# Monte Carlo Simulations of Supercoiling Free Energies for Unknotted and Trefoil Knotted DNAs

John A. Gebe,* Stuart A. Allison,‡ James B. Clendenning,* and J. Michael Schurr*
*Department of Chemistry, University of Washington, Seattle, Washington 98195; and ‡Department of Chemistry, Georgia State University, Atlanta, Georgia 30303-3083 USA

ABSTRACT A new Monte Carlo (MC) algorithm is proposed for simulating inextensible circular chains with finite twisting and bending rigidity. This new algorithm samples the relevant Riemann volume elements in a uniform manner, when the constraining potential vanishes. Simulations are performed for filaments comprising 170 subunits, each containing ~28 bp, which corresponds to a DNA length of 4770 bp. The bending rigidity is chosen to yield a persistence length, $P$ = 500 Å, and the intersubunit potential is taken to be a hard-cylinder potential with diameter $d$ = 50 Å. This value of $d$ yields the same second virial coefficient as the electrostatic potential obtained by numerical solution of the Poisson-Boltzmann equation for 150 mM salt. Simulations are performed for unknotted circles and also for trefoil knotted circles using two different values of the torsional rigidity, $C$ = (2.0 and 3.0) × $10^{-19}$ dyne cm². In the case of unknotted circles, the simulated supercoiling free energy varies practically quadratically with linking difference $\Delta\ell$. The simulated twist energy parameter $E_T$, its slope $dE_T/dT$, and the mean reduced writhe $\langle w\rangle/\Delta\ell$ for $C = 3 \times 10^{-19}$ dyne cm² all agree well with recent simulations for unknotted circles using the polygon-folding algorithm with identical $P$, $d$, and $C$. The simulated $E_T$ vs. $\Delta\ell$ data for $C = 2.0 \times 10^{-19}$ dyne cm² agree rather well with recent experimental data for p30δ DNA (4752 bp), for which the torsional rigidity, $C = 2.07 \times 10^{-19}$ dyne cm², was independently measured. The experimental data for p30δ are enormously more likely to have arisen from $C = 2.0 \times 10^{-19}$ than from $C = 3.0 \times 10^{-19}$ dyne cm². Serious problems with the reported experimental assessments of $E_T$ for pBR322 and their comparison with simulated data are noted. In the case of a trefoil knotted DNA, the simulated value, $(E_T)_{tre}$, exceeds that of the unknotted DNA, $(E_T)_{unk}$, by ≅ 1.40-fold at $|\Delta\ell| = 1.0$, but declines to a plateau about 1.09-fold larger than $(E_T)_{unk}$ when $|\Delta\ell| \geq 15$. Although the predicted ratio, $(E_T)_{tre}/(E_T)_{unk} \cong 1.40$, agrees fairly well with recent experimental measurements on a 5600-bp DNA, the individual measured $E_T$ values, like some of those reported for pBR322, are so large that they cannot be simulated using $P$ = 500 Å, $d$ = 50 Å, and any previous experimental estimate of C.

## INTRODUCTION

Monte Carlo simulations have been employed extensively to investigate both structural and thermodynamic properties of supercoiled DNAs (Vologodskii and Cozzarelli, 1994). The variation of the supercoiling free energy ($\Delta G_{sc}(\Delta\ell)$) with linking difference ($\Delta\ell$) is a topical issue (Klenin et al., 1991), for which a rigorous comparison between theory and experiment is presently lacking. Only recently were measurements of $\Delta G_{sc}(\Delta\ell)$ performed that extend over the full range of $\Delta\ell$ values from 0 to native for a single DNA, and which apply at a salt concentration where the twisting and bending rigidities and intersubunit repulsive interactions are adequately characterized (Clendenning et al., 1994). The dependence of the twist energy parameter ($E_T$), which governs the free energy of supercoiling, on the torsional rigidity ($C$) of the filament, is another topic of considerable interest. Here, too, simulated and experimental data have not been rigorously compared in any instance where $E_T$ and $C$ were independently measured for the same DNA. Furthermore, any comparisons that have been made are obscured by the considerable spread in reported $E_T$ values for different DNAs in 100–200 mM NaCl at 37°C, which range from 950 to 1320 (Pulleyblank et al., 1975; Hinton and Bode, 1975; Wu et al., 1988; Duguet, 1993; Clendenning et al., 1994; Naimushin et al., 1994). The recent measurements of $\Delta G_{sc}(\Delta\ell)$ for p30δ DNA (4752 bp) together with the independently determined value of $C$ for that same DNA (Clendenning et al., 1994) provide an experimental basis for the first rigorous test of simulated supercoiling free energies. One objective of the present study is to perform the requisite simulations for two different values of $C$ and to compare the simulation results with those experimental data.

Experimental determinations of $E_T$ were recently reported for both unknotted and trefoil knotted forms of a 5600-bp DNA at small values of $\Delta\ell$ (Shaw and Wang, 1993). A second objective of this investigation is to simulate $E_T$ for the trefoil knotted, as well as unknotted, forms of the same DNA as a function of $\Delta\ell$ and to compare both the individual values and their ratio with the experimental results.

The Monte Carlo algorithms previously employed to simulate supercoiled DNAs are not very well suited for treating anisotropic bending potentials and certain kinds of permanent bends. In addition, it is not known whether they satisfy the uniform sampling requirement discussed in the next section. A third objective of the present study is to formulate

*Received for publication 28 June 1994 and in final form 18 November 1994.*
Address reprint requests to Dr. J. Michael Schurr, Department of Chemistry, University of Washington, BG-10, Seattle, WA 98195. Tel.: 206-543-6681; Fax: 206-685-8665; E-mail: schurr@chem.washington.edu.

0006-3495/95/02/619/15 $2.00

a new Monte Carlo algorithm that satisfies the uniform sampling requirement and that can be (and already has been) extended to treat anisotropic bending potentials.

## Uniform sampling requirement

The molecular system is defined by a set of coordinates $\{q_i, i = 1, N\}$, which are not necessarily Cartesian. Because the computer truncates all numbers, the accessible $q_i$ values define a regular lattice, which we call the bin lattice, with a bin spacing $\delta q_i$ and associated bin volume, $\delta V = \prod_{i=1}^{N} \delta q_i$, which is the same everywhere. The configuration of the molecular system can be specified no more precisely than by stating the particular bin volume that it occupies. The actual Riemann volume element that is associated with a particular bin volume element is denoted by $\delta\Gamma$. When the $q_i$ are all Cartesian coordinates, then $\delta\Gamma = \delta V$, and any set of random moves on the $q_i$ that 1) accesses every bin lattice point (or bin volume) and 2) satisfies microscopic reversibility will sample the bin volumes $\delta V$, and hence the Riemann volume elements $\delta\Gamma$, uniformly on the average in the absence of potentials and constraints. However, for non-Cartesian coordinates, generally $\delta\Gamma \neq \delta V$, and moreover the $\delta\Gamma$ values associated with different bin volumes vary with one or more of the $q_i$. For example, consider a simple solid body, whose configuration is defined by the Euler angles $(\alpha\beta\gamma)$ that orient it in the laboratory frame. The bin volume is $\delta V = \delta\alpha\, \delta\beta\, \delta\gamma$, but the actual Riemann (or Euler) volume element associated with that bin volume is $\delta\Gamma = \delta\alpha \sin\beta\, \delta\beta\, \delta\gamma = \sin\beta\, \delta V$, which varies with $\sin\beta$. In this case, any protocol that accesses every bin volume in a microscopically reversible manner will uniformly sample the $\delta V$, but that precludes uniform sampling of the $\delta\Gamma$, and is therefore incorrect. The origins of the uniform sampling requirement are described briefly below.

The classical configuration integral for the molecular system is

$$Z = \int \exp\left(-\frac{\mathcal{U}(\mathbf{q})}{kT}\right) d\Gamma \qquad (1)$$

where $\mathbf{q}$ simply denotes the list of coordinates $\{q_i, i = 1, N\}$ that specifies the molecular configuration, and $\mathcal{U}(\mathbf{q})$ is the potential energy. To generate a canonical distribution of configurations, the MC sampling procedure must generate numbers of configurations in the different bin volumes (labeled 1 and 2) such that

$$\frac{n_2}{n_1} = \exp\left(-\frac{\mathcal{U}(\mathbf{q}_2) - \mathcal{U}(\mathbf{q}_1)}{kT}\right)\frac{\delta\Gamma_2}{\delta\Gamma_1} \qquad (2)$$

In particular, when the potential energy is set to 0 and any constraints are removed, the MC sampling protocol must yield $n_2/n_1 = \delta\Gamma_2/\delta\Gamma_1$ or $n_2/\delta\Gamma_2 = n_1/\delta\Gamma_1$. That is, the MC sampling of the Riemann volume elements associated with different bin volumes must be uniform in the sense of uniform "concentration" in those Riemann volume elements. Hence for non-Cartesian coordinates, when $\delta V \neq \delta\Gamma$, the sampling of the bin volumes must perforce be non-uniform.

Although considerable attention has been devoted to the question of ergodicity of various sampling schemes in which the random moves take place on Cartesian lattices (Madras and Sokal, 1987, 1988), we are not aware of any previous discussion of the uniform sampling problem that arises when the random moves are made in polar or Euler angle spaces or in other curvilinear coordinate spaces. In such cases, the sampling protocol must not only be ergodic in the sense of accessing every finite Riemann volume element, but it must also sample those uniformly in the sense of sample concentration, which is a more stringent requirement. In this work we propose an MC sampling protocol that achieves everywhere locally uniform sampling of the relevant Riemann volume elements. This algorithm also is presumably ergodic, as described below, although a rigorous proof of that is lacking.

## Alternative protocols

Previous MC determinations of supercoiling free energies proceeded by simulating the distribution of writhe in nicked circles, which (in the case of isotropic bending potentials) could be modeled simply as circular chains of bond vectors connected by bending springs without regard for twisting deformations. The supercoiling free energy of a particular closed circular topoisomer is obtained from its configuration integral, the relevant part of which can be expressed simply as a convolution of the distribution of writhe ($w$ (turns)) in the nicked circle with the (analytical) gaussian distribution of net twist ($t$ (turns)) in the nicked circle, subject to the topological constraint

$$\ell = t + w, \qquad (3)$$

wherein $\ell$(turns) is the topologically invariant linking number of the topoisomer in question.

Presently, two main methods are employed to simulate the nicked circles. One procedure is to simulate an enormous number of linear half-chains, and pair those that exhibit nearly matching end-to-end distances and suitable end-vector orientations by minor adjustments to complete the "circle" (Chen, 1981; Levene and Crothers, 1986; Taylor and Hagerman, 1990). Provided that the small random rotations of each individual bond vector sample either the solid angle volume elements $\delta(\cos\beta)\,\delta\alpha$ or the volume elements $\delta R_x\, \delta R_y$ of small rotations around (subunit) body-fixed transverse axes in a uniform manner, and that the required adjustments are sufficiently small, this protocol will be satisfactory. However, it is too inefficient for most applications. The most popular procedure currently is the polygon-folding algorithm adopted by Frank-Kamenetskii et al. (1985) and LeBret (1986), and subsequently employed by others with various embellishments (Shimada and Yamakawa, 1988; Klenin et al., 1989, 1991; Vologodskii et al., 1992; Vologodskii and Cozzarelli, 1994). The polygon represents the bond vectors that constitute the circular chain, and which consequently must sum to 0. In this algorithm, two vertices are selected at random and all of the bond vectors comprising the "arc" of the polygon on one side of the two vertices are rotated around the chord connecting the two vertices by a random amount

subject to the Metropolis et al. (1953) criterion. Certainly, on any given trial, this algorithm does not sample the Riemann volume elements in a locally uniform manner. This is clearly seen by considering a planar polygon. The intervertex line necessarily lies in the plane, so any bond vector rotations on the next move can occur only about an in-plane axis, thus deforming the polygon out of the plane, while rotations about an axis perpendicular to the plane are forbidden on any such trial. Thus, certain Riemann volume elements are not sampled at all in such a move. In contrast, a real circular DNA can spontaneously deform from a circle to an ellipse without moving out of plane, and a regular planar polygon should likewise be able to deform to an elliptical envelope without going out of plane. Although the polygon-folding algorithm does not sample the Riemann volumes in a *locally* uniform manner on any given trial, and does not "move" like a real DNA, it can presumably reach any configuration, more precisely any bin volume in the solid angle space of the bond vectors. Each basic move of the polygon-folding algorithm is a nonlocal pivot move in the sense that an arbitrarily large number of subunits (up to $N/2$, where $N$ is the total number of subunits) may move (i.e., swing) as a consequence of the rotation. For linear self-avoiding chains on a Cartesian lattice, it was proved rigorously that a class of MC algorithms consisting of particular nonlocal pivot moves is ergodic (Madras and Sokal, 1988). Hopefully, a corresponding proof will appear soon for self-avoiding circular chains. If the polygon-folding algorithm is ergodic, then it is still possible, but not necessarily certain, that in some nonlocal manner the Riemann volume elements are sampled uniformly by this procedure. However, so far as we know, this has not been proved, nor do we presently see how to do so. This raises the disturbing possibility that the polygon-folding algorithm does not sample the Riemann volume elements in a uniform manner, especially when the equilibrium distribution of bond vector orientations is anisotropic, as it typically is for supercoiled DNAs. To improve the rate of equilibration of different branched structures, the polygon-folding algorithm was augmented to incorporate moves in which different three- and four-bond arcs of the polygon were altered slightly in shape and interchanged (Vologodskii et al., 1992). Again, no proof was given that such a procedure samples the relevant Riemann volume elements uniformly.

Simulations of the distribution of writhe in nicked circles yield supercoiling free energies, but do not provide a means to generate canonical sets of configurations for any particular topoisomer. The latter requires simulating the closed circular topoisomer per se (Vologodskii et al., 1992). Although it is straightforward to determine supercoiling free energies also from such simulations, as described below, this apparently has not been done previously.

## THEORY AND METHODS

### Model and basic theory

The closed circular DNA is modeled by a string of $N$ contiguous subunits, each of which has length $b$ and is connected to its neighbors on either side by Hookean bending and twisting springs. The subunits are labeled sequentially by the index $j$, $j = 1, \cdots, N$. Embedded in the $j$th subunit is a local coordinate frame, $x_j$, $y_j$, $z_j$, which is chosen so that the $z_j$ axis is directed along the bond vector ($\mathbf{b}_j$) from the $j$th to the $(j + 1)$th subunit. The bond vector associated with the $j$th subunit is given by $\mathbf{S}_j = (00b)$ in the $j$th frame. The Euler rotation that carries a coordinate frame from coincidence with the lab frame $(x', y', z')$ to coincidence with the $j$th frame is denoted by $\Phi_j = (\alpha_j\ \beta_j\ \gamma_j)$, where $0 \le \alpha_j \le 2\pi$, $0 \le \beta_j \le \pi$, $0 \le \gamma_j \le 2\pi$. The bond vector associated with the $j$th subunit is then given in the lab frame by (Schutle, 1984)

$$\mathbf{b}_j = \begin{pmatrix} x'_j \\ y'_j \\ z'_j \end{pmatrix} = \mathbf{R}^{-1}(\Phi_j)\mathbf{S}_j \tag{4}$$

where $\mathbf{R}^{-1}(\Phi_j)$ is the rotation matrix that transforms a vector in the $j$th frame to a vector in the lab frame, and is given by

$$\mathbf{R}^{-1}(\Phi) = \begin{bmatrix} \cos(\alpha)\cos(\beta)\cos(\gamma)-\sin(\alpha)\sin(\gamma) & -\cos(\alpha)\cos(\beta)\sin(\gamma)-\sin(\alpha)\cos(\gamma) & \cos(\alpha)\sin(\beta) \\ \sin(\alpha)\cos(\beta)\cos(\gamma)+\cos(\alpha)\sin(\gamma) & -\sin(\alpha)\cos(\beta)\sin(\gamma)+\cos(\alpha)\cos(\gamma) & \sin(\alpha)\sin(\beta) \\ -\sin(\beta)\cos(\gamma) & \sin(\beta)\sin(\gamma) & \cos(\beta) \end{bmatrix} \tag{5}$$

The position of the first subunit in the chain is taken to be (000) in the laboratory frame. The position of the head end of the $j$th vector in the chain ($\mathbf{r}_j$), which actually corresponds to the position of the $(j + 1)$th subunit ($\mathbf{R}_{j+1}$), is given in the lab frame by

$$\mathbf{r}_j = \sum_{i=1}^{j} \mathbf{b}_i = \mathbf{R}_{j+1} \tag{6}$$

For a closed circular filament, the position of the first subunit is then $\mathbf{R}_1 = \mathbf{r}_N = (000)$, since the bond vectors sum to 0.

The total potential energy of a given configuration is assumed to be given by

$$U_{tot} = U_b + U_t + U_I, \tag{7}$$

wherein $U_b$ and $U_t$ are the bending and twisting energies, respectively, and $U_I$ is a term that accounts for all intersubunit interactions that are not mediated by the twisting and bending springs. The twisting and bending energies are expressed most simply in terms of the Euler angles $\Phi_{j,j+1} = (\alpha_{j,j+1}\ \beta_{j,j+1}\ \gamma_{j,j+1})$ that orient the frame of the $(j + 1)$th subunit in the frame of the $j$th subunit. Specifically,

$$U_b = \sum_{j=1}^{N} \left(\frac{\kappa_\beta}{2}\right)\beta_{j,j+1}^2 = \sum_{j=1}^{N} \left(\frac{\kappa_\beta}{2}\right)[\cos^{-1}(\hat{\mathbf{b}}_j \cdot \hat{\mathbf{b}}_{j+1})]^2, \tag{8}$$

where $\kappa_\beta$ is the torque constant for a bending spring and $\hat{\mathbf{b}}_j \equiv \mathbf{b}_j/|\mathbf{b}_j|$, and also

$$U_t = \left(\frac{\alpha}{2}\right)\sum_{j=1}^{N}(\phi_{j,j+1} - \phi_0)^2 \tag{9}$$

where $\alpha$ (dyne cm) is the torque constant for a twisting spring, $\phi_{j,j+1} = \alpha_{j,j+1} + \gamma_{j,j+1}$ is the net twist of the Euler rotation from the $j$th to the $(j+1)$th frame, and $\phi_0$ is the intrinsic equilibrium twist from the $j$th to the $(j+1)$th subunit. The torsional rigidity of the filament is $C = b \cdot \alpha$ (dyne cm$^2$). When the cumulative net twist of the circular DNA is $t = (1/2\pi)\sum_{j=1}^{N}\phi_{j,j+1}$, then Eq. 9 can be rewritten as

$$U_t = \left(\frac{\alpha(2\pi)^2}{2N}\right)(t - \ell_0)^2 + \left(\frac{\alpha}{2}\right)\sum_{j=1}^{N}\left(\phi_{j,j+1} - \frac{2\pi t}{N}\right)^2 \tag{10}$$

wherein $\ell_0 = N\phi_0/(2\pi)$ is the intrinsic twist of the circular DNA. The first term in Eq. 10 represents the minimum energy required to introduce a cumulative net twist $t$, which is accomplished when $\phi_{j,j+1} = 2\pi t/N$ is uniform for all of the springs, $j = 1, \cdots, N$. The second term in Eq. 10 represents the increase in energy that results from fluctuations about this uniform net twist. The subunit interaction energy $U_I = U_I(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ is assumed to depend on the (relative) positions of the subunits, but is assumed to be completely independent of the state of twisting of the filament. When the interactions between subunits arise from long range electrostatic forces, as is the case here, it is expected that $U_I$ will indeed be invariant to any azimuthal rotations of the subunits around their local z axes. $U_I$ is taken to be a hard-cylinder potential. That is, $U_I$ is infinite when the shortest distance between the two line segments corresponding to two different bond vectors is less than the cylinder diameter $d$, and is 0 otherwise. The shortest distance between a pair of line segments $\mathbf{r}_{i+1} - \mathbf{r}_i$ and $\mathbf{r}_{j+1} - \mathbf{r}_j$ is calculated using the algorithm in Appendix A.

It is shown in Appendix B that the configuration integral of a circular DNA with linking difference, $\Delta\ell = \ell - \ell_0$, is given for the present model by

$$Z(\Delta\ell) = 2\pi q \int_C \cdots \int_C \prod_{i=1}^{N}\{d\alpha_i\, d\beta_i \sin\beta_i\} \exp\left[-\frac{U_b + U_I + [\alpha(2\pi)^2/2N](\Delta\ell - w)^2}{k_B T}\right], \tag{11}$$

wherein $\Delta\ell = \ell - \ell_0$ is the linking difference, and subscripts $C$ on the integrals indicate that their arguments must be constrained to preserve chain closure and the correct knot topology. The integration variables in Eq. 11 are the polar angles of the bond vectors (subunit z axes) in the laboratory frame. The factor $q$ is given in Eq. B4 (Appendix B), and is independent of $\ell$, $w$, $t$, and knot topology. Because differences in free energy between topoisomers depend only upon ratios of their configuration integrals, the constant factors $2\pi q$ will cancel out and are omitted in the sequel. The writhe $w$ in Eq. 11 is evaluated using the discretized gauss integral (Hao and Olson, 1989)

$$w = \frac{1}{4\pi}\sum_{\substack{i=1\\ i\neq j}}^{N}\sum_{j=1}^{N}\frac{(\mathbf{b}_j \times \mathbf{e}_{ij} \cdot \mathbf{b}_i)}{|\mathbf{r}_j - \mathbf{r}_i|^2} \tag{12}$$

wherein $\mathbf{e}_{ij} = (\mathbf{r}_i - \mathbf{r}_j)/|\mathbf{r}_i - \mathbf{r}_j|$ is a unit vector along $\mathbf{r}_i - \mathbf{r}_j$.

The configuration of the bond vectors, which is specified by $\{\alpha_j, \beta_j; j = 1, \cdots, N\}$ without regard for the $\{\gamma_1, \cdots \gamma_N\}$, is referred to as the *reduced configuration* of the system. The total energy in the exponent in Eq. 11 is referred to as the *reduced configuration energy*, since the energy of fluctuations about the uniform net twist, $t = \ell - w$, is omitted. Fluctuations about the uniform net twist make always the same contribution, $q$, to the configuration integral, regardless of $\ell$, $w$, $t$ and knot topology, as shown in Appendix B. Consequently, they don't need to be simulated for the present potential.

## Simulation protocol

The next objective is to devise a simulation protocol that samples the relevant angular space in a uniform manner, subject to the constraint that the circle remains closed and that the knot topology remains unaltered, and which yields a population weighting in accord with the exponential in the integrand of Eq. 11. This is done for circles with fixed linking difference, $\Delta\ell$, in the following way. *Very small* random rotations, $\delta x_i$, $\delta y_i$, $\delta z_i$, of each subunit around its three body-fixed, or local, axes are performed for every subunit in the chain at the same time. This $3N$-fold rotation constitutes a single "move." Rotations around the individual axes are chosen to be uniformly distributed on the interval $-\epsilon$ to $+\epsilon$, where $\epsilon$ is always $\leq 1.7°$. For such small values of $\delta x_i$, $\delta y_i$, $\delta z_i$, these rotations commute to very high accuracy. A subunit with initial Euler orientation $\Phi_j^\circ = (\alpha_j^\circ\ \beta_j^\circ\ \gamma_j^\circ)$ in the laboratory frame has a new Euler orientation $\Phi_j = (\alpha_j\ \beta_j\ \gamma_j)$ after the move, which is given by

$$\alpha_j = \alpha_j^\circ + \delta\alpha_j = \alpha_j^\circ + (-1/\sin\beta_j^\circ)[\delta x_j \cos\gamma_j^\circ - \delta y_j \sin\gamma_j^\circ] \tag{13a}$$

$$\beta_j = \beta_j^\circ + \delta\beta_j = \beta_j^\circ + \delta x_j \sin\gamma_j^\circ + \delta y_j \cos\gamma_j^\circ \tag{13b}$$

$$\gamma_j = \gamma_j^\circ + \delta\gamma_j = \gamma_j^\circ + \delta z_j - \cos\beta_j^\circ\, \delta\alpha_j \tag{13c}$$

Eqs. 13 a–c are valid strictly in the limit of very small $\delta x_j$, $\delta y_j$, $\delta z_j$. By performing simulations on simple systems with both the same and considerably smaller intervals ($-\epsilon$ to $+\epsilon$) for the random moves, it was established that the present choice of $\epsilon$ is sufficiently small to achieve small-$\epsilon$ limiting behavior. From the set of new Euler orientations $\{\Phi_1, \cdots, \Phi_N\}$, the corresponding new set of lab-frame bond vectors $\{\mathbf{b}_1, \cdots, \mathbf{b}_N\}$ is then calculated from Eqs. 4 and 5, and $\mathbf{r}_N$ is calculated from Eq. 6. If $|\mathbf{r}_N|$, which is supposed to vanish, exceeds $(0.002)b$, which signifies too large a gap between subunits 1 and $N$, a correction algorithm described in Appendix C is applied to bring the ends together. It is essential to perform this correction in an unbiased manner. As indi-

cated in Appendix C, the offset vector $\mathbf{r}_N$ is apportioned among the $N$ bond vectors by subtracting from each a small perpendicular vector proportional to the projection of $\mathbf{r}_N$ onto the plane perpendicular to that bond vector, while at the same time preserving bond vector length. This is accomplished by an iterative procedure that typically converges to the required limit, $|\mathbf{r}_N| < (0.002)b$, in one pass, and only rarely requires as many as three passes. From the corrected set of bond vectors $\{\mathbf{b}_1^c, \cdots, \mathbf{b}_N^c\}$, the vector head positions $\{\mathbf{r}_1, \cdots, \mathbf{r}_N\}$ are calculated from Eq. 6, and the writhe $w$ is calculated from Eq. 12. It is then ascertained whether the bond-vector line segments approach more closely than the hard-cylinder diameter $d$ by using the algorithm in Appendix A. If so, the new configuration is rejected in favor of the old configuration and, if not, then $U_I = 0$. Next, the new reduced configuration energy (with $U_I = 0$),

$$U_{rc} = U_b + (\alpha(2\pi)^2/2N)\,(\Delta\ell - w)^2 \tag{14}$$

is calculated using Eq. 8 for $U_b$. According to the Metropolis criterion (Metropolis et al., 1953), if $U_{rc}^{new} < U_{rc}^{old}$, the new configuration is provisionally kept (see below), but, if $U_{rc}^{new} > U_{rc}^{old}$, the quantity $\exp[-(U_{rc}^{new} - U_{rc}^{old})/k_BT]$ is calculated and compared to a random number between 0 and 1. If the random number is less than the Boltzmann factor, the new configuration is provisionally kept; otherwise it is rejected in favor of the old configuration. Each provisionally kept configuration is further examined for any change in knot topology, which is detected by calculating its Alexander polynomial (Frank-Kamenetskii and Vologodskii, 1981), as outlined in Appendix D. If such a change has occurred, the new configuration is rejected in favor of the old configuration and, if not, it is finally accepted.

Whenever the set of bond vectors of the accepted configuration is corrected according to the protocol in Appendix C, it is necessary to recalculate the pertinent Euler angles $\{\alpha_i, \beta_i, i = 1, \cdots, N\}$ that orient each corrected bond vector in the lab frame. In this work it is not necessary to correct the third Euler angles, $\{\gamma_1, \cdots, \gamma_N\}$, because they do not enter either the integrand of the configuration integral in Eq. 11 or the reduced configuration energy, $U_b + U_I + U_t^{min}$. However, in treating anisotropic bending potentials, or permanent bends, it is essential to correct also the $\{\gamma_1, \cdots, \gamma_N\}$, which is easily done using a procedure to be described elsewhere. In the present study, $\alpha_i$ and $\beta_i$ are determined according to

$$\beta_j = \arccos(\hat{\mathbf{z}} \cdot \hat{\mathbf{b}}_j) \tag{15a}$$

$$\alpha_j = \arccos((\hat{\mathbf{x}} \cdot \hat{\mathbf{b}}_j)/\sin\beta_j) \tag{15b}$$

In fact, $\arccos(p)$ is multiply defined on the interval, $-1 \le p \le 1$, and the computer yields $\alpha$-values only in the interval 0 to $\pi$. A more detailed analysis yields: $\alpha_j = \alpha_C$ when $\hat{\mathbf{y}} \cdot \mathbf{b}_j > 0$, and $\alpha_j = 2\pi - \alpha_C$ when $\hat{\mathbf{y}} \cdot \mathbf{b}_j < 0$, where $\alpha_C = \arccos((\hat{\mathbf{x}} \cdot \hat{\mathbf{b}}_j)/\sin\beta_j)$ is the value in the interval $0 \le \alpha_C \le \pi$ returned by the computer.

A numerical problem encountered in the use of Eqs. 13a or 15b is that $\delta\alpha_j$ or $\alpha_j$ may diverge when $\beta_j$ approaches 0 or $\pi$. This circumstance occurs when $z_j$ is nearly parallel or antiparallel to the lab z axis. We avoid this circumstance in the following way. If $\beta_j$ is less than some angle, $\Theta = 17°$, then the local coordinate frame of the $j$th subunit is rotated by $\pi/2$ around its own $y_j$ axis, in which case the new $x_j''$ and $z_j''$ axes lie along $-z_j$ and $x_j$, respectively. If $\beta$ is greater than $\pi - \Theta$, the local frame of the $j$th subunit is rotated around its $y_j$ axis by $(-)\pi/2$, in which case the new $x_j'''$- and $z_j'''$-axes lie along $z_j$ and $-x_j$, respectively. The actual bond vector lies along $-x_j''$ in the former case and along $x_j'''$ in the latter case. After such a rotation of the local frame, the new Euler rotation $\Phi_j$ must also be calculated. The code must, of course, keep track of which choice of local coordinate frame applies for each subunit, so that the set of bond vectors $\{\mathbf{b}_1, \cdots, \mathbf{b}_N\}$ is calculated correctly.

As shown in Eqs. 13 a–c, any orientation $\Phi_j$ in the neighborhood of $\Phi_j^o$ can be expressed in terms of $\alpha_j^o$, $\beta_j^o$, $\gamma_j^o$, and first powers of $\delta x_j$, $\delta y_j$, and $\delta z_j$. Thus, locally, the total configuration energy can always be expressed in terms of the variables $\delta x_j$, $\delta y_j$, and $\delta z_j$, for which the relevant Riemann volume element is $d\delta x_j\, d\delta y_j\, d\delta z_j$. By design, the proposed simulation protocol samples such Riemann volume elements uniformly in the absence of constraints, and applies the constraints in a completely unbiased manner. Since this uniform sampling holds regardless of $\Phi_j^o$, one may infer that it holds everywhere. In the present case, $U_b$, $U_I$, $w$, and the entire reduced configuration energy $U_{rc}$ actually depend only upon the bond vector orientations $\{\alpha_j, \beta_j, j = 1, \cdots, N\}$. As is apparent from Eqs. 13 a–c, $\delta\alpha_j$ and $\delta\beta_j$ depend locally only upon $\delta x_j$ and $\delta y_j$. Hence, near $\alpha_j^o$, $\beta_j^o$ the relevant Riemann volume element for the reduced configuration is $d\delta x_j\, d\delta y_j$. This is, of course sampled uniformly for all subunits in the absence of constraints, which are subsequently applied in an unbiased manner.

Each move of the present algorithm can be regarded as a superposition of multiple nonlocal pivot moves in which all of the subunit positions in Cartesian space are subject to change. Although a detailed proof is lacking, this algorithm is presumably ergodic in the sense that it accesses every allowed bin volume (or configuration) in the space of solid angles of all the bond vectors. Since any move of the polygon-folding algorithm is also a possible move of the present algorithm, the present protocol is certainly ergodic in the event that polygon-folding is ergodic. Although the converse is not necessarily true, we nevertheless strongly suspect that both algorithms are in fact ergodic. As noted above, for linear self-avoiding chains on a Cartesian lattice, ergodicity was rigorously proved for a particular class of nonlocal pivot moves (Madras and Sokal, 1988). Moreover, we have been unable to imagine any unknotted configuration of a closed self-avoiding chain that could not be unfolded to a planar circle by an appropriate set of moves of either algorithm. However, in contrast to the present protocol, the polygon-folding algorithm does not locally uniformly sample the available Riemann volume elements, and could conceivably fail to achieve uniform sampling in a nonlocal manner, even if it were ergodic. For this and other reasons, it is important to compare the results of the two algorithms wherever possible.

Simulations with different superhelix densities, $\sigma = \Delta\ell/(N(b/3.4 \times 10^{-8})/(10.45))$, were performed by progressively increasing the magnitude of $\Delta\ell$ from 0, and then "equilibrating" the system for 20,000 accepted moves in each case before keeping any new configurations.

During the simulation, the autocorrelation function of the fluctuation in writhe, $w - \langle w \rangle$, is calculated as a function of move number, and its relaxation period (in number of moves) is estimated as the number of moves to decay to $e^{-1.5}$. The durations of all simulations presented here were such that ~15–20 such relaxation periods occurred during the run. This is our operational criterion for an adequate number of trial moves.

All simulations were performed on IBM RISC 6000 computers, either model 350 or model 580. At the higher superhelix densities each successive determination of $E_T$ required two million moves, which consumed about 24 hours of CPU time on an IBM RISC model 350 computer.

## Free energy of supercoiling and twist energy parameters

At constant temperature and pressure the change in free energy to supercoil the DNA is

$$dG_{sc} = \delta w_{sc} \tag{16}$$

where $\delta w_{sc}$ is the non-pressure-volume work to increase the linking difference by $d\Delta\ell$. This work is $d\Delta\ell$ times the opposing torque that resists the increase in linking difference, namely $\langle \partial U_{rc}/\partial\Delta\ell \rangle$, where $U_{rc}$ is the reduced configuration energy, and the angular brackets denote an equilibrium ensemble average. Thus, the free energy change can be expressed as

$$dG_{sc} = \left\langle \frac{\partial U_{rc}}{\partial \Delta\ell} \right\rangle d\Delta\ell = \frac{(2\pi)^2 C}{Nb} (\Delta\ell - \langle w \rangle) d\Delta\ell \tag{17}$$

The free energy change to introduce a linking difference $\Delta\ell$ into the DNA is then calculated by numerically evaluating the integral,

$$\Delta G_{sc}(\Delta\ell) = \left( \frac{(2\pi)^2 C}{Nb} \right) \int_0^{\Delta\ell} (\Delta\ell' - \langle w \rangle)\, d\Delta\ell' \tag{18}$$

The twist energy parameter is defined by

$$E_T = N_{bp} \Delta G_{sc}(\Delta\ell)/(k_B T \Delta\ell^2), \tag{19}$$

where $N_{bp}$ is the number of base pairs in the equivalent DNA. If $\Delta G_{sc}(\Delta\ell)$ varies quadratically with $\Delta\ell$, then $E_T$ will be a constant, independent of $\Delta\ell$.

In the numerical evaluation of $\Delta G_{sc}(\Delta\ell)$ in Eq. 18, it is *assumed* that $\langle w \rangle = 0$ for the unknotted DNA at $\Delta\ell = 0$, but for the trefoil knotted DNA, the corresponding value for the relaxed molecule was simulated and found to be $|\langle w \rangle| = 3.26 \pm 0.09$ at $\Delta\ell = 0$. The latter value is close to the estimate of LeBret (1980), $\langle w \rangle = 3.30$, for a trefoil knot in a filament without excluded volume (hard-cylinder) interactions.

For the unknotted DNAs, the integrand in Eq. 18 was within statistical error proportional to $\Delta\ell'$ at the $\Delta\ell'$ values simulated, which ranged from 0.0 to −25.0 in increments of −2.5. The same proportionality relation was confirmed also for $\Delta\ell' = -1.0$ by an extended simulation. Although $\Delta G_{sc}(-1.0)$ and $E_T$ are not presented for $\Delta\ell' = -1.0$, due to the considerably larger error bars in that case, the integrand of Eq. 18 for $\Delta\ell' = -1.0$ was used in the calculation of $\Delta G_{sc}(\Delta\ell')$ and $E_T$ for all larger $|\Delta\ell'|$ values.

## Statistical errors

The standard deviation of the mean of the writhe was estimated according to $\sigma_{\langle w \rangle} = \sigma/n^{1/2}$, where $\sigma$ is the standard deviation of the writhe and $n$ is the number of writhe relaxation periods (defined above) during the simulation, which should correspond to the number of independent measurements. The standard errors in $\Delta G_{sc}(\Delta\ell)$ and $E_T$ were estimated from $\sigma_{\langle w \rangle}$ using conventional rules for propagation of errors.

## Simulation parameters

All simulations were performed with a bond length $b = 95.4$ Å, which corresponds to ~28 bp. The persistence length is taken as $P = 500$ Å, so there are ~10.5 bond vectors per Kuhn length. It was shown previously that the mean squared writhe $\langle w^2 \rangle$ of nicked circles becomes independent of the number of bond vectors per Kuhn length, when that number exceeds 10 (Vologodskii et al., 1992). All simulations were performed for filaments containing $N = 170$ subunits, which corresponds to 4770 bp. For the present simulations, the hard-cylinder diameter is taken as $d = 50$ Å. This value gives practically the same second virial coefficient for rigid cylinders as is obtained for charged cylinders (with the linear charge density of DNA) in 150 mM NaCl (Stigter, 1977). Hard-cylinder diameters chosen in this way were found to yield satisfactory results for various properties of supercoiled DNA (Vologodskii and Cozzarelli, 1994). For the torsional rigidity, we initially take $C = 2.0 \times 10^{-19}$ dyne cm$^2$, which corresponds to a torsion constant $\alpha = 5.9 \times 10^{-12}$ dyne cm between base pairs. This value is typical of those obtained from time-resolved fluorescence polarization anisotropy (FPA) measurements on plasmids by assuming our current best guess for the dynamic persistence length, $P_d = 1500$ Å (Fujimoto and Schurr, 1990; Schurr et al., 1992). This also matches the value obtained by Taylor and Hagerman (1990) from the variation of cyclization rates of small circles ($N$ = 336–350 bp) with DNA length. Additional simulations were also performed using a larger torsional rigidity, $C = 3.0 \times 10^{-19}$ dyne cm$^2$, which was obtained by analyzing the topoisomer distribution data for three small circular DNAs (Horowitz and Wang, 1984; Shimada and Yamakawa, 1985; Frank-Kamenetskii et al., 1985). The bending constant, $\kappa_\beta = 0.187 \times 10^{-12}$ dyne cm, is chosen so that the persistence length is 500 Å, when calculated according to

$$P = b/(1 - \langle \cos\theta \rangle) \tag{20}$$

and

$$\langle\cos\theta\rangle = \frac{\int_0^\pi d\theta \sin\theta \cos\theta \exp[-\kappa_\beta\theta^2/2k_BT]}{\int_0^\pi d\theta \sin\theta \exp[-\kappa_\beta\theta^2/2k_BT]} \quad (21)$$

The integral in Eq. 21 was evaluated numerically.

The simulation temperature was $T = 298$ K. The slight variation of simulated $E_T$ values with temperature, when $C$, $P$, and $d$ are held fixed, is estimated in the following way. Differentiating Eq. 19 with respect to $T$ at constant $\Delta\ell$ gives (with $R$ in place of $k_B$),

$$\frac{dE_T}{dT} = \left(\frac{N_{bp}}{(\Delta\ell^2)}\right)\left(-\frac{\Delta S_{sc}}{RT} - \frac{\Delta G_{sc}}{RT^2}\right) \quad (22)$$

For the present 4770-bp DNA with $C = 2.0 \times 10^{-19}$ dyne cm$^2$, $P = 500$ Å, $d = 50$ Å, and $\Delta\ell = -25$, we obtain $\Delta G_{sc} = 5.24 \times 10^{-19}$ J/molecule $= 3.16 \times 10^5$ J/mol, and $\Delta\langle U\rangle = 3.20 \times 10^{-19}$ J/molecule $= 1.93 \times 10^5$ J/mol, where $\langle U\rangle$ is the mean potential energy. The enthalpy change upon supercoiling is just $\Delta H_{sc} = \Delta\langle U\rangle$; hence, the entropy change upon supercoiling is $\Delta S_{sc} = (\Delta\langle U\rangle - \Delta G_{sc})/298 = -412$ J/K mol. Substituting these values for $\Delta G_{sc}$ and $\Delta S_{sc}$ into Eq. 22 yields $dE_T/dT = -2.0$ K$^{-1}$. The estimated change in $E_T$ between 298 and 310 K is then $\Delta E_T = -24$, which is only ~2.4% of the simulated $E_T$ value (1010) at 298 K. When $C$ is increased to $3.0 \times 10^{-19}$ dyne cm$^2$ and $\Delta\ell = -22.5$, but the other parameters are the same as above, we obtain $\Delta G_{sc} = 4.79 \times 10^{-19}$ J/molecule $= 2.88 \times 10^5$ J/mol, and $\Delta\langle U\rangle = 2.57 \times 10^{-19}$ J/molecule $= 1.55 \times 10^5$ J/mol, whence $\Delta S_{sc} = -446$ J/K mole. In this case we again find $dE_T/dT = -2.0$ K$^{-1}$, and the estimated change in $E_T$ between 298 and 310 K is $\Delta E_T = -24$, which is only about 2.1% of the simulated $E_T$ value (1150) at 298 K. Thus, in either case, within the statistical errors, the present simulation results apply practically as well at 310 or 293 K as at 298 K.

The range of simulated $\Delta\ell$ values was 0 to −25 for the unknotted circle and 0 to −22.5 for the trefoil knotted circle.

## RESULTS

The supercoiling free energies obtained for unknotted and trefoil knotted circles with C $= 2.0 \times 10^{-19}$ dyne cm$^2$ are plotted vs. $\Delta\ell$ in Fig. 1. This graph indicates that the supercoiling free energy of the trefoil knotted circle is slightly greater than that of the unknotted circle, but provides no quantitative estimate of the relative difference between those two species at small values of the linking difference. A clearer picture emerges from Fig. 2, where $E_T$ is plotted vs. $\Delta\ell$. For the unknotted circle, the twist energy parameter remains practically constant near the mean value $(\bar{E}_T)_{unk} \cong 1024 \pm 26$ over the whole range of $\Delta\ell$ values. This indicates that $\Delta G_{sc}(\Delta\ell)$ varies essentially quadratically with $\Delta\ell$ clear up to $|\Delta\ell| = 25$ for that species. However, for the trefoil knotted circle, the twist energy parameter extrapolates to $(E_T)_{tre} \cong 1440$ at $|\Delta\ell| \cong 1$, and decreases with increasing $|\Delta\ell|$ until it plateaus at $(E_T)_{tre} \cong 1100$ for $|\Delta\ell| \gtrsim 15$. Clearly, when $|\Delta\ell| \leq 15$, $\Delta G_{sc}(\Delta\ell)$ for the trefoil knotted circle does not vary quadratically with $\Delta\ell$. The ratio of twist

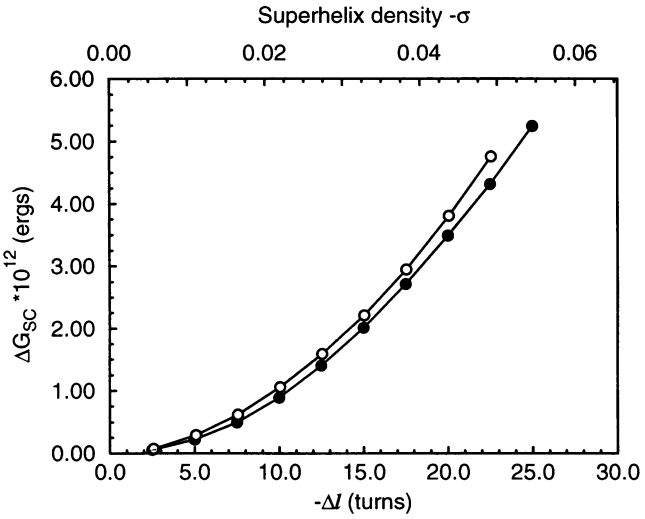


FIGURE 1 Supercoiling free energy ($\Delta G_{SC}$) vs. $-\Delta\ell$ for simulated trefoil knotted (○) and unknotted (●) circular DNA. Simulation parameters are $N = 170$ subunits (4770 bp), persistence length $P = 500$ Å, torsional rigidity $C = 2.0*10^{-19}$ dyne cm$^2$, and $d = 50$ Å. For the trefoil the abscissa actually is not $-\Delta\ell$, but instead is $-(\Delta\ell - \langle w_o\rangle)$, where $\langle w_o\rangle = -3.26$ is the nonzero writhe of the relaxed trefoil.

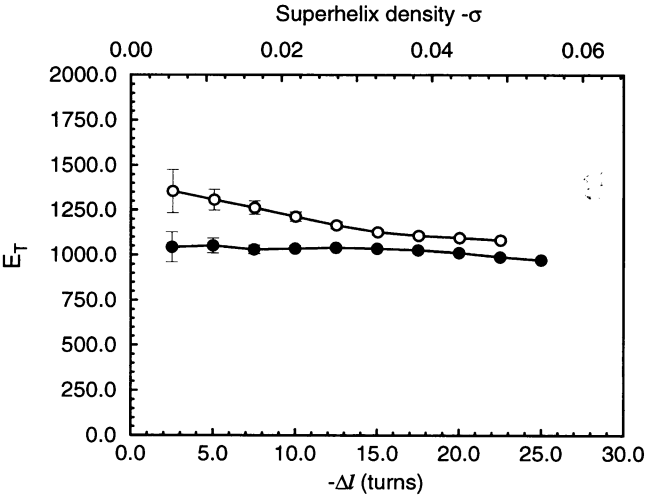


FIGURE 2 $E_T$ vs. $-\Delta\ell$ for simulated trefoil knotted (○) and unknotted (●) circular DNA. Simulation parameters are $N = 170$ subunits (4770 bp), $P = 500$ Å, torsional rigidity $C = 2.0*10^{-19}$ dyne cm$^2$, and $d = 50$ Å. For the trefoil the abscissa actually is not $-\Delta\ell$, but instead is $-(\Delta\ell - \langle w_o\rangle)$, where $\langle w_o\rangle = -3.26$ is the nonzero writhe of the relaxed trefoil.

energy parameters, $(E_T)_{tre}/(E_T)_{unk}$, is plotted vs. $\Delta\ell$ in Fig. 3. In the region of small linking difference ($|\Delta\ell| \leq 2$), which is sampled with greatest accuracy in a typical topoisomer distribution experiment, one finds $(E_T)_{tre}/(E_T)_{unk} \cong 1.40$. This ratio declines with increasing $|\Delta\ell|$ and plateaus at ~1.09 for $|\Delta\ell| \gtrsim 15$. As is evident from Fig. 2, the decline in this ratio is due almost entirely to the decrease in $(E_T)_{tre}$ with increasing linking difference from $|\Delta\ell| = 0$ to 15.

Simulations were also performed with $C = 3.0 \times 10^{-19}$ dyne cm$^2$, which is the value employed in practically all previous simulations. The resulting $E_T$ values are displayed

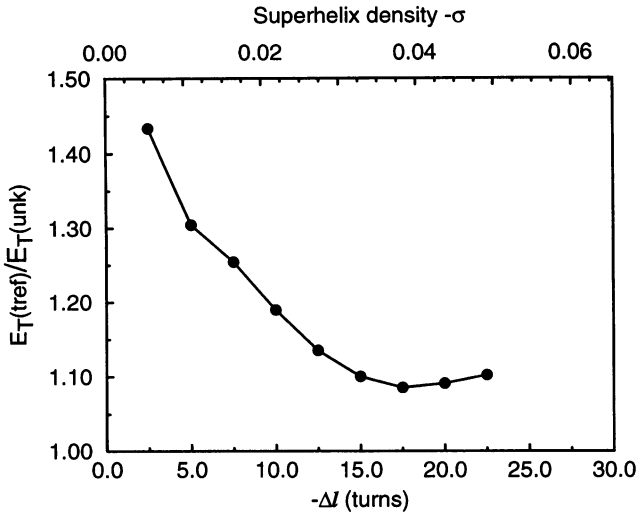


FIGURE 3 Ratio of twist energy parameters ($E_T$(tre)/$E_T$(unk)) for trefoil knotted and unknotted DNAs vs. $-\Delta\ell$. Simulation parameters are $N = 170$ subunits (4770 bp), $P = 500$ Å, torsional rigidity $C = 2.0*10^{-19}$ dyne cm$^2$, and $d = 50$ Å.

in Fig. 4. For the trefoil knotted circle, the twist energy parameter extrapolates to $E_T \cong 1670$ at $|\Delta\ell| \cong 1$, and again declines with increasing $|\Delta\ell|$. For the unknotted circle the twist energy parameter is again nearly constant near the mean value $(\bar{E}_T)_{unk} \cong 1153 \pm 47$ over the whole range of $\Delta\ell$ values. It should be noted that as $|\Delta\ell|$ decreases toward 0, the fluctuations in writhe become much larger with much longer relaxation periods, so the statistical errors become much larger, and much longer simulation times are required to achieve satisfactory accuracy. Thus, within the simulation errors it appears that $\Delta G_{sc}(\Delta\ell)$ also varies nearly quadratically with $\Delta\ell$, when $C = 3.0 \times 10^{-19}$ dyne cm$^2$. If we assume that $(E_T)_{unk} = 1190$ for all $|\Delta\ell| \lesssim 15$, then the ratio of twist

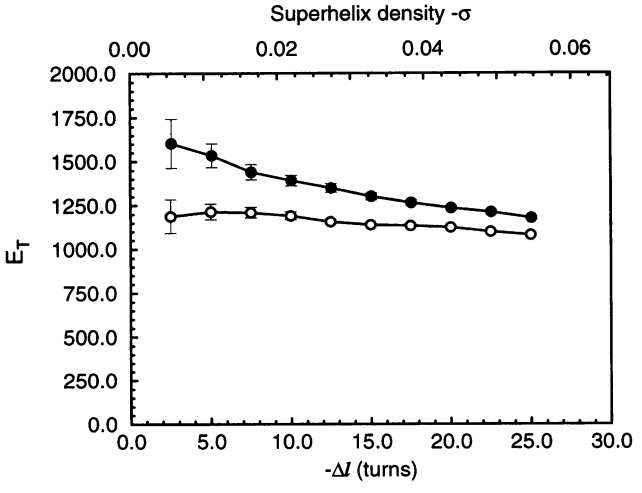


FIGURE 4 $E_T$ parameter vs. $-\Delta\ell$ for simulated trefoil knotted (●) and unknotted (○) circular DNA. Simulation parameters are $N = 170$ subunits (4770 bp), $P = 500$ Å, torsional rigidity $C = 3.0*10^{-19}$ dyne cm$^2$, and $d = 50$ Å. For the trefoil the abscissa actually is not $-\Delta\ell$, but instead is $-(\Delta\ell - \langle w_o\rangle)$, where $\langle w_o\rangle = -3.26$ is the nonzero writhe of the relaxed trefoil.

energy parameters in the region of small $\Delta\ell$ ($|\Delta\ell| \lesssim 3$) is again $(E_T)_{tre}/(E_T)_{unk} \cong 1.40$. The effect of increasing $C$ by 1.5-fold from (2.0 to 3.0) $\times 10^{-19}$ dyne cm$^2$ is to increase $E_T$ by about 1.15-fold for all $\Delta\ell$ values for *both* the unknotted and trefoil knotted circles, but the ratio is practically unchanged.

Fig. 5 presents the reduced writhe, $\langle w\rangle/\Delta\ell$, plotted vs. $\Delta\ell$ for both unknotted and knotted DNAs for $C = (2.0$ and $3.0) \times 10^{-19}$ dyne cm$^2$.

Typical structures obtained in the simulations with $C = 2.0 \times 10^{-19}$ dyne cm$^2$ at particular superhelix densities are displayed in Figs. 6 and 7.

## Comparison with results of previous simulations

Vologodskii and Cozzarelli (1994) report simulated values of $\Delta G_{sc}(\Delta\ell)$ vs. $-\sigma$ for a 5200 -bp filament with $P = 500$ Å, $C = 3.0 \times 10^{-19}$ dyne cm$^2$, and d = 50 Å at T = 310 K. Although the diameter $d = 50$ Å, is ascribed to 200 mM $Na^+$ by those authors, it is clearly associated with 150 mM $Na^+$ by Stigter (1977). In any case, they report that $\Delta G_{sc}(\Delta\ell)$ varies nearly quadratically with $\Delta\ell$, and from their data we estimate that $E_T \cong 1110$. The present simulations for a 4770-bp filament with the same parameters at $T = 298$ K yield an essentially quadratic variation of $\Delta G_{sc}(\Delta\ell)$ with $\Delta\ell$, and the mean value $\bar{E}_T \cong 1153 \pm 47$. Applying the temperature correction estimated above, $dE_T/dT = -2.00$ K$^{-1}$, gives finally $\bar{E}_T = 1129 \pm 47$ at 310 K. Thus, the present and previous simulations yield $E_T$ values in rather good agreement.

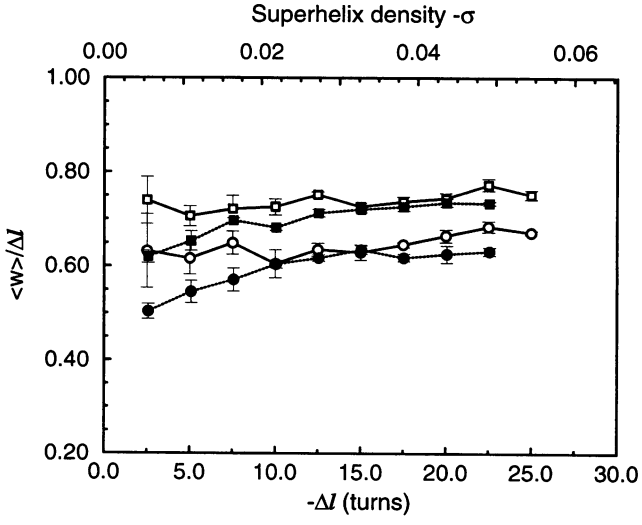


FIGURE 5 Reduced writhe ($\langle w\rangle/\Delta\ell$) vs. $-\Delta\ell$ for unknotted and trefoil knotted DNAs. Simulation parameters are $N = 170$ subunits (4770 bp), $P = 500$ Å, and $d = 50$ Å. The torsional rigidity is $C = 2.0*10^{-19}$ dyne cm$^2$ for trefoil knotted (●), $C = 2.0*10^{-19}$ dyne cm$^2$ for unknotted (○), $C = 3.0*10^{-19}$ dyne cm$^2$ for trefoil knotted (■), and C = $3.0*10^{-19}$ dyne cm$^2$ for unknotted (□) DNAs. The reduced writhe for the trefoil was calculated as $(\langle w\rangle - \langle w_o\rangle)/(\Delta\ell - \langle w_o\rangle)$, where $\langle w_o\rangle = -3.26$ is the nonzero writhe of the relaxed trefoil. For the trefoil the abscissa actually is not $-\Delta\ell$, but instead is $-(\Delta\ell - \langle w_o\rangle)$.

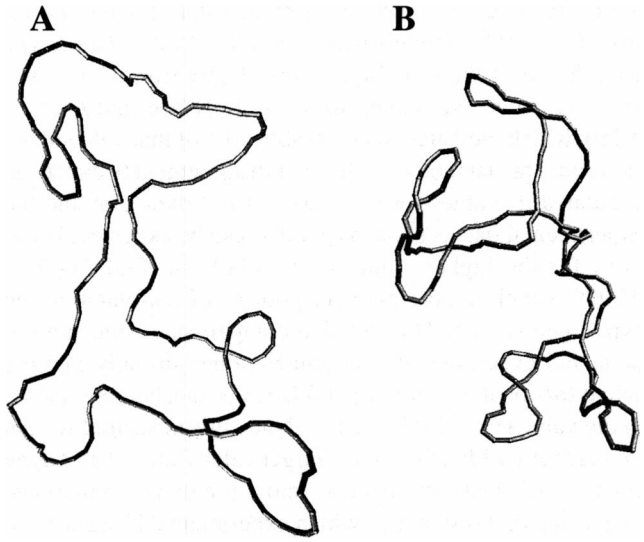


FIGURE 6 Configurations of simulated unknotted supercoiled DNA. Simulation parameters are $N = 170$ subunits (4770 bp), $P = 500$ Å, torsional rigidity $C = 2.0*10^{-19}$ dyne cm$^2$, $d = 50$ Å, $-\Delta\ell = 5.00$ ($A$), and $-\Delta\ell = 22.5$ ($B$).

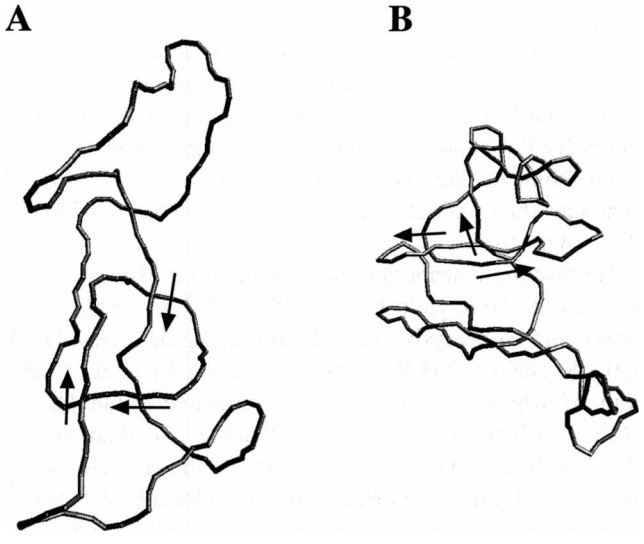


FIGURE 7 Configurations of simulated trefoil knotted supercoiled DNA. Simulation parameters are $N = 170$ subunits (4770 bp), $P = 500$ Å, torsional rigidity $C = 2.0*10^{-19}$ dyne cm$^2$, $d = 50$ Å and $-\Delta\ell = 5.00$ (A), and $-\Delta\ell = 22.5$ (B). The arrows denote the three nodes of the trefoil knot and are also parallel to the overcrossing strand.

The simulation results of Vologodskii and Cozzarelli (1994) for the aforementioned 5200-bp DNA at $\sigma = -0.06$ at 310 K, yield $\Delta G_{sc} \cong 117{,}500$ cal/mol, $\Delta S_{sc} = -129$ e.u. and $dE_T/dT = -2.3$ K$^{-1}$. Although this $dE_T/dT$ value is 15% larger than the value ($-2.0$ K$^{-1}$) estimated above for the 4770-bp DNA with the same $C$, $P$, and $d$ at $\sigma = -0.05$ at 298 K, the difference might not be significant.

Vologodskii and Cozzarelli (1994) also report reduced writhes, $\langle w\rangle/\Delta\ell \cong 0.75$, for DNAs containing either 3500 or 7000 bp with $C = 3.0 \times 10^{-19}$ dyne cm$^2$, $P = 500$ Å, and $d = 50$ Å at $\sigma = -0.05$ at 310 K. Within statistical error, we find the same value of the reduced writhe for the 4770-bp DNA with the same $C$, $P$, and $d$ at 298 K at (nearly) the same superhelix density, which corresponds to $\Delta\ell = -22.8$ turns. The simulated data of Vologodskii and Cozzarelli (1994) also suggest that the curve of $\langle w\rangle/\Delta\ell$ should be rather flat for a DNA with 4770 bp, and indeed the $\langle w\rangle/\Delta\ell$ values for the unknotted DNAs in Fig. 5 both define relatively flat "curves".

Typical structures obtained in the present simulations, which are shown in Figs. 6 and 7, resemble those reported by Vologodskii and coworkers (1992).

The evidence cited above suggests that the different sampling protocols used in the previous and present simulations cause no substantial differences in $E_T$ values, $dE_T/dT$ values, reduced writhe values, or typical structures. However, a much more rigorous comparison of other quantities, such as mean-squared radius of gyration, which may be more sensitive to any non-uniform sampling of the Riemann volume elements, will be required to ascertain whether the previous and present sampling protocols are completely equivalent.

Studies of the potential energy and writhe as a function of $\Delta\ell$ along the minimum energy path reveal multiple discontinuous transitions in writhe, which are accompanied by discontinuous changes in the slope of the minimum potential energy vs. $\Delta\ell$ (Hao and Olson, 1989; Schlick and Olson, 1992; Yang et al., 1993; Schlick et al., 1994). The present and previous simulations show no significant manifestation of such transitions in $\langle w\rangle$ vs. $\Delta\ell$, in $\Delta G_{sc}(\Delta\ell)$ vs. $\Delta\ell$, or in $E_T$ vs. $\Delta\ell$, for DNAs as long as those considered here. As noted above, the integrand in Eq. 18, namely $\Delta\ell' - \langle w\rangle$, is practically proportional to $\Delta\ell'$. In particular, its value at $\Delta\ell' = -1.0$, well below the putative circle-to-figure-8 transition at $|\Delta\ell'| \cong 1.6$, follows the same proportionality relation obeyed by the integrand values at larger $\Delta\ell'$. Hence, even that most prominent circle-to-figure-8 transition is effectively masked for such a long DNA at our present level of statistical precision. Evidently, the substantial fluctuations around the minimum energy configuration that occur at finite $T$ act to broaden and smoothe such transitions to the point where they are practically invisible. In fact, the standard deviation of the writhe is rather large and varies from about $\sigma_w = 0.85$ at $\Delta\ell = -1.0$ to $\sigma_w = 0.70$ at $\Delta\ell = -25.0$ turns.

## Comparison with experimental data for unknotted DNAs

Only a few experimental investigations of supercoiling free energies have been performed under conditions where the present simulations are expected to apply, specifically in 100–200 mM NaCl in the absence of $Mg^{2+}$. Recent measurements on p30δ DNA (4752 bp) in 0.1 M NaCl at 310 K indicate that $\Delta G_{sc}(\Delta\ell)$ varies nearly quadratically with $|\Delta\ell|$ from 0 up to 24 turns ($\sigma = -0.053$), as shown in Fig. 8, where 18 experimental $E_T$ values are plotted vs. $\Delta\ell$. The simulated $E_T$ values for $C = (2.0$ and $3.0) \times 10^{-19}$ dyne cm$^2$ are co-plotted vs. $\Delta\ell$ in the same figure. The torsional rigidity of

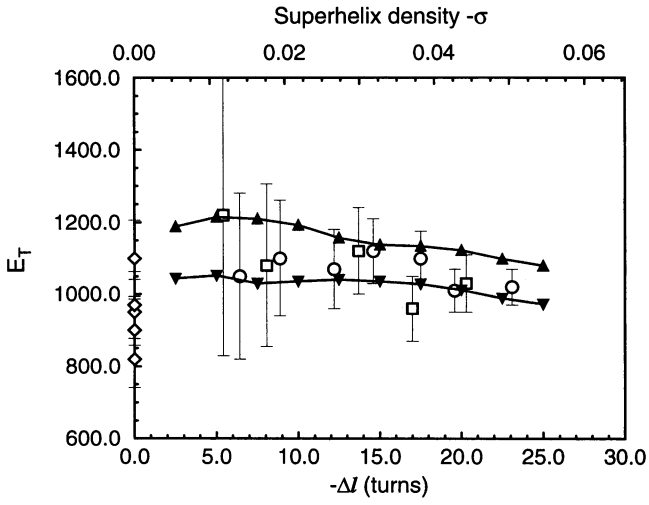


FIGURE 8 Experimental and simulated twist energy parameters ($E_T$) vs. $-\Delta\ell$. Experimental $E_T$ values were obtained by the topoisomer distribution method (◇) and ethidium-binding method (○) and (□) for p30δ DNA (4752 bp) in 100 mM NaCl, 10 mM Tris, 1 mM $Na^2$ EDTA, 1.48 mM sodium phosphate at 310 K (Clendenning et al., 1994). There are two identical values superposed at $E_T = 950$ at $\Delta\ell = 0$, so there are 18 total data points. Simulated values apply for $N = 170$ subunits (4770 bp), $P = 500$ Å, $d = 50$ Å, and either $C = 2.0*10^{-19}$ dyne cm$^2$ (▼) or $C = 3.0*10^{-19}$ dyne cm$^2$ (▲).

p30δ is obtained from time-resolved FPA (by assuming $P_d = 1500$ Å) and found to be $C = 2.07 \times 10^{-19}$ dyne cm$^2$. The average measured twist energy parameter is $\bar{E}_T = 1030 \pm 90$ (Clendenning et al., 1994). The latter quantity is quite close to the mean of the simulated values for $C = 2.0 \times 10^{-19}$ dyne cm$^2$, namely $\bar{E}_T = 1024$, and somewhat below that for $C = 3.0 \times 10^{-19}$ dyne cm$^2$, namely $\bar{E}_T = 1153$. To effect a detailed comparison between theory and experiment, we *assume* that the simulated values at $\Delta\ell = 0$ are just the respective weighted means of the simulated values, namely $(\bar{E}_T)_w = 1001$ for $C = 2.0 \times 10^{-19}$ dyne cm$^2$ and $(\bar{E}_T)_w = 1114$ for $C = 3.0 \times 10^{-19}$ dyne cm$^2$. The simulated $E_T$ vs. $\Delta\ell$ "curve" for $C = 2.0 \times 10^{-19}$ dyne cm$^2$ lies within one (experimental) standard deviation of 16 of the experimental data, and exceeds the remaining two by, respectively, 1.25 and 1.4 standard deviations. In contrast, the simulated $E_T$ vs. $\Delta\ell$ "curve" for $C = 3.0 \times 10^{-19}$ dyne cm$^2$ lies within one standard deviation of nine experimental data, and exceeds the nine other data by, respectively, 1.1, 1.6, 1.7, 1.9, 2.0, 2.0, 2.2, 2.7, and 3.9 standard deviations. A quantitative estimate of the relative likelihood that the weighted mean $E_T = 1001$ or $E_T = 1114$ gives rise to the experimental data is made as follows. We first compute

$$\chi^2 = \sum_{j=1}^{m} \frac{(E_{T_j} - E_T)^2}{\sigma_j^2}, \qquad (23)$$

where $E_{T_j}$ is the $j$th experimental value, $\sigma_j^2$ is its variance, and the index $j$ runs over all 18 experimental data. When $E_T = 1001$, we obtain $\chi_1^2 = 14.5$, and when $E_T = 1114$, we obtain $\chi_2^2 = 40.0$. The relative probability is then $P(1001)/P(1114) = \exp[-(\chi_1^2 - \chi_2^2)/2] = 3.4 \times 10^5$. Thus, it is enormously more probable that these experimental data for p30δ arise from $E_T = 1001$, or equivalently from $C = 2.0 \times 10^{-19}$ dyne cm$^2$, than from $E_T = 1114$, or equivalently from $C = 3.0 \times 10^{-19}$ dyne cm$^2$. Moreover, for $E_T = 1001$, we find $F(\chi^2) = 0.250$, which indicates an acceptable "fit" of that value to the experimental data. Indeed, the overall agreement between the simulated $E_T$ values for $C = 2.0 \times 10^{-19}$ dyne cm$^2$ and the experimental data is about as good as can be expected. However, for the higher value, $E_T = 1114$, we find $F(\chi^2) = 0.9979$, which indicates a very poor fit of that value to the experimental data. This detailed comparison of the simulations and experimental data argues rather strongly that the actual torsional rigidity of p30δ DNA is substantially closer to the value ($C = 2.07 \times 10^{-19}$ dyne cm$^2$) measured for the same DNA by FPA than to the larger value ($3.0 \times 10^{-19}$ dyne cm$^2$), which was employed in most previous simulations. This is the first instance in which experimental $E_T$ values are accurately matched over a broad range of $\Delta\ell$ values by simulations using (essentially) the independently measured torsional rigidity for the same DNA. Possible origins of the higher torsional rigidity, which was manifested in the small circle data of Horowitz and Wang (1984) and Shore and Baldwin (1983), will be discussed in a subsequent communication.

Pulleyblank et al. (1975) reported that, in 200 mM NaCl at 310 K, $E_T = 985$, 1115, and 1065 for, respectively PM2, ColE1, and Minicol DNAs near zero superhelix density. The values for PM2 and Minicol DNAs also agree better with the mean simulated value ($\bar{E}_T = 1024 \pm 26$) for $C = 2.0 \times 10^{-19}$ dyne cm$^2$ than with that ($E_T = 1153 \pm 47$) for $C = 3.0 \times 10^{-19}$ dyne cm$^2$.

Topoisomer distribution measurements on pBR322 in 0.1 M NaCl at 310 K yield $E_T = 1155 \pm 65$ (Naimushin et al., 1994), and an analysis of ethidium binding data on λ DNA in 0.1 M NaCl at 293 K yields $E_T = 1246$ (Wu et al., 1988), both of which are closer to the simulated value ($E_T = 1153 \pm 47$) for $C = 3.0 \times 10^{-19}$ dyne cm$^2$ than to that ($E_T = 1024 \pm 26$) for $C = 2.0 \times 10^{-19}$ dyne cm$^2$. However, the torsional rigidity of pBR322 obtained from FPA data (by assuming $P_d = 1500$ Å) is only $C = 2.2 \times 10^{-19}$ dyne cm$^2$, so the larger value of $E_T$ in that case most probably cannot be attributed to an appreciably greater torsional rigidity, although it could conceivably be ascribed to an enhanced persistence length.

Two additional problems are encountered with pBR322. First, ethidium binding data and FPA measurements indicate the presence of significant metastable secondary structure even in relaxed circles (Naimushin et al., 1994; Wu et al., 1988, 1991; Wu and Schurr, 1989; Schurr et al., 1992). Long-lived metastable secondary structure was also observed in freshly linearized pBR322 (Song et al., 1990). Secondly, measurements of $E_T$ by the topoisomer distribution method in the presence of $Mg^{2+}$ give mixed results. Horowitz and Wang (1984) obtain $E_T = 1130$ in 50 mM Tris, 10 mM $Mg^{2+}$ at 310 K, which is similar to the results of Naimushin et al. (1994) noted above. In contrast, Duguet (1993) obtains a significantly larger value, $E_T = 1320$, in 30 mM KCl, 30 mM

Tris, 5.5 mM $Mg^{2+}$ at 308 K, and Shore and Baldwin (1983) obtain a larger value yet, $E_T = 1610$, in 50 mM NaCl, 10 mM Tris, 10 mM $Mg^{2+}$ at 293 K. By using the temperature dependence reported by Duguet (1993), $\Delta E_T/\Delta T \cong (-)\,15.6$/K, the Shore-Baldwin value can be "corrected" to give $E_T = 1345$ at 310 K. For still unknown reasons, the pBR322 DNAs of Shore and Baldwin (1983) and Duguet (1993) are significantly stiffer than those of Horowitz and Wang (1984) and Naimushin et al. (1994). Indeed, *all* of the DNAs studied by Shore and Baldwin (1983), from 250 to 4363 bp, exhibit about 1.4-fold larger $E_T$ values than the DNAs of Horowitz and Wang (1984) (Clendenning and Schurr, 1994). Whether this is due to enhanced torsional and/or bending rigidity, to enhanced intersubunit repulsions, or to other causes, is not yet known. In any case, such large $E_T$ values cannot be simulated by the present model using $P = 500$ Å, $d = 50$ Å, and any previously reported estimate of $C$. If $C$ is constrained to lie within the range of previously measured values ($C \leq 3.7 \times 10^{-19}$), then a significantly enhanced value of $P$ or $d$ would be required to obtain $E_T \geq 1320$.

A notable discrepancy between the simulations, both past and present, and the experimental data concerns the magnitude of $dE_T/dT$. The simulations predict $dE_T/dT \cong -(2.0\text{–}2.3)\ K^{-1}$, whereas the experimental results of Duguet (1993) yield $dE_T/dT \cong -15.6\ K^{-1}$. Calorimetric studies of Seidl and Hinz (1984) also suggest that $dE_T/dT$ may be several times greater than the values predicted by simulations. These large negative values of $dE_T/dT$ reflect a large *positive* entropy of supercoiling in the experiments, in contrast to the small negative entropies of supercoiling in the simulations (Vologodskii and Cozzarelli, 1994). A substantial temperature dependence of the torsional and/or bending rigidity, or of the intersubunit interactions, would be required to account for such large negative $dE_T/dT$ values. However, previous measurements indicate that both the bending and twisting rigidities of ordinary linear DNAs are practically independent of $T$ (Gray and Hearst, 1968; Wilcoxon and Schurr, 1983; Soda and Wada, 1984; Robinson et al., 1980; Thomas and Schurr, 1983), and the effective hard-cylinder diameter is also expected to be practically constant (Stigter, 1977). Recent dynamic light scattering (DLS) studies on a nicked circular DNA containing 1876 bp also indicate no significant change in bending rigidity over the range from 5 to 40° C (J.A. Gebe and J.M. Schurr, unpublished results). It may be significant that Duguet's results pertain to topoisomer distributions of relaxed pBR322, which has been found repeatedly to exhibit metastable secondary structure (Wu et al., 1988, 1991; Wu and Schurr, 1989; Schurr et al., 1992; Naimushin et al., 1994). Unfortunately, there are no experimental data concerning $dE_T/dT$ for p30δ, which exhibits a significantly lower value of $E_T$, and which in our lab shows little or no evidence of metastable secondary structure. It might, then, be premature to generalize Duguet's results to p30δ or other circular DNAs that exhibit lower values of $E_T$. Also relevant to the present discussion is the fact that incomplete removal of contaminating proteins can cause the torsional rigidity to decrease substantially with increasing temperature (J. C. Thomas, unpublished results). A proper understanding of the origins of the large negative $dE_T/dT$ value reported by Duguet (1993) will require an independent assessment of any changes in torsional and/or bending rigidity of relaxed pBR322 with increasing $T$, e.g., by FPA and DLS, as well as measurements of $dE_T/dT$ for other DNAs with lower $E_T$ values such as p30δ.

Despite their rather different behavior, the sequences of p30δ and pBR322 are identical over 4087 bp. The only difference is that a 276-bp segment of pBR322 between the *Bam*HI site (position 375) and the *Sal*I site (position 657) is replaced by a 665-bp segment of yeast DNA.

### Comparison with experimental studies of knotted DNAs

Shaw and Wang (1993) determined $E_T$ values for both the trefoil knotted and unknotted forms of a 5600-bp DNA at 307 K. They obtained $E_T = 1310$ for the unknotted DNA and $E_T = 1760$ for the trefoil knotted DNA, both of which apply for small linking differences. The observed ratio, $(E_T)_{tre}/(E_T)_{unk} = 1.34$, agrees well with the simulated ratios, which are near 1.40 for $C = (2.0 \text{ and } 3.0) \times 10^{-19}$ dyne cm$^2$ at small $|\Delta\ell|$. However, the simulations predict that this ratio declines to ~1.1 at higher $|\Delta\ell| \gtrsim 15$, corresponding to superhelix densities, $|\sigma| \gtrsim 0.03$, as noted above. Unfortunately there are no relevant experimental data with which to compare this prediction.

Inspection of molecular configurations for the trefoil knotted DNA suggests a possible reason for the decline in $E_T$ with increasing $|\Delta\ell|$ from 0 up to ~15. It appears that the knot nodes occur preferentially at branch points of the superhelical structure, where they are evidently accommodated with the least distortion. Since the probability of branching increases with $\Delta\ell$, the ease of accommodating knot nodes also increases with $\Delta\ell$, and this acts to reduce the free energy cost of additional supercoiling up to $|\Delta\ell| \cong 15$.

Although the experimental and theoretical ratios agree well, the individual measured $E_T$ values are too large to be simulated by the present model using $P = 500$ Å, $d = 50$ Å, and any previously measured value of $C$. Significantly enhanced values of one or more of the input parameters, or some alteration of the model, will be required to satisfactorily simulate such elevated $E_T$ values. Indeed, the measured $E_T$ values are comparable to the higher values reported for pBR322 (Shore and Baldwin, 1983; Duguet, 1993).

## APPENDIX A. DISTANCE OF CLOSEST APPROACH OF CYLINDERS OF FINITE LENGTH

We consider the bond vector $\mathbf{b}_j$, which originates from $\mathbf{r}_{j-1}$ and terminates on $\mathbf{r}_j$, and the bond vector $\mathbf{b}_i$, which originates from $\mathbf{r}_{i-1}$ and terminates on $\mathbf{r}_i$. We consider also the two lines coincident with $\mathbf{b}_j$ and $\mathbf{b}_i$ that extend across the entire space. The location of an arbitrary point on the line coincident with $\mathbf{b}_j$ is given by $\mathbf{r}_{j-1} + \lambda_j \mathbf{b}_j$. If $0 \leq \lambda_j \leq 1.0$, then the point lies within the bond vector $\mathbf{b}_j$, but if $\lambda_j < 0$ or $\lambda_j > 1.0$, then it lies

on the extended line either behind or in front of the bond vector. Similarly, the location of an arbitrary point on the line coincident with $\mathbf{b}_i$ is given $\mathbf{r}_{j-1} + \lambda_i \mathbf{b}_i$. The squared distance between these two points, one on each line, is

$$D(\lambda_i, \lambda_j) = [(\mathbf{r}_{j-1} + \lambda_j \mathbf{b}_j) - (\mathbf{r}_{i-1} + \lambda_i \mathbf{b}_i)]^2 \quad \text{(A1)}$$

If $D(0.5, 0.5) > (b + d)^2$, then there is no possibility of any contact closer than $d$. However, if $D(0.5, 0.5) < (b + d)^2$, then the distance of closest approach is determined as follows. The minimum distance between the two lines is obtained by solving simultaneously the equation, $\partial D/\partial \lambda_j = 0$, which gives

$$\lambda_j = (\lambda_i(\mathbf{b}_i \cdot \mathbf{b}_j) - \mathbf{r}_{j-1,i-1} \cdot \mathbf{b}_j)/b^2 \quad \text{(A2)}$$

and $\partial D/\partial \lambda_i = 0$, which gives

$$\lambda_i = (\lambda_j(\mathbf{b}_j \cdot \mathbf{b}_i) + \mathbf{r}_{j-1,i-1} \cdot \mathbf{b}_i)/b^2 \quad \text{(A3)}$$

The minimum distance solutions are given by

$$(\lambda_i)_{\min} = \frac{b^2(\mathbf{b}_i \cdot \mathbf{r}_{j-1,i-1}) - (\mathbf{b}_j \cdot \mathbf{b}_i)(\mathbf{b}_j \cdot \mathbf{r}_{j-1,i-1})}{b^4 - (\mathbf{b}_i \cdot \mathbf{b}_j)^2} \quad \text{(A4)}$$

$$(\lambda_j)_{\min} = \frac{-b^2(\mathbf{b}_j \cdot \mathbf{r}_{j-1,i-1}) + (\mathbf{b}_j \cdot \mathbf{b}_i)(\mathbf{b}_j \cdot \mathbf{r}_{j-1,i-1})}{b^4 - (\mathbf{b}_i \cdot \mathbf{b}_j)^2} \quad \text{(A5)}$$

where $\mathbf{r}_{j-1,i-1} = \mathbf{r}_{j-1} - \mathbf{r}_{i-1}$. If $(\lambda_i)_{\min}$ and $(\lambda_i)_{\min}$ both lie in the interval 0–1.0, then they can be immediately substituted into Eq. A1 to obtain $D_{\min}$. If either $(\lambda_i)_{\min}$ or $(\lambda_j)_{\min}$ or both lie outside this range, then either or both of $\lambda_i$, $\lambda_j$ must take an extreme value of 0 or 1 within its bond vector. The minimum distance is found by evaluating all four of the following possibilities.

1. Set $\lambda_i = 0$, and solve Eq. A2 for $\lambda_j$. If $0 \le \lambda_j \le 1.0$, evaluate $D(0, \lambda_j)$. If $\lambda_j < 0$, set $\lambda_j = 0$, or if $\lambda_j > 1.0$, set $\lambda_j = 1.0$, and evaluate $D(0, \lambda_j)$. This procedure determines the distance from the nearest point within $\mathbf{b}_j$, specified by $0 \le \lambda_j \le 1.0$, to the point specified by $\lambda_i = 0$.
2. Set $\lambda_i = 1$, and solve Eq. A2 for $\lambda_j$. If $0 \le \lambda_j \le 1.0$, evaluate $D(1, \lambda_j)$. If $\lambda_j < 0$, set $\lambda_j = 0$, or if $\lambda_j \ge 1.0$, set $\lambda_j = 1.0$, and evaluate $D(1, \lambda_j)$. This procedure determines the distance from the nearest point within $\mathbf{b}_j$ specified by $0 \le \lambda_j \le 1.0$ to the point specified by $\lambda_i = 1.0$.
3. Set $\lambda_j = 0$, and solve Eq. A3 for $\lambda_i$. If $0 \le \lambda_i \le 1.0$, evaluate $D(\lambda_i, 0)$. If $\lambda_i \le 0$, set $\lambda_i = 0$, or if $\lambda_i > 1.0$, set $\lambda_i = 1.0$, and evaluate $D(\lambda_i, 0)$. This procedure determines the distance from the nearest point within $\mathbf{b}_i$ specified by $0 \le \lambda_i \le 1.0$ to the point specified by $\lambda_j = 0$.
4. Set $\lambda_j = 1.0$, and solve Eq. A3 for $\lambda_i$. If $0 \le \lambda_i \le 1.0$, evaluate $D(\lambda_i, 1.0)$. If $\lambda_i \le 0$, set $\lambda_i = 0$, or if $\lambda_i > 1.0$, set $\lambda_i = 1.0$, and evaluate $D(\lambda_i, 1.0)$. This procedure determines the distance from the nearest point within $\mathbf{b}_i$ specified by $0 \le \lambda_i \le 1.0$ to the point specified by $\lambda_j = 1.0$.

The smallest of the $D(\lambda_i, \lambda_j)$ values resulting from steps 1–4 yields the desired minimum distance.

Eqs. A4 and A5 cannot be used for parallel or antiparallel bond vectors, in which case their denominators diverge. For the parallel case, we first calculate $M = (\mathbf{b}_i \cdot \mathbf{r}_{j,i-1})(\mathbf{b}_i \cdot \mathbf{r}_{j-1,i})$ where $\mathbf{r}_{j,i-1} = \mathbf{r}_j - \mathbf{r}_{i-1}$ and $\mathbf{r}_{j-1,i} = \mathbf{r}_{j-1} - \mathbf{r}_i$, and evaluate one of the following four possible cases.

1. If $M > 0$ and $\mathbf{b}_i \cdot \mathbf{r}_{j,i-1} < 0$, calculate D(0, 1).
2. If $M > 0$ and $\mathbf{b}_i \cdot \mathbf{r}_{j,i-1} > 0$, calculate D(1, 0).
3. If $M < 0$ and $|\mathbf{b}_i \cdot \mathbf{r}_{j-1,i}| > |\mathbf{b}_i \cdot \mathbf{r}_{j,i-1}|$, set $\lambda_i = 0$, and find $\lambda_j^{\min}$ $(0 < \lambda_j^{\min} < 1.0)$ from Eq. A2. Then calculate $D(0, \lambda_j^{\min})$.
4. If $M < 0$ and $|\mathbf{b}_i \cdot \mathbf{r}_{j-1,i}| < |\mathbf{b}_i \cdot \mathbf{r}_{j,i-1}|$, set $\lambda_i = 1$, and find $\lambda_j^{\min}$ $(0 < \lambda_j^{\min} < 1.0)$ from Eq. A2. Then calculate $D(1, \lambda_j^{\min})$.

For the antiparallel case we calculate $M = (\mathbf{b}_i \cdot \mathbf{r}_{j,i}) \cdot (\mathbf{b}_i \cdot \mathbf{r}_{j-1,i-1})$ and evaluate one of the following four possible cases.

1. If $M > 0$ and $\mathbf{b}_i \cdot \mathbf{r}_{j-1,i-1} < 0$, calculate $D(0, 0)$.
2. If $M > 0$ and $\mathbf{b}_i \cdot \mathbf{r}_{j-1,i-1} > 0$, calculate $D(1, 1)$.
3. If $M < 0$ and $|\mathbf{b}_i \cdot \mathbf{r}_{ji}| > |\mathbf{b}_i \cdot \mathbf{r}_{j-1,i-1}|$, set $\lambda_i = 0$, and find $\lambda_j^{\min}$ $(0 < \lambda_j^{\min} < 1.0)$ from Eq. A2. Then calculate $D(0, \lambda_j^{\min})$.
4. If $M < 0$ and $|\mathbf{b}_i \cdot \mathbf{r}_{j,i}| < |\mathbf{b}_i \cdot \mathbf{r}_{j-1,i-1}|$, set $\lambda_i = 1.0$, and find $\lambda_j^{\min}$ $(0 < \lambda_j^{\min} < 1)$ from Eq. A2. Then calculate $D(1, \lambda_j^{\min})$.

## APPENDIX B. DERIVATION OF THE REDUCED PARTITION FUNCTION

The configuration integral for a closed circular DNA with linking number $\ell$ and a particular knot topology can be written formally as

$$Z = \int d\alpha_1\, d\beta_1 \sin\beta_1\, d\gamma_1 \int_C \cdots \int_C \prod_{i=1}^{N-1} \{d\alpha_{i,i+1}\, d\beta_{i,i+1} \sin\beta_{i,i+1}\, d\gamma_{i,i+1}\} \exp\left[-\frac{\mathcal{U}}{k_B T}\right] \quad \text{(B1)}$$

where the subscript $C$ on the integrals indicates that they must be constrained to closed molecules with the specified linking number $\ell$ and knot topology. In fact, the rotation $\alpha_{N1}$, $\beta_{N1}$, $\gamma_{N1}$ is completely fixed by the other rotations and the closure condition. The set of bond vectors $\{\mathbf{b}_i, i = 1, \cdots, N\}$ and relative positions $\{\mathbf{r}_i, i = 1, \cdots, N\}$ of the bond vector heads are completely determined by the angles $\alpha_1$, $\beta_1$, $\gamma_1$, and $\{\alpha^o_{i,i+1}, \beta_{i,i+1}, i = 1, \cdots, N-1\}$, where

$$\alpha^o_{12} = \alpha_{12} \quad \text{(B2a)}$$

$$\alpha^o_{i,i+1} = \alpha_{i,i+1} + \gamma_{i-1,i}, \qquad i = 2, \cdots, N-1 \quad \text{(B2b)}$$

For $i \ge 2$, $\alpha^o_{i,i+1}$ is the net rotation around the $z$ axis of the $i$th subunit between the $\beta_{i-1,i}$ and $\beta_{i,i+1}$ bends. The remaining angles $\alpha^o_{N1}$, $\beta_{N1}$ are determined from the others by the closure constraint. Thus, the spatial configuration of the chain, the value of its writhe, $w$, and its bending and subunit interaction energies, $U_b$ and $U_I$, respectively, are all determined by these angles. For any given spatial configuration of the chain, its state of twist is specified by the set of angles $\{\phi_{i,i+1}, i = 1, \cdots, N\}$, where $\phi_{i,i+1} = \alpha_{i,i+1} + \gamma_{i,i+1}$. For a spatial con-

figuration with net writhe $w$, $\sum_{i=1}^{N} \phi_{i,i+1} = 2\pi t = 2\pi(\ell - w)$. The fluctuation part of the twisting energy in Eq. 8 depends only on the set of angles $\{\mu_{i,i+1} = \phi_{i,i+1} - t/(2\pi N), i = 1, \cdots, N\}$. Hence, we introduce the coordinate transformation $\{\alpha_{i,i+1}, \gamma_{i,i+1}, i = 1, \cdots, N-1\} \rightarrow \{\alpha^{o}_{i,i+1}, \mu_{i,i+1}, i = 1, \cdots, N-1\}$ for which the Jacobian is 1.0. This gives

$$Z = 8\pi^2 \int_C \cdots \int_C \prod_{i=1}^{N-1} \{d\alpha^{o}_{i,i+1}\, d\beta_{i,i+1} \sin\beta_{i,i+1}\} \exp\left[-\frac{(U_b + U_I + (\alpha(2\pi)^2/2N)(\Delta\ell - w)^2}{k_B T}\right] \times \int \cdots \int d\mu_{12}\, d\mu_{23} \cdots d\mu_{N-1,N} d\mu_{N,1} \exp\left[(-)\left(\frac{\alpha}{2k_B T}\right)\sum_{i=1}^{N} \mu^2_{i,i+1}\right] \delta\left(\sum_{j=1}^{N} \mu_{j,j+1} - 0\right) \quad \text{(B3)}$$

wherein $\Delta\ell = \ell - \ell_0$ is the linking difference, so $\Delta\ell - w = t - \ell_0$. The constraint on the $d\mu_{12} \cdots d_{N-1,N}$ integral is incorporated into the integrand by including $d\mu_{N,1}$ and a $\delta$-function that enforces the relation $\sum_{j=1}^{N} \phi_{j,j+1} = 2\pi(\ell - w)$. The resulting $d\mu_{12} \cdots d\mu_{N,1}$ integral was evaluated previously (Wu et al., 1988) to give

$$q = (2\pi k_B T/\alpha)^{N/2} (2\pi N k_B T/\alpha)^{-1/2}, \quad \text{(B4)}$$

which is independent of $\ell$, $t$, $w$, or knot topology. Thus,

$$Z = 8\pi^2 q \int_C \cdots \int_C \prod_{i=1}^{N-1} \{d\alpha^{o}_{i,i+1}\, d\beta_{i,i+1} \sin\beta_{i,i+1}\} \exp\left[-\frac{U_b + U_I + (\alpha(2\pi)^2/2N)(\Delta\ell - w)^2}{k_B T}\right] \quad \text{(B5)}$$

It is clear from Eq. B5 that the fluctuations in local twist about the minimum energy deformation merely contribute the factor $q$, which is completely independent of $\ell$, $t$, $w$, and knot topology. The constant linking number constraint has been incorporated in the remaining integrand through the minimum twist energy term (proportional to $(\Delta\ell - w)^2$) in the exponent. The remaining constraints on the integrals pertain to chain closure and the knot topology. Since the integrand depends only on the set of bond vectors $\{\mathbf{b}_1, \cdots, \mathbf{b}_N\}$ and the integral is over the orientations of the $N - 1$ independent vectors in the successive subunit frames, we can change the integration variables to the orientations of those same vectors in the laboratory frame to obtain

$$Z = 2\pi q \int_C \cdots \int_C \prod_{i=1}^{N} \{d\alpha_i\, d\beta_i \sin\beta_i\} \exp\left[-\frac{U_b + U_I + (\alpha(2\pi)^2/2N)(\Delta\ell - w)^2}{k_B T}\right] \quad \text{(B6)}$$

where the factor $4\pi = d\alpha_1\, d\beta_1 \sin\beta_1$ has been reintroduced into the integral. Again, only N − 1 of the bond-vector orientations $\{\alpha_i, \beta_i, i = 1, \cdots, N\}$ are independent due to the closure constraint.

## APPENDIX C. RING CLOSURE ALGORITHM

We propose here an algorithm for absorbing the offset vector $\mathbf{r}_N$ into the $N$ bond vectors $\mathbf{b}_i$, $i = 1, \cdots N$, in an unbiased manner by making appropriate small changes in their orientations. This is done by subtracting from each bond vector a vector quantity that is proportional to the projection of $\mathbf{r}_N$ onto the plane perpendicular to that bond vector. The new corrected bond vector $\mathbf{b}^c_j$ is given by

$$\mathbf{b}^c_j = N_j(\mathbf{b}_j - a(\mathbf{1} - \hat{\mathbf{b}}_j\hat{\mathbf{b}}_j) \cdot \hat{\mathbf{u}}) \quad \text{(C1)}$$

where $\hat{\mathbf{u}}$ is a unit vector along $\mathbf{r}_N$, $\mathbf{b}_j$ is a unit vector along the original bond vector $\mathbf{b}_j$, $N_j$ is a normalization factor to maintain constant bond length, $|\mathbf{b}^c_j| = |\mathbf{b}_j| = b$, and $a$ is a proportionality constant that applies to all bond vectors and determines how much of the projection is to be subtracted off. The normalization is obtained by taking the dot product of Eq. C1 with itself and solving for $N_j$,

$$N_j = \left[1.0 + \left(\frac{a}{b}\right)^2 (1 - (\hat{\mathbf{u}} \cdot \hat{\mathbf{b}}_j)^2)\right]^{-1/2} \quad \text{(C2)}$$

The value of $a$ is determined by noting that the sum of the differences between the corrected and uncorrected bond vectors must be $(-)\mathbf{r}_N$, which can be written as

$$\sum_{j=1}^{N} (\mathbf{b}^c_j - \mathbf{b}_j) = \sum_{j=1}^{N} N_j[\mathbf{b}_j - a(\mathbf{1} - \hat{\mathbf{b}}_j\hat{\mathbf{b}}_j) \cdot \hat{\mathbf{u}}] - \mathbf{b}_j = -|\mathbf{r}_N|\,\hat{\mathbf{u}} \quad \text{(C3)}$$

Taking the dot product of $\hat{\mathbf{u}}$ with Eq. C3 gives

$$\sum_{j=1}^{N} [(N_j - 1)\mathbf{b}_j \cdot \hat{\mathbf{u}} - N_j a(1 - (\hat{\mathbf{u}} \cdot \hat{\mathbf{b}}_j)^2)] = -|\mathbf{r}_N| \quad \text{(C4)}$$

Since the correction to $\mathbf{b}_j$ is small, $N_j$ will be close to 1.0. Therefore, the expansion $(1 + x)^{-1/2} = 1 - x/2$ can be employed to write $N_j$ in Eq. C2 as

$$N_j \equiv 1 - \frac{1}{2}\left(\frac{a}{b}\right)^2 (1 - (\hat{\mathbf{u}} \cdot \hat{\mathbf{b}}_j)^2) \quad \text{(C5)}$$

Upon substituting this expression for $N_j$ into Eq. C4 and solving for $a$ to lowest order yields,

$$a = \frac{|\mathbf{r}_N|}{\sum_{j=1}^{N} (1 - (\hat{\mathbf{u}} \cdot \hat{\mathbf{b}}_j)^2)} \quad \text{(C6)}$$

The corrected bond vectors are calculated as follows. If $|\mathbf{r}_N| > (0.002)b$, then $a$ is calculated from Eq. C6, $N_j$ is calculated from Eq. C2, and $\mathbf{b}^c_j$ is calculated from Eq. C1. The new offset vector is calculated according to $\mathbf{r}^c_N = \sum_{j=1}^{N} \mathbf{b}^c_j$, and this procedure is iterated until $|\mathbf{r}^c_N| < 0.002\, b$. Typically, only one iteration is required for $\mathbf{r}^c_N$ to "converge" to or below the allowed limit. The potential energy is evaluated using the corrected bond vectors $\mathbf{b}^c_j$, $j = 1, \cdots N$, and

a new set of Euler angles $\Phi_j^c = (\alpha_j^c, \beta_j \gamma_j^c)$ is calculated only when the new configuration is to be kept.

## APPENDIX D. DETECTING CHANGES IN KNOT TYPE

The following summarizes a lengthier discussion presented by Frank-Kamenetskii and Vologodskii (1981). The knot type of a closed space curve can be characterized by its Alexander polynomial, $\Delta(t)$. This polynomial is invariant to changes in shape of the space curve, provided the curve is not broken and does not pass through itself. Not only is the Alexander polynomial an invariant of the knot, but for the simplest knots with which we are concerned here it is unique. In particular $\Delta(t) = 1$ for the unknotted molecule (trivial knot), and $\Delta(t) = t^2 - t + 1$ for the trefoil knot, which has three nodes. The distinct polynomials for all higher knots with up to eight nodes (minimum number of crossings in a planar projection) are given by Frank-Kamenetskii and Vologodskii (1981). This allows one to detect whether a change in knot type has occurred during a simulation move.

The Alexander polynomial associated with a given closed space curve is determined from a square Alexander matrix (**A**), which is constructed from a projection of that space curve onto any plane according to the following rules. As one passes along the projected space-curve in a given direction, one encounters successive underpasses (and overpasses). The underpasses are enumerated, beginning with 1. The segment of the contour between the $(k - 1)$th and $k$th underpasses is called the $k$th generator of the projected space curve. Associated with the $k$th underpass is its overpassing generator, denoted as the $i$th. The underpass can be either of two types. When oriented so that the overpassing $i$th generator passes south to north, the underpass from the $k$th to $(k + 1)$th generator is type I when it passes from west to east, and type II when it passes from east to west. The dimension of the Alexander matrix is $n$, the total number of undercrossings. The $k$th row corresponds to the $k$th underpass. All elements $a_{kj}$ in the $k$th row are zero, except for $a_{kk}$, $a_{k,k+1}$, and $a_{ki}$. The nonzero elements of the $k$th row are defined as follows:

1. When $i = k$ or $i = k + 1$, then $A_{kk} = -1$, $A_{k,k+1} = 1$, regardless of the type of underpass.
2. When $i \neq k$, $i \neq k + 1$, then $A_{kk} = 1$, $A_{k,k+1} = -t$, $A_{ki} = t - 1$ for a type I underpass, and $A_{kk} = -t$, $A_{k,k+1} = 1$, $A_{ki} = t - 1$ for a type II underpass.

The Alexander polynomial is obtained from the Alexander matrix by computing any $(n - 1)$th-order minor and multiplying that by $t^{-m}$ (integral $m$) so that the resulting polynomial has no negative powers and a positive constant term. The final result is independent of the particular minor used. Evaluation of the polynomial at $t = -1$ allows us to readily distinguish the unknotted (trivial knot) DNA ($\Delta(-1) = 1$) from the trefoil knotted DNA ($\Delta(-1) = +3$), and to further distinguish either of these from other simple knots that might arise during a single simulation move.

This work was supported in part by grants DMB 90-06651 and MCB 93-17402 from the National Science Foundation, an NIH Molecular Biophysics Training Grant to J. A. G., and a grant from the U. W. Royalty Research Fund.

## REFERENCES

Chen, Y. 1981. Monte Carlo study of freely jointed ring polymers. II. The writhing number. *J. Chem. Phys.* 75:2447–2453.

Clendenning, J. B., A. N. Naimushin, B. S. Fujimoto, D. W. Stewart, and J. M. Schurr. 1994. Effect of ethidium binding and superhelix density on the supercoiling free energy and torsion and bending constants of p30δ. *Biophys. Chem.* 52:191–218.

Clendenning, J. B., and J. M. Schurr. 1994. Circularization of small DNA in the presence of ethidium: a theoretical analysis. *Biopolymers.* 34:849–868.

Duguet, M. 1993. The helical repeat of DNA at high temperature. *Nucleic Acids Res.* 21:463–468.

Frank-Kamenetskii, M. D., A. V. Lukashin, V. V. Anshelevich, and A. V. Vologodskii. 1985. Torsional and Bending Rigidity of the Double Helix from Data on Small DNA Rings. *J. Biomol. Struct. Dyn.* 2:1005–1012.

Frank-Kamenetskii, M. D., and A. V. Vologodskii. 1981. Topological aspects of the physics of polymers: the theory and its biophysical applications. *Sov. Phys. Usp.* 24:679–696.

Fujimoto, B. S., and J. M. Schurr. 1990. Dependence of the torsional rigidity of DNA on base composition. *Nature.* 344:175–178.

Gray, H. B., and J. E. Hearst. 1968. Flexibility of native DNA from the sedimentation behavior as a function of molecular weight and temperature. *J. Mol. Biol.* 35:111–129.

Hao, M., and W. K. Olson. 1989. Global equilibrium configurations of supercoiled DNA. *Macromolecules.* 22:3292–3303.

Hinton, D. M., and V. C. Bode. 1975. Ethidium binding affinity of circular λ deoxyribonucleic acid determined fluorometrically. *J. Biol. Chem.* 250: 1061–1070.

Horowitz, D. S., and J. C. Wang. 1984. Torsional rigidity of DNA and length dependence of the free energy of supercoiling. *J. Mol. Biol.* 173:75–94.

Klenin, K. V., A. V. Vologodskii, V. V. Anshelevich, A. M. Dykhne, and M. D. Framk-Kamenetskii. 1991. Computer Simulation of DNA Supercoiling. *J. Mol. Biol.* 217:413–419.

Klenin, K. V., A. V. Vologodskii, V. V. Anshelevich, V. Y. Klishko, A. M. Dykhne, and M. D. Frank-Kamenetskii. 1989. Variance of writhe for wormlike DNA rings with excluded volume. *J. Biomol. Struct. Dyn.* 6:707–714.

LeBret, M. 1980. Monte Carlo computation of the supercoiling energy, the sedimentation constant, and the radius of gyration of unknotted and knotted circular DNA. *Biopolymers.* 19:619–637.

LeBret, M. 1986. Computation of the radius of gyration of a closed Ring to which linear side chains are attached. *Biochemistry.* 25:5734–5735.

Levene, S. D., and D. M. Crothers. 1986. Topological distributions and the torsional rigidity of DNA. A monte Carlo study of DNA circles. *J. Mol. Biol.* 189:73–83.

Madras, N., and A. D. Sokal. 1987. Non-ergodicity of local, length-conserving Monte Carlo algorithms for the self-avoiding walk. *J. Stat. Phys.* 47:573–595.

Madras, N., and A. D. Sokal. 1988. The pivot algorithm: a highly efficient Monte Carlo method for the self-avoiding walk. *J. Stat. Phys.* 50: 109–186.

Metropolis, N., A. W. Rosenbluth, A. H. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21:1087–1092.

Naimushin, A. N., J. B. Clendenning, U. Kim, L. Song, B. S. Fujimoto, D. W. Stewart, and J. M. Schurr. 1994. Effect of ethidium binding and superhelix density on the supercoiling free energy and torsion constant of pBR322. *Biophys. Chem.* 52:219–236.

Pulleyblank, D. E., M. Shure, D. Tang, J. Vinograd, and H. P. Vosberg. 1975. Action of nicking-closing enzyme on supercoiled and non-supercoiled closed circular DNA: Formation of a Boltzmann distribution of topological isomers. *Proc. Natl. Acad. Sci. USA.* 72:4280–4284.

Robinson, B. H., L. S. Lerman, A. Beth, H. L. Frisch, L. R. Dalton, and C. Auer. 1980. Analysis of double-helix motions with spin-labeled probes:

binding geometry and the limit of torsional elasticity. *J. Mol. Biol.* 139: 19–44.

Schlick, T., and W. K. Olson. 1992. Supercoiled DNA energetics and dynamics by computer simulation. *J. Mol. Biol.* 223:1089–1119.

Schlick, T., W. K. Olson, T. Wescott, and J. C. Greenberg. 1994. On higher buckling transitions in supercoiled DNA. *Biopolymers.* 34: 565–597.

Schurr, J. M., B. S. Fujimoto, P. Wu, and L. Song. 1992. Topics in Fluorescence Spectroscopy, Vol. 3. Plenum Press, New York. 137–229.

Schutle, C. J. H. 1984. Angular momentum and euler angle conventions in molecular spectroscopy. *Appl. Spectrosc. Rev.* 20:347–371.

Seidl, A., and H. J. Hinz. 1984. The free energy of DNA supercoiling is enthalpy determined. *Proc. Natl. Acad. Sci. USA.* 81:1312–1316.

Shaw, S. Y., and J. C. Wang. 1993. Knotting of a DNA chain during ring closure. *Science.* 260:533–536.

Shimada, J., and H. Yamakawa. 1985. Statistical mechanics of DNA topoisomers. The helical worm-like chain. *J. Mol. Biol.* 184:319–329.

Shimada, J., and H. Yamakawa. 1988. Moments for DNA topoisomers: the helical wormlike chain. *Biopolymers.* 27:657–673.

Shore, D., and R. L. Baldwin. 1983. Energetics of DNA twisting II. Topoisomer analysis. *J. Mol. Biol.* 170:983–1007.

Soda, K., and A. Wada. 1984. Dynamic light scattering studies on thermal motions of native DNAs in solution. *Biophys. Chem.* 20:185–200.

Song, L., B. S. Fujimoto, P. Wu, J. C. Thomas, J. H. Shibata, and J. M. Schurr. 1990. Evidence for allosteric transitions in secondary structure induced by superhelical stress. *J. Mol. Biol.* 214:307–326.

Stigter, D. 1977. Interactions of highly charged colloidal cylinders with applications to double stranded DNA. *Biopolymers.* 16:1435–1448.

Taylor, W. H., and P. J. Hagerman. 1990. Application of the method of T4 phage DNA ligase-catalyzed ring-closure to the study of DNA structure II. NaCl-dependence of DNA flexibility and helical repeat. *J. Mol. Biol.* 212:363–376.

Thomas, J. C., and J. M. Schurr. 1983. Fluorescence depolarization and temperature dependence of the torsion elastic constant of linear $\phi 29$ DNA. *Biochemistry.* 22:6194–6198.

Vologodskii, A. V., and N. R. Cozzarelli. 1994. Conformational and Thermodynamic Properties of Supercoiled DNA. *Annu. Rev. Biophys. Biomol. Struct.* 23:609–643.

Vologodskii, A. V., S. D. Levene, K. V. Klenin, M. D. Frank-Kamenetskii, and N. R. Cozzarelli. 1992. Conformational and thermodynamic properties of supercoiled DNA. *J. Mol. Biol.* 227:1224–1243.

Wilcoxon, J. P., and J. M. Schurr. 1983. Temperature dependence of the dynamic light scattering of linear $\pi 29$ DNA: implications for spontaneous opening of the double-helix. *Biopolymers.* 22:2273–2321.

Wu, P., L. Song, J. B. Clendenning, B. S. Fujimoto, A. S. Benight, and J. M. Schurr. 1988. Interaction of chloroquine with linear and supercoiled DNAs. Effect on the torsional dynamics, rigidity, and twist energy parameter. *Biochemistry.* 27:8128–8144.

Wu, P. G., and J. M. Schurr. 1989. Effects of chloroquine on the torsional dynamics and rigidities of linear and supercoiled DNAs at low ionic strength. *Biopolymers.* 28:1695–1703.

Wu, P. G., B. S. Fujimoto, L. Song, and J. M. Schurr. 1991. Effect of ethidium on the torsion constants of linear and supercoiled DNAs. *Biophys. Chem.* 41:217–236.

Yang, Y., I. Tobias, and W. K. Olson. 1993. Finite element analysis of DNA supercoiling. *J. Chem. Phys.* 98:1673–1686.