Monte Carlo Implementation of Supercoiled Double-Stranded DNA

Zhang Yang,*1 Zhou Haijun,*‡ and Ou-Yang Zhong-can*§

*Institute of Theoretical Physics, Academia Sinica, P.O. Box 2735, Beijing 100080, China, †Institut für Theoretische Physik, FU Berlin, Arnimallee 14, 14195 Berlin, Germany, ‡The State Key Laboratory of Scientific and Engineering Computing, Beijing 100080, China, and §Center for Advanced Study; Tsinghua University, Beijing 100084, China

ABSTRACT Metropolis Monte Carlo simulation is used to investigate the elasticity of torsionally stressed double-stranded DNA, in which twist and supercoiling are incorporated as a natural result of base-stacking interaction and backbone bending constrained by hydrogen bonds formed between DNA complementary nucleotide bases. Three evident regimes are found in extension versus torsion and force versus extension plots: a low-force regime in which over- and underwound molecules behave similarly under stretching; an intermediate-force regime in which chirality appears for negatively and positively supercoiled DNA and extension of underwound molecule is insensitive to the supercoiling degree of the polymer; and a large-force regime in which plectonemic DNA is fully converted to extended DNA and supercoiled DNA behaves quite like a torsionless molecule. The striking coincidence between theoretic calculations and recent experimental measurement of torsionally stretched DNA (Strick et al., Science, 271:1835, 1996; Biophys. J. 74:2016, 1998) strongly suggests that the interplay between base-stacking interaction and permanent hydrogen-bond constraint takes an important role in understanding the novel properties of elasticity of supercoiled DNA polymer.

INTRODUCTION

Recent years have witnessed a remarkably intense experimental and theoretical activity in searching for the elasticity of a single supercoiled DNA molecule (see, e.g. Strick et al., 1996, 1998; Fain et al., 1997; Vologodskii and Marko, 1997; Moroz and Nelson, 1997; Bouchiat and Mezard, 1998; Zhou et al., 1999). Within a cell, native double-stranded DNA (dsDNA) often exists as a twisted and heavily coiled, closed circle. Differing amount of supercoiling, in addition to affecting the packing of DNA within cells, influences the activities of proteins that participate in processes—such as DNA replication and transcription—that require the untwisting of dsDNA (Wu et al., 1988). It is believed that changes in supercoiling can also promote changes in DNA secondary structure that influences the binding of proteins and other ligands (Morse and Simpson, 1988).

In recent experiments (Strick et al., 1996, 1998) on single torsonally constrained DNA molecules it was found that the supercoiling remarkably influences the mechanical property of DNA molecules. When applied with relatively low stretching force, a supercoiled molecule can reduce its torque by writhing, forming structures known as plectonemes. Therefore, the distance between two ends of the polymer decreases with increasing supercoiling. However, above a certain critical force \( f_c \), this dependence of extension on supercoiling disappears. More strikingly, the value of \( f_c \) is significantly different for positively and negatively supercoiled DNA, i.e. \( f_c \) ~ 0.8 pN for underwound molecule and \( f_c \) ~ 4.5 pN for overwound ones. On the theoretical side, harmonic twist elasticity and bending energy according to the wormlike chain model have been used to understand supercoiling of DNA polymer (Fain et al., 1997; Vologodskii and Marko, 1997; Moroz and Nelson, 1997; Bouchiat and Mezard, 1998), and some qualitative mechanic features of plectonemic structures of supercoiled DNA polymer have been described by the harmonic twist elasticity. However, because of the chiral symmetry of harmonic twist elasticity, the asymmetry of elastic behaviors of supercoiled DNA cannot be understood by this model, and, especially, the three obvious mechanic regimes observed in experiment of supercoiling DNA (Strick et al., 1996, 1998) still need better understanding. To understand the supercoiling property and the highly extensibility of DNA, we have developed a more realistic model in which the double-stranded nature of DNA is taken into account explicitly (Zhou et al., 1999). The supercoiling property of highly extended DNA was investigated analytically. Here, we aim at performing a thorough and systematic investigation into the property of supercoiled DNA by using Monte Carlo simulations based on this model.

As we have known, the bending energy of DNA polymer is mainly associated with the covalent bonding between neighboring atoms of DNA backbone (Nossel and Lecar, 1991). In our previous work (Zhou et al., 1999), van der Waals interactions between adjacent basepairs was introduced, and this helps to explain the highly cooperative extensibility of overstretched DNA (Cruzel et al., 1996; Smith et al., 1996). It has also been shown that the short-range base-stacking interaction takes a significant role in determining the elastic property of DNA. Lennard–Jones type potential between adjacent basepairs can be written as

\[
U(\theta) = \begin{cases} 
\epsilon \left( \frac{\cos \theta_0}{\cos \theta} \right)^{12} - 2 \left( \frac{\cos \theta_0}{\cos \theta} \right)^{6}, & \text{for } \theta > 0, \\
\epsilon [\cos^2 \theta_0 - 2 \cos^2 \theta_0], & \text{for } \theta \leq 0,
\end{cases}
\]  

(1)
(see also Fig. 1). The folding angle $\theta$ of the sugar-phosphate backbones around DNA central axis is associated with the steric distance $r$ of adjacent basepairs by $r = r_0 \cos \theta$, where $r_0$ is the backbone arc length between adjacent bases. The asymmetric potential related to positive and negative folding angle $\theta$ in Fig. 1 ensures a native DNA to take a right-handed double-helix configuration with its equilibrium folding angle $\theta_{eq} \sim \theta_0$. This double-helix structure is anticipated to be very stable because $\epsilon (\sim 14 k_B T)$ is much higher than thermal energy $k_B T$ according to the results of quantum chemical calculations (Saenger, 1984).

In the case in which DNA polymer is torsionally constrained, the basepair folding angle will deviate from the equilibrium position $\theta_{eq}$. However, if the stretching force is very small, the folding angle may deviate from $\theta_{eq}$ only slightly. This is because, as we can infer from Fig. 1, the base-stacking potential is very sharp around $\theta_0$, and a relatively large force is needed to make $\theta$ deviate considerably from its equilibrium value. It is reasonable for us to anticipate that a supercoiled DNA under low stretching force will convert its excess or deficit linking number into positive or negative writhing of its central axis. Because the central axis is symmetric with respect to positive or negative writhing, the elastic response of DNA at this force regime will certainly be symmetric with a positive or negative degree of supercoiling. Only when the stretching force becomes large enough will the chirality of supercoiled DNA appear. In this regime, it becomes more and more difficult for the central axis to writhe to absorb the linking number, and an increasing portion of the linking number will be converted to a twisting number of the backbones, which will certainly change the twisting manner of dsDNA. Because Eq. 1 shows that, for dsDNA, untwisting is much easier than overtwisting, chiral behavior is anticipated to emerge. This opinion is consistent with the experimental result of Strick et al. (1996).

In this paper, we investigate the mechanical properties of supercoiled DNA by numerical Monte Carlo method. Base-stacking van der Waals interactions between adjacent basepairs are incorporated by introducing the new degree of freedom, namely the folding angle $\theta$. A fundamental difference from the previous approaches (see, e.g., Vologodskii and Marko, 1997), which try to include the twist degrees of freedom by adding extra terms to the free energy, is that the twist and supercoiling are treated as the cooperative result of base stacking and backbone bending constrained from permanent basepairs. The striking coincidence between theoretical calculations and experimental data of supercoiling DNA (Stick et al., 1996, 1998) indeed confirms this treatment.

**MODEL AND METHOD OF CALCULATION**

In the simulation, the double-stranded DNA molecule is modeled as a chain of discrete cylinders, or two discrete wormlike chains constrained by basepairs of fixed length $2R$ (Fig. 2). The conformation of DNA molecule of $N$ straight cylinder segments is specified by the space positions of vertices of its central axis, $r_i = (x(i), y(i), z(i))$ in three-dimensional Cartesian coordinate system, and the folding angle of the sugar-phosphate backbones around the central axis, $\theta_i, i = 1, 2, \ldots, N$. Each segment is
assigned the same amount of basepairs, \( n_{bp} \), so that the length of the \( i \)th segment satisfies
\[
\Delta s_i = \| \mathbf{r}_i - \mathbf{r}_{i-1} \| = 0.34 n_{bp} \frac{\cos \theta_i}{\left( \cos \theta_0 \right)^i},
\]
where \( \langle \cdot \rangle_0 \) means the thermal average for a relaxed DNA molecule. Moreover, bearing in mind the experimental fact that there are \( \approx 10.5 \) basepairs for each turn of a native double helix DNA and the average distance between the adjacent basepairs is about \( d_0 = 0.34 \text{ nm} \), we have set the basepair length as \( 2R = 10.5d_0/\pi(\tan \theta_0) \) in our model.

Metropolis Monte Carlo method (Metropolis et al., 1953) is used to simulate the equilibrium evolution procedure of torsionally stretched dsDNA molecules. At each step of the simulation procedure, a trial conformation of the chain is generated by a movement from the previous one. The starting configuration is chosen arbitrarily (except that some topology and boundary conditions should be satisfied, see below) and the averaged results of equilibrium ensemble are independent of the initial choice after numerous movements. The probability of acceptance of the movement depends on the difference in energy between the trial and the current conformations, according to the Boltzmann weight. When a trial movement is rejected, the current conformation should be counted once more. This procedure is repeated numerous times to obtain an ensemble of conformations that, in principle, is representative of the equilibrium distribution of DNA conformation.

### The DNA model

As we have known, double strand DNA is formed by winding two polynucleotide backbones right-handedly around a common central axis. Between the backbones, nucleotide basepairs are formed with the formation of hydrogen bonds between complementary bases. In our continuous model (Zhou et al., 1999), the embeddings of two backbones are defined by \( \mathbf{r}_1(s) \) and \( \mathbf{r}_2(s) \). The ribbon structure of DNA is enforced by having \( \mathbf{r}_2(s') \) separated from \( \mathbf{r}_1(s) \) by a distance \( 2R \), i.e. \( \mathbf{r}_2(s') = \mathbf{r}_1(s) + 2R \mathbf{b}(s) \) where the hydrogen-bond-director unit vector \( \mathbf{b}(s) \) points from \( \mathbf{r}_1(s) \) to \( \mathbf{r}_2(s') \). As the result of the wormlike backbones, the bending energy of two backbones can be written as
\[
E_B = \frac{\kappa}{2} \int_0^L \left[ \left( \frac{d\mathbf{r}_1}{ds} \right)^2 + \left( \frac{d\mathbf{r}_2}{ds} \right)^2 \right] ds.
\]

The formation of basepairs leads to rigid constraints between the two backbones and, at the same time, hinder considerably the bending freedom of DNA central axis because of the strong steric effect. In the assumption of permanent hydrogen bonds (Everaers et al., 1995; Liverpool et al., 1998; Zhou et al., 1999), \( |s' - s| = 0 \). The relative sliding of backbones is prohibited, and the basepair orientation lies perpendicular to the tangent vectors \( \mathbf{t}_1 = \frac{d\mathbf{r}_1}{ds} \) and \( \mathbf{t}_2 = \frac{d\mathbf{r}_2}{ds} \) of the two backbones and that of the central axis, \( \mathbf{t} : \mathbf{b} = -\mathbf{t}_1 \times \mathbf{t}_2 = \mathbf{b} \times \mathbf{t} = 0 \). By defining the folding angle as half of the rotation angle from \( \mathbf{t}_2(s) \) to \( \mathbf{t}_1(s) \), i.e., the intersection angle between tangent vector of backbones \( \mathbf{t}_{1(2)} \) and DNA central axis \( \mathbf{t} \), we have
\[
\begin{align*}
\mathbf{t}_1 &= \cos \theta \mathbf{t} + \sin \theta \mathbf{b} \times \mathbf{t} \\
\mathbf{t}_2 &= \cos \theta \mathbf{t} - \sin \theta \mathbf{b} \times \mathbf{t}.
\end{align*}
\]

Therefore, the bending energy of the two backbones can be rewritten as
\[
E_B = \int_0^L \left[ \kappa \left( \frac{d\mathbf{t}_1}{ds} \right)^2 + \kappa \left( \frac{d\theta}{ds} \right)^2 + \kappa \frac{\sin^2 \theta}{R^2} \right] ds,
\]
where \( ds \) denotes arc-length element of the backbones, \( L \) the total contour length of each backbone, and \( \kappa \) is the persistence length of one DNA backbone. Bearing in mind that the pairing and stacking enthalpy of the bases significantly increase bending stiffness of the polymer axis, the experimental value of persistent length of dsDNA polymer is considerably larger than that of a DNA single strand (see, e.g. Smith et al., 1996). To incorporate the steric effect, and also considering the typical experimental value of persistent length of dsDNA \( p = 53 \text{ nm} \), the simplest way is to substitute \( k \) in the first term of Eq. 5 with a phenomenological parameter \( k^* = 53.0/2(\cos \theta_0) \text{ nm} k_B T \) (Zhou et al., 1999), hereafter this is assumed.

Taking into account Eqs. 1 and 5, the total energy of dsDNA molecule with \( N \) segments in our discrete computational model is expressed as
\[
E = \alpha \sum_{i=1}^{N-1} \gamma_i^2 + \alpha' \sum_{i=1}^{N-1} (\theta_{i+1} - \theta_i)^2
\]
\[
+ \frac{\kappa}{R} \sum_{i=1}^{N} \Delta s_i \sin^3 \theta_i \tan \theta_i
\]
\[
+ \sum_{j=1}^{N_{bp}} U(\theta_j) - f_z(N),
\]
where \( \gamma_i \) is the bending angle between the \((i - 1)\)th and the \( i \)th segments (Fig. 2), \( N_{bp} \) the total number of basepairs of DNA polymer, and \( f_z(N) \) is the total extension of the DNA central axis along the direction of the external force \( f \) (assumed in the \( z \)-direction).

Because the Kuhn statistical length of dsDNA polymer is associated with its bending stiffness (the Kuhn length is twice the persistence length of dsDNA polymer according to the wormlike chain model), one can decide bending rigidity parameter \( \alpha \) of the discrete chain accordingly. Suppose that we take the \( N \) discrete segments to simulate the behaviors of a dsDNA polymer of \( n \) Kuhn statistical length, the length of \( m = N n \) segments should correspond to one Kuhn statistical length. Therefore, for any chosen value \( m \), we can decide the bending rigidity parameter \( \alpha \) in the way (see Appendix)
\[
m = \frac{1 + (\cos \gamma)}{1 - (\cos \gamma)},
\]
where
\[
(\cos \gamma) = \frac{\int_0^\infty \cos \gamma \exp(-\alpha \gamma^2) \sin \gamma \, dy}{\int_0^\infty \exp(-\alpha \gamma^2) \sin \gamma \, dy}.
\]

In principle, the discrete DNA model becomes continuous only when \( m \) approaches infinity. The CPU time needed for a simulation, however, increases approximately as \( N^2 = (nm)^2 \). So it is necessary to choose a value of \( m \) that is large enough to ensure reliable results but small enough to keep the computational time within reasonable bounds. Our calculation and also previous work (Vologodskii et al., 1992) showed that simulated properties do not depend on \( m \) if it exceeds 8. Therefore, \( m = 8 \) was used in the current calculation, for which the bending constant \( \alpha = 1.895 k_B T \). Furthermore, we have chosen \( N = 160 \) in consideration of the feasible computer time. Because Kuhn statistical length of dsDNA is taken as 106 nm, the B-form length of the polymer in our simulation corresponds to
Equilibrium distance between a DNA dimer. The base-stacking intensity should be associated with stiffness of the DNA single strand. As an crude between adjacent basepairs (see Eq. 1),

The fourth term in Eq. 6 accounts for van der Waals interactions between adjacent basepairs (see Eq. 1). \( \theta_B (= 62^\circ) \) is related to the equilibrium distance between a DNA dimer. The base-stacking intensity \( \epsilon \) is generally influenced by the composition and sequence of nucleotide chains. For example, the solubility experiments in biphasic systems show that stacking interactions between purine and pyrimidine bases follow the trend

\[
\text{purine–purine} \succ \text{pyrimidine–purine} \succ \text{pyrimidine–pyrimidine}.
\]

Because we do not distinguish the specific base sequence of purine and pyrimidine in our DNA model, we take the statistic average of stack energies as \( \epsilon = 14k_B T \) according to the result of quantum chemical calculations (Saenger, 1984).

To simulate the extension of the stretched DNA chain, we fixed one of its ends at the original point in the three-dimensional Cartesian system and applied a force \( f \) directed along the \( z \) axis to the second end, which corresponds to the fifth term of Eq. 6.

### Calculation of link number

The number of times the two strands of DNA double helix are interwound, i.e., the link number \( L_k \), is a topologic invariant quantity for closed DNA molecule and also for linear DNA polymer in the case in which the orientations of two extremities of the linear polymer are fixed and any part of the polymer is forbidden to go around the extremities of the polymer. An unstrained B-DNA molecule has one right-handed twist per 3.4 nm along its length, i.e., \( L_k = L_0/3.4 \). Under some twist stress, the link number of a DNA polymer may be different from its torsionally relaxed value. In all cases, when \( \Delta L_k = L_k - L_0 \neq 0 \), the DNA polymer is called supercoiled (Vologodskii and Cozzarelli, 1994). The relative difference in link number

\[
s = \frac{L_k - L_0}{L_0}
\]

signifies the degree of supercoiling that is independent of the length of DNA polymer. The native DNA of organisms living at physiological environment are found always slightly underwound, and its supercoiling degree is between \(-0.03\) and \(-0.09 \) (Bauer, 1978; Vologodskii and Cozzarelli, 1994), which is believed to be significantly relevant in some fundamental biological processes (Wu et al., 1988; Morse and Simpson, 1988).

In addition to directly counting the number of times the two strands are interwound, the link number of the closed DNA circle can be conveniently calculated by White’s theorem (White 1969)

\[
L_k = T_W + W_r.
\]

The twist, \( T_W \), is the number of times basepairs twist around the central axis and does not depend upon the configuration of molecule axis. The writhe, \( W_r \), of the molecule is a simple function of only the molecule axis vector \( \mathbf{r}(s) \) (White, 1969; Fuller, 1971)

\[
W_r = \frac{1}{4\pi} \int \int ds \, ds' \frac{\partial_s \mathbf{r}(s) \times \partial_{s'} \mathbf{r}(s') \cdot [\mathbf{r}(s) - \mathbf{r}(s')] }{[\mathbf{r}(s) - \mathbf{r}(s')]^3}.
\]

\( Wr \) is scale invariant and dimensionless and changes sign under reflection or inversion of \( r \), reflecting the cross product in the formula above. Therefore, \( Wr = 0 \) if \( r(s) \) is planar or otherwise reflection symmetric.

To control and measure experimentally the supercoiling degree of linear DNA polymer, Strick et al. (1996, 1998) attached one end of a DNA molecule to a glass cover slip by DIG-anti-DIG links and the other end to a paramagnetic bead by biotin-streptavidin links. Bearing in mind that the diameter of the magnetic bead (~4.5 μm) is far beyond that of a polymer, the anchoring points can be considered as on impenetrable walls, and ~16-μm-long DNA (Strick et al., 1996), in fact, is prohibited from passing around the ends of the polymer. A magnetic field pointing in the plane of the microscope slide was applied to fix the orientation of the bead. Therefore, by rotating the magnets and counting the time of turns, the link number \( L_k \) of the linear DNA molecule can be controlled and measured experimentally.

In Monte Carlo calculations, we restrict the DNA chain by two impenetrable parallel walls crossing the chain ends, which is to simulate the above-mentioned experimental equipment of the magnet bead and the microscope slide (see also the treatment in Vologodskii and Marko, 1997). The walls are always parallel to the \( xy \) plane in our Cartesian coordinate system and thus perpendicular to the direction of the force applied to the chain ends.

One way to calculate the link number \( L_k \) of DNA molecule in our Monte Carlo simulation is to use White’s formula Eq. 10. However, the writhe \( Wr \) is defined only for closed chain. To solve the problem, we add three long flat ribbons to the two ends of the DNA chain in each conformation during the simulation procedure. The axes of these ribbons are kept in the same planar and consist of a closed circle together with the linear DNA chain. Because there is not any twist in the added three flat ribbons, it is not difficult to verify from Fig. 3 that the number of times two strands interwind \( L_k \) in Fig. 3 a is equal to the link number of new closed polymer \( L_k \) in Fig. 3 d. Therefore, we only calculate \( L_k \) of the closed chain in our simulations according to Eqs. 10 and 11.

Quite similar to the model by Tan and Harvey (1989) in which the twist of each basepair of DNA chain is explicitly specified, the folding angle of backbones in each segment has been given in our model. So the twist can be directly calculated by

\[
T_W = \frac{1}{2\pi R} \sum_{i=1}^{N} \Delta \theta_i \tan \theta_i.
\]

The writhe \( Wr \) of the new DNA circle can be calculated according to Eq. 11.

### Simulation procedure

For any given force, equilibrium sets of conformations of DNA chain are constructed using the Metropolis MC procedure (Metropolis et al., 1953). Three kinds of movements have been considered in our simulations (see Fig. 4).

In the first type of movement, a random chosen segment is undertwisted or overtwisted by an angle \( \lambda_i \). In other words, the folding angle \( \theta_i \) of the chosen segment is modified into a new value \( \theta'_i = \theta_i + \lambda_i \). When \( \theta'_i \) is beyond the setting interval \([-\theta_{min}, \theta_{max}]\) from one side, it will re-enter the interval from the opposite side according to the periodicity assumption. Although the geometric limit of folding angle of DNA backbone is \( \theta_{min} = \pi/2 \), we set \( \theta_{max} = 85^\circ \) here to avoid the possible divergence in numerical calculation of potential of Eq. 1. It should be mentioned that this movement modifies not only the folding angle of the chosen segment but also the coordinates of all the behind vertices \( \mathbf{r}_j \), \( j = i+1, \cdots, N \) along the length, because, when the folding angle \( \theta_i \) is changed, we have also changed the length of the segment \( \Delta s_j \) according to Eq. 2. So we should translate all those segments behind this one to make the chain match up (Fig. 4 a).
In the second type of movement, an interval subchain containing an arbitrary amount of segments will be rotated by an angle of $\lambda_i$ around the straight line connecting the vertices bounding the subchain (Fig. 4 b). The third type of movement involves a rotation of the subchain between chosen vertices and the free end by an angle of $\lambda_j$ around an axis with arbitrary orientation (Fig. 4 c). All three types of movements satisfy the basic requirement of the Metropolis procedure of microscopic reversibility, i.e., the probability of trial conformation $B$ when current conformation is $A$ must be equal to the probability of trial conformation $A$ when current conformation is $B$.

All three types of movements change the configurations of the DNA chain. However, from the viewpoint of energy, their functions are quite different. Although the first type of movement is concerned mainly with modifying twist and stacking energy, the second one changes only the bending energy and the third modifies both bending energy and extension of the DNA chain. Each of them is performed with a probability of $\frac{1}{3}$. The value of $\lambda_1$, $\lambda_2$, $\lambda_3$ are uniformly distributed over interval $(-\lambda_0^c, \lambda_0^c)$, $(-\lambda_0^b, \lambda_0^b)$, and $(-\lambda_0^t, \lambda_0^t)$, respectively, and $\lambda_0^b$, $\lambda_0^t$, and $\lambda_0^c$ are chosen to guarantee that about half of the trial moves of each type are accepted.

The starting conformation of DNA chain is unknotted. But the configurations after numerous steps of movements may become knotted, which violates the topological invariance of chain and is incorporeal. Especially, both ends of molecule are anchored in the experiment and knots never occur. To avoid knotted configuration, we should check the knot status for both ends of molecule are anchored in the experiment and knots never occur. To avoid knotted configuration, we should check the knot status for each trial conformation. The most effective way to clarify the knot category is to deconstruct the DNA chain into the subchains, i.e., $L_k_i = L_k_c = Tw + Wr$. So, by adding three straight ribbons, the link number of linear double-helix DNA can be obtained by calculating the link number of the new DNA circle, i.e., $L_k = L_k_c = Tw + Wr$.

RESULT OF MONTE CARLO SIMULATION

To obtain an equilibrium ensemble of DNA evolution, $10^7$ elementary displacements are produced for each chosen applied force $f$. The relative extension $x$ and supercoiling degree $\sigma$ of each accepted conformation of the DNA chain are calculated. When the trial movement is rejected, the current conformation is counted up twice (see Metropolis et al., 1953).

To see the dependence of mechanics property of DNA upon supercoiling degree, the whole sample is partitioned into 15 subsamples according to the value of the supercoiling degree $\sigma$. For each subsample, we calculate the averaged extension,

$$\langle x_j \rangle = \frac{1}{N} \sum_{i=1}^{N} \frac{z_i}{L_i}, \quad j = 1, \ldots, 15,$$

and the averaged torsion

$$\langle \sigma \rangle = \frac{1}{N} \sum_{i=1}^{N} \sigma_i, \quad j = 1, \ldots, 15,$$
where $N_j$ is the number of movements supercoiling that belong to $j$th subsample.

We display the force versus relation extension for all positive and negative supercoiling in Fig. 5, a and c, respectively. As a comparison, the experimental data (Strick et al., 1998) are shown in Fig. 5, b and d. In Fig. 6 is shown the averaged extension as a function of supercoiling degree for three typical applied forces. At low force, the extension in our Monte Carlo simulation saturates at a value greater than zero because of the impenetrable walls, which strait the vertical coordinate of the free end always higher than that of any other points of the DNA chain. The same effect of the impenetrable walls was found in earlier works (see Fig. 9 of Vologodskii and Marko, 1997). For conciseness, we did not show the relative extension points of less than 0.15 in Figs. 5 and 6.

Despite quantitative difference between Monte Carlo results and experimental data, the qualitative coincidence is striking. Especially, three evident regimes exist in both experimental data and our Monte Carlo simulations:

1. At a low force, the elastic behavior of DNA is symmetrical under positive or negative supercoiling. This is understandable, because the DNA torsion is the cooperative result of hydrogen-bond constrained bending of DNA backbones and the base-stacking interaction in our model. At very low force, the contribution from applied force and the thermodynamic fluctuation perturbate the folding angle $\theta$ of the basepair to derive very little from the equilibrium position $\theta_0$. Therefore, the DNA elasticity is achiral at this region (see Introduction). For a fixed applied force, the increasing torsion stress tends to produce a plectonemic state, which shortens the distance between the two ends; and therefore, the relative extension of linear DNA polymer. These features can also be understood by the traditional approaches with harmonic twist and bending elasticity (Vologodskii and Marko, 1997; Bouchiat and Mezard, 1998).

2. At intermediate force, the folding angle of basepairs is pulled slightly further away from equilibrium value $\theta_0$ where van der Waals potential is not symmetric around $\theta_0$. So the chiral nature of elasticity of the DNA molecule appears. In the negative supercoiling region, i.e., $\theta < \theta_0$, the contribution of applied force dominates that of potential force because of the low plateaus of $U(\theta)$. So, the extension is insensitive to negative supercoiling degree. In contrast, the positive supercoiling still tends to contract the molecule.

3. At higher force, the contribution of the applied force to the energy dominates that of van der Waals potential in both over- and underwound DNA. The extension of DNA accesses to its B-form length. Therefore, the plectonemic DNA is fully converted to extended DNA, the writhe is essentially entirely converted to twist and the force-extension behavior reverts to that of untwisted ($\sigma = 0$) DNA as expected from a torsionless worm-like chain model (Smith et al., 1992; Marko and Siggia, 1995; Zhou et al., 1999). Because of the effect of the impenetrable wall, however, the extension of the DNA molecule in our calculation is slightly higher than in experimental data.
In conclusion, the elasticity of supercoiled double-stranded DNA is investigated by Monte Carlo simulations. Instead of introducing an extra twist energy term, twist and supercoiling are led into as a natural result of cooperative interplay between base-stacking interaction and sugar-phosphate backbones bending constrained by permanent hydrogen bonds. Without any adjustable parameter, the theoretic results on the correlations among DNA extension, supercoiling degree, and applied force agree qualitatively with recent experimental data by Strick et al. (1996, 1998).

It should be mentioned that there is an upper limit of supercoiling degree for extended DNA in the current model, i.e., $\sigma_{\text{max}} = 0.14$, which corresponds to $\theta = 90^\circ$ of folding angle. In recent experiments, Allemand et al. (1998) twisted the plasmid up to the range of $-5 < \sigma < 3$. They found that, at this “unrealistically high” supercoiling, the curves of force versus extension for different $\sigma$ split again at higher stretch force (>3 pN). As argued by Allemand et al. (1998), in the extremely under- and overwound torsion stress, two new DNA forms, denatured-DNA and P-DNA with exposed bases, will appear. In fact, if the deviation of the angle that specifies DNA twist from its equilibrium value exceeds some threshold, the corresponding torsional stress causes local distraction of the regular double-helix structure (Vo-logodskii and Cozzarelli, 1994). So, the emergence of these two striking forms is essentially associated with the broken processes of some basepairs under super-highly torsional stress. In this case, the permanent hydrogen constraint will be violated, and the configuration of base-stacking interactions vary considerably. We hope, with incorporation of these effects at high supercoiling degree, our model should reproduce the novel elastic behavior of DNA. This part of work is in progress.

**APPENDIX: KUHN STATISTICAL LENGTH OF DISCRETE CHAIN**

Let us consider a discrete chain of $N$ segments with each of length $l_0$, the end-to-end vector of which is written as

$$\mathbf{R} = l_0 \sum_{i=1}^{N} \mathbf{t}_i$$

(A1)

where $\mathbf{t}_i = (\mathbf{R}_i - \mathbf{R}_{i-1})/||\mathbf{R}_i - \mathbf{R}_{i-1}||$.

For chains with bending stiffness, e.g., the DNA model described in Eq. 6, $(\mathbf{t}_{i+k} \cdot \mathbf{t}_i)$ does not vanish for $k \neq 0$. $\mathbf{t}_{i+k}$ can be expressed relative to $\mathbf{t}_i + \mathbf{t}_{i+k}$.
where $\gamma_{i+k-1}$ is the bending angle between $(i + k - 1)$th and $(i + k)$th segments as defined in Eq. 6, and $n_{i+k-1}$ is the unit vector coplanar with $t_{i+k}$ and $t_{i+k-1}$ but perpendicular to the latter. If the average of $t_{i+k}$ is taken with the rest of the chain (i.e., $t_i, t_{i+1}, \ldots, t_{i+k-1}$) fixed, one obtains

\[
\langle t_{i+k} \cdot t_i \rangle = \langle \cos \gamma \rangle \langle t_{i+k-1} \cdot t_i \rangle, \tag{A4}
\]

where $\langle \cos \gamma \rangle$ is not specific to segments and is given by Eq. 8. This recursion equation, with the initial condition $t^t = 1$, is solved by

\[
\langle t_{i+k} \cdot t_i \rangle = \langle \cos \gamma \rangle^k. \tag{A5}
\]

Thus, for large $N$, $\langle R^2 \rangle$ is given by

\[
\langle R^2 \rangle = l_0^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \langle t_i \cdot t_j \rangle = l_0^2 \left( N + 2 \sum_{i=1}^{N-1} \sum_{k=1}^{N-i} \langle t_i \cdot t_{i+k} \rangle \right) = Nl_0^2 \left( 1 + \langle \cos \gamma \rangle \right) / \left( 1 - \langle \cos \gamma \rangle \right). \tag{A6}
\]

Therefore, the Kuhn statistical length of the discrete chain can be written as

\[
b \equiv \frac{\langle R^2 \rangle}{R_{\text{max}}} = l_0 + \frac{\langle \cos \gamma \rangle}{1 - \langle \cos \gamma \rangle}.
\]

where $R_{\text{max}}$ is the maximum length of the end-to-end vector.

Parts of the computer calculations of this work were performed in the Computer Cluster of Institut für Theoretische Physik, Berlin and the State Key Laboratory of Scientific and Engineering Computing, Beijing, to which our thanks are due.

One of the authors (Z.Y.) would like to thank U.H.E. Hansmann, B.-L. Hao, L.-S. Liu and W.-M. Zheng for discussions and help.

REFERENCES


