as a measure of temperature. We then determine the constants self-consistently by demanding the constraints of Eq. (3) to hold on average, where the thermal average is calculated by using the total Hamiltonian $\mathcal{H} + \mathcal{H}_m$. Self-consistency leads to the following set of equations for the constants b and c: $(c+u^2v^2)[(b-uv) + \sqrt{c+u^2v^2}] = 9v^2/32; (1/4\sqrt{2b}) + \frac{1}{3}(1/v)\sqrt{c+u^2v^2} = \frac{1}{3}$ and $e = 0$. The above equations, which are nonlinear and difficult to solve exactly, determine the behavior of b and c functions of u and v. One can solve the equations analytically in some limiting cases corresponding to $T \to 0$ and $T \to \infty$. For $v \to \infty (T \to 0)$ and $H \neq 0$, we find $b \sim \frac{9}{32}[1/(1-u)^2], c \sim \frac{9}{16}[(1/u^2) - (1/(1-u)^2)]uv$ and for $v \to 0 (T \to \infty)$ and $H \neq 0$, we obtain $b \sim \frac{9}{8}, c \sim [(1/4u^2) - 1]u^2v^2$. A full solution requires a numerical treatment and shows simple monotonic behavior for both c and b [15].

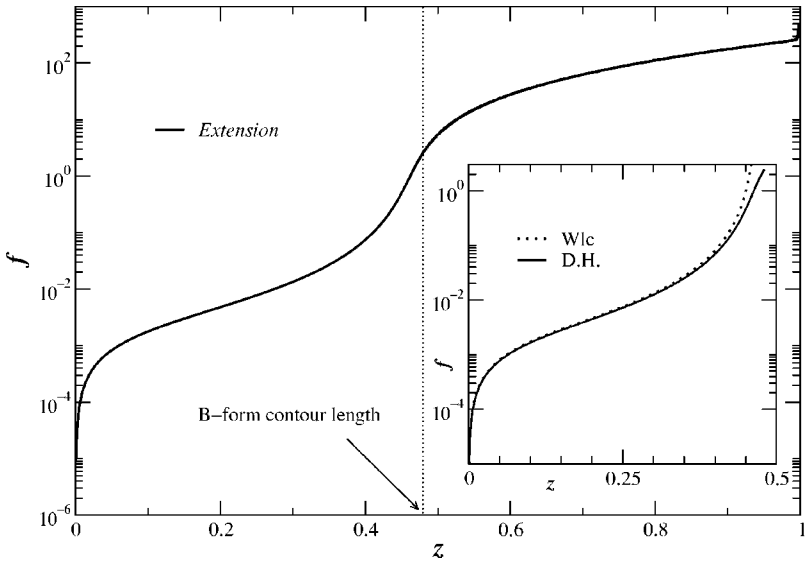

FIG. 2. Force-extension relation for the double-helical polymer. This is a log plot with $u=0.77$ and $v=115000$, typical values for B-DNA which corresponds to a low temperature $T<T_c$ (note $f=Fa/8$). The dashed line shows the $T=0$ contour length $\sqrt{1-a^2\Theta^2/4}$. Inset: comparison with wormlike chain at small forces. f and z are dimensionless quantities.

We can then calculate the correlation functions. For the tangent–tangent correlation one obtains

$$\langle \mathbf{t}(s)\cdot\mathbf{t}(0)\rangle = \frac{3}{4\sqrt{2b}} e^{-\sqrt{2b}(s/\ell)} \overset{T\to 0}{\approx} (1-u)e^{-[3/4(1-u)](s/\ell)}, \quad (4)$$

whereas for the bond-director field one obtains

$$\langle \mathbf{b}(s)\cdot\mathbf{b}(0)\rangle = \frac{3i}{2a^2}\frac{\ell^2}{\ell^2\sqrt{D}}\left[\frac{e^{i\sqrt{(uv-b)+\sqrt{D}\ell^2}(s/\ell)}}{\sqrt{(uv-b)+\sqrt{D}\ell^2}} + \frac{e^{-i\sqrt{(uv-b)-\sqrt{D}\ell^2}s/\ell}}{\sqrt{(uv-b)-\sqrt{D}\ell^2}}\right] \overset{T\to 0}{\approx} e^{-(3/8u)(s/\ell)}\cos\left(\sqrt{\frac{2H}{a}}s\right). \quad (5)$$

We have defined a discriminant $D(u,v)=(1/\ell^2)\sqrt{(b-uv)^2-(c+u^2v^2)}$. The tangent–tangent correlation [Eq. (4)] of the double-helical polymers is similar to that of a wormlike chain but with a length rescaling factor, and decays exponentially at all temperatures. Equation (5) on the other hand, indicates a change of behavior at $D=0$, i.e., $(b-uv)^2=c+u^2v^2$ for the bond-director correlation. The correlation is *overdamped* for $(b-uv)^2>c+u^2v^2$ (high temperatures), while it is *underdamped* (oscillatory) for $(b-uv)^2<c+u^2v^2$ (low temperatures). The interesting point $(b-uv)^2=c+u^2v^2$ happens for $v_c\simeq 9[9/4(1+4u)]^3$, which leads to the value for T_c quoted above.

We emphasize that it is not a thermodynamic phase transition in the sense of long-range ordering and broken symmetry. It is a crossover that appears due to competing effects, and the transition is from a state with some short-range order to a state with a different short-range order. The crossover (*transition*) point corresponds to a "*Lifshitz line*" for a 1D system with competing interactions [6,14]. A Lifshitz line is the boundary between a state with short-range antiferromagnetic order and no order found in 1D systems with competing interactions [6,14]. At high temperatures, in addition to the effects we describe we also expect the H bonds between the bases to break, leading to *denaturation* which we cannot treat in this model. In the low T regime, we can read off the bend (tangent) persistence length $L_{TP}=\frac{4}{3}\ell(1-Ha/2)^{-1}$ and twist (bond) persistence length, $L_{BP}=\frac{4}{3}\ell Ha$ and helical frequency $(2H/a)^{1/2}$. Both the bend and twist rigidity of the double-helical polymer are given by the same energy scale $B=k_BT\ell$ and have little to do with the stiffness of the individual strands.

Force extension: unwinding the helix. Under a stretching force, which without loss of generality we orient along the z axis $\mathbf{F}=F\hat{\mathbf{z}}$, the Hamiltonian becomes $\mathcal{H}_F=\mathcal{H}-\mathbf{F}\cdot\int ds\partial_s\mathbf{r}$. Now $\langle t_z\rangle=\langle \mathbf{t}\cdot\hat{\mathbf{z}}\rangle\neq 0$ and we must use connected correlation functions to calculate the self-consistent equations. Defining the dimensionless force $f=(Fa/8k_BT)$ we obtain the equations

$$\frac{1}{4\sqrt{2b}}+\frac{1}{3}\left(\frac{f}{b}\right)^2 v+\frac{1}{3}\frac{1}{v}\sqrt{c+u^2v^2}=\frac{1}{3},$$

$$(c+u^2v^2)[(b-uv)+\sqrt{c+u^2v^2}]=\frac{9v^2}{32}. \quad (6)$$

We can solve Eqs. (6) for $b(u,v,f)$ and $c(u,v,f)$ and the extension per unit length $z(u,v,f)$ is simply

$$z=\frac{1}{L}\int_0^L \langle t(s)\rangle ds=\frac{f}{b}\sqrt{v}. \quad (7)$$

We can therefore obtain the force-extension relation for the double-helical polymer. Equation (6) has been solved numerically and in Fig. 2 we have plotted z against f. In the inset we compare the force-extension curve to that of the WLC and find they are identical for small forces. We note the interesting fact that the mean-field model can also be used to calculate the force-extension behavior of the WLC. It has the simple compact form $fL_p=[3z/2(1-z^2)^2]$, where $f=F/k_BT$ and L_p is the persistence length. We find that this simple expression has an error of less than 4% compared to

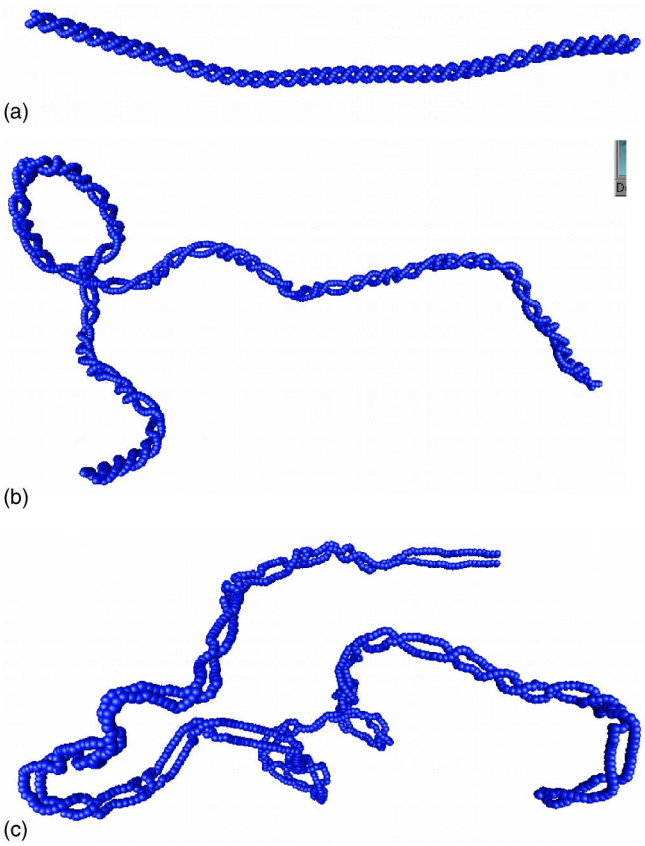

FIG. 3. Typical conformations at (a) low, (b) intermediate, and (c) high temperatures.

the "exact" value of Marko and Siggia [11]. We take this as evidence for the utility of the mean-field technique for force-extension calculations. The response of the polymer to external torque, we leave as a subject for further work [9].

Simulations. We also performed Monte Carlo simulations of the double-helical polymers. We discretized the "helical ribbon" and implemented all the constraints in units of length Δ (see Fig. 1). The discretized Hamiltonian is given by

$$\beta\mathcal{H} = \ell \sum_{n=1}^{N-1} \left(\frac{2}{\Delta} + H^2\Delta - \sum_{j=1}^{2} \left[\frac{\mathbf{t}_{n+1}^j \cdot \mathbf{t}_n^j}{\Delta} + (-1)^j H \mathbf{b}_n \cdot (\mathbf{t}_{n+1}^j - \mathbf{t}_n^j) \right] \right), \quad (8)$$

where n is the monomer label (position along the chain) and $j \in \{1,2\}$ the strand label (see Fig. 1). Since the Hamiltonian is local, we grew chains with a local algorithm. In our simulation the number of monomers N was taken to be 1000 for each chain, $\Delta = 1$ and $a = 3$. Typical conformations at high and low temperatures are shown in Fig.3 .

Typical bond correlation functions are plotted in Fig. 4. From the bond and tangent correlation functions, we can obtain the bond and tangent persistence lengths. The "generalized stiffnesses" are obtained from the slope of the log of the (envelope) correlation functions. These stiffnesses are plotted against temperature ($\ell \propto T^{-1}$) in order to obtain the

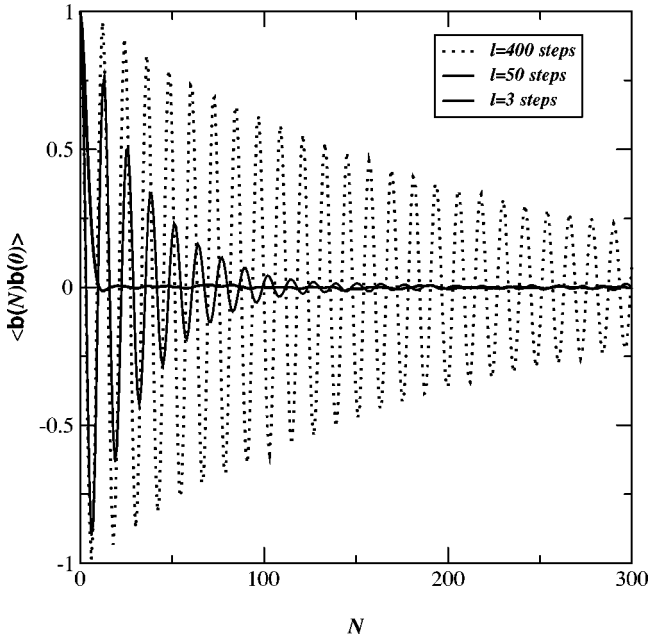

FIG. 4. The $\langle \mathbf{b}(s) \cdot \mathbf{b}(0) \rangle$ correlation function measured in the simulations for temperatures $\ell = 400, 50, 3$ and $u = 0.6$ corresponding to $T < T_c$ and $T > T_c$. The averages were done over 200 statistically independent samples.

bond correlation length as $L_{BP} = C_B u \ell$ and the tangent correlation length as $L_{TP} = C_T \ell$ with the constants given $C_B = 1.09 \pm 0.09$ and $C_T = 2.46 \pm 0.73$, respectively. A comparison with the mean-field results shows that the results are the same to within a constant. This difference can be considered as due to finite corrections to the mean-field solution due to higher order terms in the $1/d$ expansion.

In conclusion, we have studied the properties of a well-defined model of a double-helical, double-stranded semiflexible polymer using a mean-field analytical approach as well as extensive Monte Carlo simulations. We have shown nontrivial differences between the high, low, and zero temperature behavior. A detailed comparison with the simulations and the mean-field solution shows qualitative and almost quantitative agreement. The fine structure studied in this model gives a different behavior compared to a WLC. We see this most clearly when the double-helical polymer is subject to an external pulling force. It shows two distinct regimes. For low forces the dynamical response is similar to a WLC, with the most relevant energetic contribution given by the bending rigidity. At higher forces, when the $T = 0$ contour length is reached, the extension changes abruptly and increases steeply as the double-helical structure is *unwound*. This *qualitatively* agrees with the experimental results on overstretched DNA [4,5], which nevertheless have a sharper transition and a flatter plateau at a larger z than the ground-state value. The graph is plotted on a log scale, making clearer the complex behavior of the force-extension curve. There are two obvious ways in which a sharper transition could be obtained. First, we have used a simple quadratic potential for the base-stacking interaction corresponding only to an expansion about the minimum energy conformation. This leads to the lack of cooperativity we observe for

the base-stacking transition in our model. A more realistic potential with a cutoff and/or several minima could change the sharpness and number of transitions, as well as the value of the extension at which the transition takes place. In addition one could imagine, as in mechanical unzipping experiments on DNA, there could be cooperative effects due to the effects of sequence disorder. Finally the effects of solvent could play an important role in cooperativity.

After this work was submitted we became aware of the work by B. Mergell *et al.* [16] which finds no **b** oscillations in a ribbon model similar to the LGK model. We emphasize that ribbon models are very different to the model presented here which unlike the ribbon has a nonzero helical pitch at $T=0$.

We have benefited from discussions with R. Everaers, R. Golestanian, and K. Kremer. The financial support of the Royal Society and the National Science Foundation under Grant No. PHY-99-07949 (at KITP) is gratefully acknowledged.

[1] O. Kratky and G. Porod, Recl. Trav. Chim. Pays-Bas **68**, 1106 (1949).
[2] W. K. Olson and V. B. Zhurkin, Curr. Opin. Struct. Biol. **10**, 286 (2000).
[3] J. F. Marko and E. D. Siggia, Macromolecules **27**, 981 (1994); C. Bouchiat and M. Mezard, Phys. Rev. Lett. **80**, 1556 (1998); S. Panyukov and Y. Rabin, *ibid.* **85**, 2404 (2000).
[4] S. B. Smith, Y. Cui, and C. Bustamante, Science **271**, 795 (1996).
[5] P. Cluzel *et al.*, Science **271**, 792 (1996).
[6] T. B. Liverpool, R. Golestanian, and K. Kremer, Phys. Rev. Lett. **80**, 405 (1998); T. B. Liverpool and R. Golestanian, Phys. Rev. E **62**, 5488 (2000).
[7] R. Everaers, R. Bundschuh, and K. Kremer, Europhys. Lett. **29**, 263 (1995); I. A. Nyrkova, A. N. Semenov, and J-F. Joanny, J. Phys. II **6**, 1411 (1996); S. Kumar and J. Singh, J. Stat. Phys. **89**, 981 (1997); E. J. Janse van Rensburg *et al.*, *ibid.* **85**, 103 (1996).
[8] Z. Haijun, Y. Zhang, and Z.-C. Ou-Yang, Phys. Rev. Lett. **82**, 4560 (1999).
[9] T. B. Liverpool (unpublished).
[10] J. B. Lagowski, J. Noolandi, and B. Nickel, J. Chem. Phys. **95**, 1266 (1991); A. M. Gupta, and S. F. Edwards, *ibid.* **98**, 1588 (1993); B-Y Ha and D. Thirumalai, *ibid.* **103**, 9408 (1995).
[11] J. F. Marko and E. D. Siggia, Macromolecules **28**, 8759 (1995); H. Flyvbjerg, e-print cond-mat/0103417; D. Thirumalai and B-Y Ha, in *Theoretical and Mathematical Models in Polymer Research*, edited by A. Grosberg (Academic, New York, 1998).
[12] F. David and E. Guitter, Europhys. Lett. **5**, 709 (1988); T. B. Liverpool and S. F. Edwards, J. Chem. Phys. **103**, 6716 (1995).
[13] T. B. Liverpool and A. C. Maggs, Macromolecules **34**, 6064 (2001); D. C. Morse, Phys. Rev. E **58**, R1237 (1998).
[14] R. M. Hornreich, R. Liebmann, H. G. Schuster, and W Selke, Z. Phys. B **35**, 91 (1979).
[15] A. De Col and T. B. Liverpool (unpublished).
[16] B. Mergell, M. R. Ejtehadi, and R. Everaers, Phys. Rev. E **66**, 011903 (2002).