

Sequence-Dependent Effects in the Cyclization of Short DNA

Luke Czaplá,[†] David Swigon,^{†,‡} and Wilma K. Olson^{*,†}

Department of Chemistry & Chemical Biology, Rutgers, the State University of New Jersey, Wright-Rieman Laboratories, 610 Taylor Road, Piscataway, New Jersey 08854, and Department of Mathematics, University of Pittsburgh, 519 Thackeray Hall, Pittsburgh, Pennsylvania 15260

Received January 17, 2006

Abstract: A new, computationally efficient Monte Carlo approach has been developed to estimate the ring-closure properties of short, realistically modeled DNA chains. The double helix is treated at the level of base-pair steps using an elastic potential that accounts for the sequence-dependent variability in the intrinsic structure and elastic moduli of the base-pair steps, including the known coupling of conformational variables. Rather than using traditional Metropolis-Monte Carlo techniques to generate representative configurations, a Gaussian sampling method is introduced to construct three-dimensional structures from linear combinations of the rigid-body parameters defining the relative orientation and displacement of successive base pairs. The computation of the J factor, the well-known ratio of the equilibrium constants for cyclization vs bimolecular association of a linear molecule, takes into account restrictions on the displacement and directions of the base pairs joined in ring closure, including the probability that the end-to-end vector is null and the terminal base pairs coincide. The increased sample sizes needed to assess the likelihood that very short chains satisfy these criteria are attained using the Alexandrowicz half-chain sampling enhancement technique in combination with selective linkage of the two-half-chain segments. The method is used to investigate the cyclization properties of arbitrary-length DNA with greatly enhanced sampling sizes, i.e., $O(10^{14})$ configurations, and to estimate J factors lower than 0.1 pM with high accuracy. The methodology has been checked against classic theoretical predictions of the cyclization properties of an ideal, inextensible, naturally straight, DNA elastic rod and then applied to investigate the extent to which one can account for the unexpectedly large J factors of short DNA chains without the need to invoke significant distortions of double helical structure. Several well-known structural features of DNA—including the presence of intrinsic curvature, roll-twist coupling, or enhanced pyrimidine-purine deformability—bring the computed J factors in line with the observed data. Moreover, periodically distributed roll-twist coupling reduces the magnitude of oscillations in J , seen in plots of J vs chain length, to the extent found experimentally.

Introduction

The representation of chain molecules with ends confined to a fixed separation and orientation is a long standing

problem in polymer physical chemistry that can be attacked from several points of view. In one approach the configurations of unconstrained linear molecules which meet preassigned geometric criteria, e.g., the distance and relative orientation of chain ends, are collected through exhaustive simulation studies.¹ This method has been used extensively in analyses of the kinetics of chain cyclization, i.e., ring

* Corresponding author fax: (732) 445-5958; e-mail: olson@rutchem.rutgers.edu.

[†] Rutgers, the State University of New Jersey.

[‡] University of Pittsburgh.

closure,^{2–6} and the formation of closed loops.^{7–9} An alternative approach to the ring-closure problem is to derive an analytical expression for the configurational partition function of a simplified polymer, such as a freely jointed (Gaussian) chain or a freely rotating (wormlike) chain, that satisfies the requisite end conditions.^{10–13} Such theories have provided important insights into experimental findings and serve as valuable benchmarks for simulations of the cyclization propensities of polymer chains with similar molecular properties.

The probability of polymer ring closure is typically described in terms of the Jacobson-Stockmayer J factor, or cyclization constant, defined as the ratio of the equilibrium constants for cyclization and bimolecular association of a linear molecule.¹⁰ If the chain is sufficiently long and exhibits ideal Gaussian (random-coil) behavior, the J factor decreases with chain length N as $N^{-3/2}$.¹⁰ At short chain lengths near the rigid-rod limit, where there is very little variation in the overall molecular structure, the likelihood of ring closure drops off sharply with a decrease in N . Thus the J factor exhibits a maximum at intermediate chain lengths comparable to 2–3 times the persistence length.^{12,14,15}

The unexpected, spontaneous cyclization of DNA molecules much shorter than the persistence length^{16,17} has renewed interest in ring-closure measurements and the intrinsic structure and deformability of double helical DNA. DNA chain segments less than 100 bp in length are cyclized up to 5 orders of magnitude more efficiently than expected from theoretical predictions based on the conventional representation of DNA as an ideal, inextensible, naturally straight elastic rod, i.e., helical wormlike chain model with a persistence length of ~ 500 Å.¹⁵ New theories account for these discrepancies in terms of the spontaneous melting of isolated base pairs,¹⁸ the presence of occasional sharp kinks in the chain backbone,^{19,20} and/or the reduction of the torsional modulus.¹⁷ The published models, however, do not consider the unique sequence-dependent structural features of the short minicircles and the possibility that the wide range of J factors which have been reported^{6,16,17,20} may be a natural consequence of the differences in intrinsic structure and deformability of the selected sequences. For example, the DNA molecules found to close most easily into tight minicircles contain a well-known nucleosome-positioning sequence²¹ with periodically repeating chemical features. This sequence includes two notable patterns: (i) TA base-pair steps, which recur at increments of ~ 10 bp, i.e., roughly a complete double helical turn, and (ii) AT-containing dimers (AA, TA, or TT), which alternate at half helical turns with GC base-pair steps. The TA steps are among the most easily deformed of all base-pair steps,²² and the separation of tracts of AA and TT dimers by GC-rich steps is implicated in DNA intrinsic curvature.²³ Furthermore, complementary base pairs remain intact even in the most severely deformed protein-mediated DNA bends, such as the ~ 80 bp wrapped a complete superhelical turn around the histone proteins in the nucleosome core particle,²⁴ and even the very sharpest (40 – 50°) known protein-induced DNA bends^{25,26} are typically smaller than the ($\sim 90^\circ$) values posited to account for the cyclization tendencies of short chains.

The occurrence of small rings thus presents new technical challenges to the simulation of DNA ring closure. It is difficult to accumulate a meaningful sample of closed configurations from the random sampling of a relatively stiff, naturally straight molecule. The approximation of DNA as an ideal elastic rod has only limited success in predicting the behavior of the molecule in the short-length regime, where configurational fluctuations do not result in the ends having sufficiently random alignment for the persistence length to account entirely for the cyclization efficiency. This is despite the fact that most mixed-sequence DNA is expected to have only negligible intrinsic curvature, and hence a naturally straight model would be considered appropriate. Some sequences, such as those containing certain motifs beginning with four to six adenines (A-tracts), however, possess significant intrinsic curvature that is known to facilitate ring closure.^{12,27}

Here we present the details of a new computationally efficient technique to estimate the J factors of short, realistically modeled DNA chains. We treat the double helix at the level of base-pair steps, i.e., dimers, making use of an elastic potential which governs the fluctuations in the relative orientation and displacement of successive base pairs from their intrinsic values. We account for sequence-dependent variability in both the intrinsic structure and the elastic moduli of the base-pair steps, including the known coupling of conformational variables and anisotropy of bending. In place of conventional Metropolis-Monte Carlo methods²⁸ for sampling configurations of constrained molecules, we utilize the quadratic form of the DNA elastic energy and sample unconstrained configurations using a standard Gaussian random number generator. We increase sample size considerably by combining the Alexandrowicz half-chain sampling enhancement technique,²⁹ with a novel method of selectively linking the pairs of half-chain segments that are likely to satisfy the end-to-end ring-closure criteria. The method can be used to estimate J factors lower than 0.1 pM with high accuracy and hence is suited for the investigation of short DNA molecules. DNA polymers with different degrees of intrinsic curvature and different patterns of dimeric flexibility are investigated with the new technique. After checking the methodology against classic theoretical predictions of the cyclization properties of an ideal, elastic DNA rod, we use the approach to investigate the extent to which one can account for the observed ring-closure properties of short chains without the need to invoke significant distortions of double helical structure. We consider the effects of intrinsic curvature, anisotropic bending, roll-twist coupling, and enhanced pyrimidine-purine deformability and compare the computed J factors against values found experimentally for DNA chains of the same length and chemical content.

Methods

DNA Model. We make use of a dimeric representation of DNA which incorporates the known effects of base sequence on the intrinsic structure and deformability of the constituent dinucleotide steps.²² The rest state of each dimer is described by six independent step parameters which specify the preferred orientation and displacement of neighboring base-

pair planes—three angular variables termed tilt, roll, and twist and three variables called shift, slide, and rise with dimensions of distance.³⁰ The configuration of a base-pair step is denoted by the vector Θ , with components θ_i ($i = 1\dots 6$) corresponding respectively to the instantaneous values of tilt, roll, twist, shift, slide, and rise at the given step.

We use an elastic potential to describe the fluctuations in structure at each base-pair step. The potential is parametrized in terms of the differences between the values of the base-pair step parameters for the assumed geometry of the step and the intrinsic base-pair step parameter values assigned to the minimum-energy state of the step. Coupling is incorporated in the model through off-diagonal terms in the energy expression (see below).

Defining Θ^0 as the vector containing the intrinsic base-pair step parameters, the potential of a single XZ base-pair step of the DNA is expressed by the double summation:

$$\Psi = \frac{1}{2} \sum_{i=1}^6 \sum_{j=1}^6 f_{ij} \Delta\theta_i \Delta\theta_j$$

$$\Delta\theta_i = \theta_i - \theta_i^0 \quad (1)$$

Here the f_{ij} are elements of the symmetric 6×6 elastic force constant matrix \mathbf{F} , which contains the noncoupled elastic constants along its diagonal and the coupled terms in its off-diagonal elements. The θ_i^0 are the step parameters of the minimum-energy reference state stored in Θ^0 and the θ_i are the corresponding values of the instantaneous configurational state stored in the vector Θ . Hence, a set of intrinsic base-pair step parameters Θ^0 and an elastic constant matrix \mathbf{F} are sufficient to monitor the fluctuations of a single base-pair step. The configuration of a DNA chain segment depends on the choice of parameters at each of the N base-pair steps, and the total energy U is a sum of the energies over all steps:

$$U = \sum_{n=1}^N \Psi_n \quad (2)$$

Configurational Sampling. The sampled configurations are chosen to satisfy a Boltzmann distribution. Here, rather than using traditional Metropolis-Monte Carlo sampling techniques²⁸ to generate representative configurations, we take advantage of the quadratic form for the energy of DNA in eq 1 and the consequent simplification of the configurational partition function as a product of Boltzmann terms for individual base-pair steps. Hence, the probability that a base-pair step will adopt configuration Θ is proportional to the Boltzmann factor of the dimer deformation energy

$$p(\Theta) \propto \exp(-\beta\Psi) = \exp\left[-\frac{\beta}{2} \sum_{i=1}^6 \sum_{j=1}^6 f_{ij} (\theta_i - \theta_i^0) (\theta_j - \theta_j^0)\right] \quad (3)$$

where $\beta = 1/k_B T$, k_B is the Boltzmann constant, and T is the absolute temperature. Here, for simplicity, we omit indices that denote the sequential location n and chemical identity XZ of the base-pair step.

To express the probability density function as a product of independent terms, we write the dimeric energy Ψ in

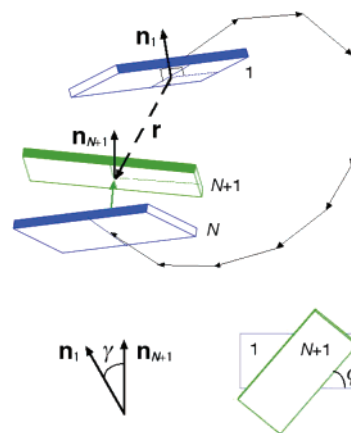


Figure 1. Linear DNA segment of N base pairs (blue blocks) in a configuration approaching the requirements for cyclization. The end-to-end vector \mathbf{r} (thick dashed arrow) joins the first base pair to a hypothetical (green) base pair $N+1$, which coincides with the first base pair in a perfectly closed chain. The net bending angle γ (lower left image) is defined by the normals \mathbf{n}_1 and \mathbf{n}_{N+1} of base pairs 1 and $N+1$ (thick solid arrows) and the end-to-end twist ϕ (lower right image) by the long axes and normals of the same base pairs.

matrix form, i.e., $\Psi = 1/2 \Delta\Theta^T \mathbf{F} \Delta\Theta$, diagonalize the force-constant matrix \mathbf{F} , and rewrite the energy in terms of a diagonal matrix \mathbf{D} and a basis variable Ω , with elements ω_i ($i = 1\dots 6$) given by linear combinations of the base-pair step parameters:

$$\Psi(\Theta) = \frac{1}{2} \Delta\Theta^T \mathbf{F} \Delta\Theta = \frac{1}{2} \Delta\Theta^T \mathbf{Q} (\mathbf{Q}^T \mathbf{F} \mathbf{Q}) \mathbf{Q}^T \Delta\Theta$$

$$= \frac{1}{2} (\Delta\Theta^T \mathbf{Q}) \mathbf{D} (\mathbf{Q}^T \Delta\Theta) = \frac{1}{2} \Omega^T \mathbf{D} \Omega \quad (4)$$

Here \mathbf{Q} is the eigenvector matrix specifying the directions of the principal axes of deformation. The superscript T is used to denote the transpose of a matrix or vector.

Elimination of the cross terms in the energy expression allows us to write the probability density function for a single base-pair step, including normalization, as a product of Gaussians:

$$p(\Theta) = \prod_{i=1}^6 \left(\frac{1}{2\pi\beta D_{ii}} \right)^{1/2} \exp\left(-\frac{\beta}{2} D_{ii} \omega_i^2 \right) \quad (5)$$

To sample this function, we modify a standard Gaussian random number generator³¹ and collect a Boltzmann distribution of dimeric states without the necessity of using the Metropolis method. This approach, which we term Gaussian sampling, is superior to the Metropolis method in that it is computationally more efficient and does not suffer from correlations between sample points or incomplete coverage of phase space.

Calculation of the J Factor. To monitor the closure of a DNA chain of N base pairs, we add a virtual ($N+1$)th base pair (of the same type as the first base pair) and consider the DNA to be closed when the ($N+1$)th base pair coincides with the first base pair (see Figure 1). Following Flory et al.,³² we express the J factor as a product of probabilities which describe the contribution of the spatial configuration

to the cyclization equilibrium constant:

$$J = \frac{4\pi}{N_A} W(\mathbf{r} = \mathbf{0}) \Gamma_r(\cos \gamma = 1) \Phi_{r,\cos\gamma}(\phi = 0) \quad (6)$$

Here $W(\mathbf{r} = \mathbf{0})$ is the probability that the configuration is closed, i.e., the end-to-end vector \mathbf{r} , which connects the first and $(N+1)$ th base pairs of the DNA, is zero. The factor $\Gamma_r(\cos \gamma = 1)$ is the conditional probability that the normals of terminal base pairs are aligned when the chain ends coincide, i.e., the cosine of the net bending angle is unity ($\cos \gamma = 1$) when the end-to-end separation $r = |\mathbf{r}|$ is zero ($r = 0$). The term $\Phi_{r,\cos\gamma}(\phi = 0)$ is the conditional probability that the first and last, i.e., $(N+1)$ th, base pairs coincide, i.e., the end-to-end twist ϕ is zero when both $r = 0$ and terminal base pairs are coplanar ($\cos \gamma = 1$). The quotient $N_A/4\pi$, where N_A is Avogadro's number, is the normalization constant associated with the uniformly distributed probability density of bimolecular association

For a DNA configuration specified by the angular parameters Θ_n , $n = 1 \dots N$, we calculate the end-to-end vector and orientation angles from the serial product of 4×4 generator matrices \mathbf{A}_n that incorporate the 3×1 displacement vector \mathbf{r}_n and the 3×3 transformation matrix $\mathbf{T}_{n,n+1}$, which relates coordinate frames on successive base pairs (n , $n+1$):

$$\begin{aligned} \mathbf{A}_{1:N} &= \mathbf{A}_1 \mathbf{A}_2 \dots \mathbf{A}_{N-1} \mathbf{A}_N \\ \mathbf{A}_n &= \begin{bmatrix} \mathbf{T}_{n,n+1} & \mathbf{r}_n \\ \mathbf{0} & 1 \end{bmatrix} \end{aligned} \quad (7)$$

The dependence of $\mathbf{T}_{n,n+1}$ and \mathbf{r}_n on Θ_n follows the formulation introduced by Zhurkin et al.³³ and subsequently developed by El Hassan and Calladine,³⁴ in which the angular step parameters are defined in terms of a sequence of symmetric Euler rotations and the translational components are expressed in the "middle" base-pair frame corresponding to the axis positions generated by half the rotational operation that brings adjacent base-pair frames into coincidence. See the literature^{33–35} for further details.

The end-to-end vector $\mathbf{r} = \mathbf{r}_{1:N}$ is accumulated in the far right column of $\mathbf{A}_{1:N}$:

$$\mathbf{r} = [\mathbf{I}_3 \quad \mathbf{0}] \mathbf{A}_{1:N} \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \quad (8)$$

Here \mathbf{I}_3 is the identity matrix of order three, and the $\mathbf{0}$'s are null matrices of orders necessary to fill the 3×4 premultiplication and 4×1 postmultiplication vectors.

The cosine of the angle between the normals of terminal base pairs is the (3,3) element of $\mathbf{A}_{1:N}$:

$$\cos \gamma = [0 \quad 0 \quad 1 \quad 0] \mathbf{A}_{1:N} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad (9)$$

The end condition imposed on the twisting of terminal base pairs is extracted from the trace of the end-to-end transformation matrix $\mathbf{T}_{1:N+1}$ accumulated in $\mathbf{A}_{1:N}$:

$$\text{Tr}(\mathbf{T}_{1:N+1}) = \cos \phi(1 + \cos \gamma) + \cos \gamma \quad (10)$$

Sampling Criteria. Because there is negligible likelihood that all three constraints on base-pair juxtaposition and orientation will be met exactly in numerical calculations, we relax the end conditions and consider as closed those configurations which meet the following criteria: (i) the length of the end-to-end vector \mathbf{r} is less than r_ϵ ; (ii) the cosine of the angle γ between the normals of terminal base pairs is larger than $1 - \nu_\epsilon$; and (iii) the magnitude of the end-to-end twist ϕ is less than τ_ϵ . Each generated configurational state is checked against these criteria, and each satisfaction is recorded. The probability densities are obtained as Bernoulli trials maximum-likelihood estimates, with each probability normalized by the volume of the enclosed phase space (see below).

The radial probability density $W(\mathbf{r} = \mathbf{0})$ is calculated as the number of configurations M_r obeying criterion (i), normalized by the sample size M and the phase space volume given as $4/3 \pi r_\epsilon^3$. The contribution $\Gamma_r(\cos \gamma = 1)$ from the alignment of normals is calculated as the number of configurations $M_{r,\cos\gamma}$ for which both criteria (i) and (ii) hold, normalized by M_r and the phase space volume ν_ϵ . Finally, the twist alignment density $\Phi_{r,\cos\gamma}(\phi = 0)$ is calculated as the number of configurations $M_{r,\cos\gamma,\tau}$ satisfying all three criteria, normalized by $M_{r,\cos\gamma}$ and its phase space $2\tau_\epsilon$. The estimated J factor, obtained by substitution of these values in eq 6, is thus given by

$$J = \frac{M_{r,\cos\gamma,\tau}}{QM} \quad (11)$$

where $Q = 4\pi N_A r_\epsilon^3 \nu_\epsilon \tau_\epsilon / 3$.

The standard deviation of this estimate, assuming sampled configurations are uncorrelated, is given by

$$\sigma = \frac{1}{Q} \sqrt{\frac{M_{r,\cos\gamma,\tau}(M - M_{r,\cos\gamma,\tau})}{M^3}} \quad (12)$$

When $M_{r,\cos\gamma,\tau}$ is much smaller than M , as is the case for closure probability calculations, the relative error is given by the approximate formula $\sigma/J \cong (M_{r,\cos\gamma,\tau})^{-1/2}$. Because the half-chain sampling method (see below) introduces slight correlation between sampled configurations, the true relative error (estimated as the standard deviation of multiple runs) is about 2.4 times larger than σ/J . Thus, the relative error is significant if $M_{r,\cos\gamma,\tau}$ is less than 100. For reported values of J greater than 0.1 pM, i.e., $\log J > -13$, the relative error is consistently below 10% (around 3% whenever feasible). The largest relative error occurs for smaller J factors (below 0.1 pM); data shown in this molar range are generally obtained with less accuracy, with as much as 32% relative error when the ends of short, naturally straight chains are out of phase, e.g., 10 accepted configurations for a 90 bp ideal, elastic B-DNA chain.

Sampling Enhancement. It is computationally expensive to generate ensembles of more than $O(10^7)$ configurations. Since DNA segments shorter than the persistence length are stiff and the probability that a randomly generated configuration satisfies all three ring-closure criteria is very small, the computed J factors are subject to large error if straightforward sampling approaches are used. Therefore, for such segments we use a sampling enhancement technique intro-

duced by Alexandrowicz²⁹ and subsequently applied by Levene and Crothers⁴ in Monte Carlo simulations of DNA ring closure. Rather than generate the configurations of the entire DNA with N base pairs, we divide the chain into two equal (or nearly equal) pieces and sample L configurations of each half-chain segment separately. By taking all pairwise combinations of both halves we can theoretically achieve L^2 configurations of the full-length chain. It should be noted that, even though the configurations of half-chain segments are uncorrelated, the multiplicative combination of a finite number of states introduces some bias in the full-chain ensemble.

If we let m be a base pair close to the midpoint of the chain, the end-to-end matrix $\mathbf{A}_{1:N}$ can be factored into two submatrices containing the structural details of the two-half-segments of the molecule:

$$\mathbf{A}_{1:N} = \mathbf{A}_{1:m} \mathbf{A}_{m:N} \quad (13)$$

Elements of these two submatrices are stored during the simulation and multiplied when half-segments combine successfully to form a closed, full-length chain. It follows from eq 13 and the definition of \mathbf{A}_n in eq 7 that the end-to-end vector can similarly be divided into two vectors that connect the chain ends via a central base pair.

$$\mathbf{r}_{1:N} = \mathbf{r}_{1:m} + \mathbf{T}_{1:m} \mathbf{r}_{m:N} \quad (14)$$

The requirement of ring closure, i.e., $\mathbf{r}_{1:N} = \mathbf{0}$, imposes a constraint on the vectors in the two halves of the DNA, namely

$$\mathbf{r}_{m:N} = -\mathbf{T}_{1:m}^{-1} \mathbf{r}_{1:m} \quad (15)$$

To reduce the number of unnecessary half-chain combinations, we join only those pairs of segments that are likely to satisfy the end-to-end ring-closure criterion. The current algorithm keeps track of the values of $-\mathbf{T}_{1:m}^{-1} \mathbf{r}_{1:m}$ for the first half-segments and $\mathbf{r}_{m:N}$ for the second half-segments and combines only those for which eq 15 is approximately satisfied. This is implemented as follows: the three-dimensional Cartesian space of \mathbf{r} is divided into cubes of volume V . As first and second half-configurations are being sampled we record cubes in which those configurations terminate in the sense that the values of $-\mathbf{T}_{1:m}^{-1} \mathbf{r}_{1:m}$ and $\mathbf{r}_{m:N}$ fall within the cube boundaries. Then, for each cube we combine all first half-configurations terminating in that cube with all second half-segments terminating in this same (primary) cube. We also combine those first half-configurations with the second half-segments terminating in all cubes that are less than r_c away from any point in the primary cube, where r_c is the radius of the sphere used to approximate the radial probability density $W(\mathbf{r} = \mathbf{0})$. First and second half-segments terminating in cubes that are farther than r_c apart need not be combined because they cannot satisfy the initial closure criterion, i.e., $\mathbf{r} = \mathbf{0}$.

This method discards the majority of possible configurational combinations, especially when sampling chain segments with a very low J factor. Thus, the total sample size can be, effectively, as large as 10^{14} , when necessary, to achieve a good estimate of the contribution of the conditional twist probability density $\Phi_{r,\cos\gamma}(\phi = 0)$. The latter quantity

is the most difficult to sample reliably and, except for the recent biased Monte Carlo calculations of the Vologodskii group,⁶ which unfortunately are not applicable to the present study of sequence-dependent effects, has not been considered in previous Monte Carlo simulations of DNA ring closure. Published estimates of the J factors obtained from Brownian dynamics simulations have also omitted the contribution of terminal base-pair overlap to the likelihood of chain closure.³⁶

The calculations were performed on a single 2.1 Ghz AMD Athlon(tm) processor with 2 gigabytes of available RAM. Numerical estimation of the J factor of a single molecule takes 1–10 CPU h depending upon DNA chain length and intrinsic structure. Simulations are longest for short, naturally straight chains where there is the least variation in overall molecular structure.

Results and Discussion

Intrinsically Straight DNA. We used the Gaussian sampling method in combination with the improved half-chain generation technique to investigate the cyclization probabilities of various models of double helical DNA. We tested the computational approach by first determining the ring-closure properties of an ideal, inextensible, naturally straight DNA molecule over the entire range of chain lengths between 90 and 450 bp. We chose a minimum-energy rest state with an intrinsic helical repeat of 10.5 bp/turn ($\sim 34.3^\circ$ twist at every base-pair step) and a pitch of 35.7 Å (3.4 Å rise at every base-pair step). All other step parameters were equated to zero.

In this simple *ideal* model, analogous to the Shimada-Yamakawa twisted wormlike chain representation of DNA¹⁵ (where the molecule is naturally straight, inextensible, subject to isotropic bending, and able to undergo independent fluctuations in twist), we allow deformations in tilt, roll, and twist ($\theta_1, \theta_2, \theta_3$) but fix the translational parameters shift, slide, and rise ($\theta_4, \theta_5, \theta_6$) near their intrinsic values (by use of very large force constants). The root-mean-square (RMS) fluctuation of tilt is equated to that of roll (isotropic bending), i.e., $\langle \Delta\theta_1^2 \rangle = \langle \Delta\theta_2^2 \rangle$, and is assigned a value of 4.84° , corresponding to a persistence length $a = 2\Delta s / (\langle \Delta\theta_1^2 \rangle + \langle \Delta\theta_2^2 \rangle)$ of nearly 500 Å (if Δs , the per residue base-pair displacement, is taken as 3.4 Å). The assumed RMS fluctuation in twist $\langle \Delta\theta_3^2 \rangle = 4.09^\circ$ corresponds to a global twisting constant $C = k_B T / \langle \Delta\theta_3^2 \rangle$ somewhat larger in magnitude than the global bending constant A , i.e., $C/A = 1.4$, where $A = ak_B T$. The choice of C is compatible with measurements of the equilibrium topoisomer distributions of DNA minicircles³⁷ and the fluorescence depolarization anisotropy of ethidium bromide molecules intercalated in DNA minicircles.³⁸

To estimate J , we set a radial bound r_c of 30 Å on the sphere used to monitor the end-to-end vector of the full-length chain, a limit of 30° or less on the angle between the normals of the first and last base pairs ($\cos\gamma > \sqrt{3}/2$), and a restriction on the magnitude of the end-to-end twist τ_c to values less than 30° . The same values are adopted in all subsequent simulations unless otherwise noted. As is clear from Table 1, the computed J factors are generally insensitive to the magnitude of the selected bounds. More restrictive bounds, however, lead to smaller numbers of acceptable configurations and hence to larger errors.

Table 1. Effect of Boundary Conditions on the Computed J Factor of a 210 bp, Ideal DNA Molecule^a

r_c (Å)	γ_c (deg)	τ_c (deg)	M	J (M)	relative error ^b (%)
30	30	30	3.6×10^{13}	1.26×10^{-8}	0.41
20	30	30	3.6×10^{13}	1.25×10^{-8}	0.74
10	30	30	3.6×10^{13}	1.27×10^{-8}	2.11
10	10	10	3.6×10^{13}	1.03×10^{-8}	12.5
10	10	10	1.44×10^{14}	0.92×10^{-8}	6.38

^a See legend to Figure 2 for details of model. ^b The relative error is estimated from multiple runs and equal to roughly $2.4\sigma/J$.

As is clear from Figure 2, the computed dependence of J on DNA chain length mimics the predictions of the Shimada-Yamakawa theory,¹⁵ showing a close fit with the observed cyclization efficiencies of chains of 240 bp or more^{39–42} and accounts satisfactorily for the reported ring-closure tendencies of other short (105–130 bp) DNA sequences.²⁰ The model, like the Shimada-Yamakawa theory, underestimates J by up to 3 orders of magnitude for some very short (89–105 bp) chains.^{16,17} These discrepancies have prompted our extension of the calculations to more realistic models of DNA with concomitant analysis of the effects of chemical features on cyclization properties as described below.

Influence of Natural Curvature. Individual base-pair steps adopt characteristic spatial forms and show different deformational tendencies in high-resolution DNA structures.²² These local turns and twists, if appropriately concatenated and then repeated in phase with the double helical repeat, can introduce a natural “static” curvature or superhelicity in the DNA.^{27,43,44} Indeed, it is rare to find segments of intrinsically straight DNA,⁴⁵ particularly in fragments as short as those employed in recent DNA ring-closure measurements.^{16,17,20}

Here we study an idealized, naturally curved DNA made up of two types of base-pair steps arranged in such a way that a molecule of 150 bp forms a stress-free, nearly circular configuration, or O-ring.³⁵ Half of the steps in the 10 bp repeating sequences are intrinsically straight ($\theta_1^0 = \theta_2^0 = 0$) with twist ($\theta_3^0 = 36^\circ$) corresponding to a 10 bp/turn helical repeat. The remaining steps have a different rest state with a positive roll angle ($\theta_2^0 = 7.41^\circ$) and an intrinsic twist ($\theta_3^0 = 35.57^\circ$) slightly lower than that of the preceding steps. The elastic properties at all steps, however, mimic those of an ideal DNA rod, i.e., the force constants are the same as those applied above in the simulation of naturally straight DNA. The sequence is constructed such that the first five steps assume the ideal B-DNA rest state and the last five steps adopt the perturbed rest state, i.e., an X_5Z_5 repeating pattern with XX and XZ dimer steps assigned the ideal B-DNA step parameters and ZZ and ZX steps assigned the set of modified parameters. As above, the intrinsic shift and slide are fixed at 0 Å, the rise at 3.4 Å, and the tilt at 0° at all base-pair steps.

Not surprisingly, the computed J factors of chain segments of the 150 bp O-ring are many orders of magnitude greater than those of a naturally straight molecule with the same elastic properties, chain length, and double helical repeat (compare the results reported in Figure 3 for naturally curved molecules of 90–210 bp and ideal, straight DNA chains of the same chain length and with the same 10 bp helical

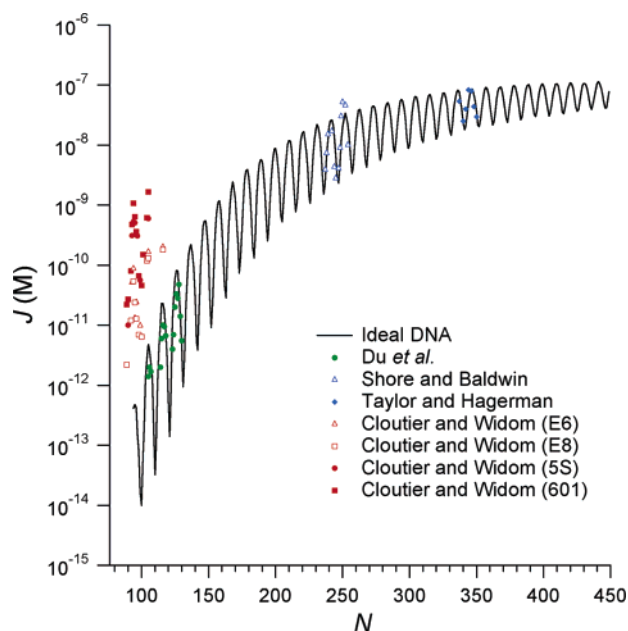


Figure 2. Monte Carlo estimates of the dependence on chain length of the J factors of ideal DNA, which has properties analogous to those of the Shimada-Yamakawa twisted worm-like chain:¹⁵ the double helix is assumed to be naturally straight in its equilibrium rest state, inextensible, and capable of isotropic bending and independent twisting at the base-pair level. The intrinsic twist and bending fluctuations are consistent with experimental measurements, i.e., 10.5 bp per helical turn and a persistence length of approximately 500 Å. Observed data points are denoted by color-coded symbols: filled and open red symbols (short nucleosome-binding and control sequences, respectively);^{16,17} filled green circles (other short sequences);²⁰ open blue triangles (*EcoRI* restriction fragments);^{40,41} and filled blue diamonds (*EcoRI* (E)-ended constructs).⁴² The curve is made up of a series of line segments connecting the computed J factors of molecules that differ in chain length by one base pair.

repeat). The enhancement in J is greatest for curved segments containing an integral number of helical turns. The ends of such chains readily meet in perfect register for successful ring closure. Interestingly, the chain segment which closes most easily into a circle is somewhat shorter than the perfect O-ring (130 vs 150 bp), reflecting the effects of thermal fluctuations on the configuration of DNA. Moreover, the cyclization probability maximum at 130 bp and the accompanying enhancement of the J factor compared to that of ideal, naturally straight DNA, i.e., $\log [J/J_0] = 7.48$, where J_0 is the Jacobson-Stockmayer factor of the latter model, agree remarkably well with values predicted for a naturally curved DNA with the same ratio of intrinsic curvature and persistence length.¹² Finally, although the curvature enhances the likelihood of DNA ring closure, the greater amplitude of the sawtooth oscillations in $\log J$ with N in the naturally curved DNA compared to those of a naturally straight model is counter to the observed changes in J at short chain lengths, where the amplitude is lower than that expected for a naturally straight DNA.¹⁷ The computed variation of $\log J$ vs N in Figure 3, however, is exaggerated by the lower radial bound imposed on the curved DNA sequences ($r_c = 10$ Å) compared to that imposed on the straight chain ($r_c = 30$ Å).

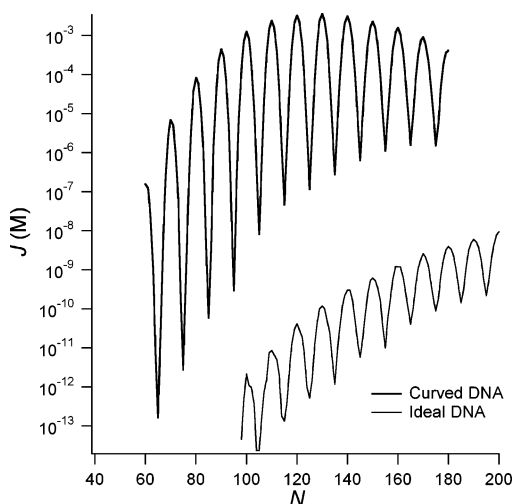


Figure 3. Predicted effects of intrinsic curvature on the J factors of DNA chains of 60–180 bp (thick solid curve). The intrinsically curved DNA is made up of two types of base-pair steps, each subject to isotropic bending and arranged such that a molecule of 150 bp forms a stress-free, nearly circular configuration (O-ring).³⁵ The corresponding variation of $\log J$ vs N for an ideal DNA with the same assumed helical repeat (10 bp per turn) is shown as the thin solid line.

Table 2 reports the J factors for a series of short (95, 100, 105 bp) DNA molecules of different intrinsic curvature, all with a 10.5 bp helical repeat and all determined with the same radial bound ($r_\epsilon = 30 \text{ \AA}$). The chains are constructed, as described above, from two types of base-pair steps: one with $\theta_2^0 = 0^\circ$ and $\theta_3^0 = 34.3^\circ$ and the other with nonzero values of intrinsic roll and slightly reduced dimeric twist (see the legend to the table). Only a small intrinsic roll in the latter steps ($\theta_2^0 \approx 2^\circ$) is required to bring the computed J factors in line with those reported for very short nucleosome-positioning sequences, e.g., $J_{\text{obs}} = 6.39 \times 10^{-10} \text{ M}$ at $N = 95$ bp and $1.66 \times 10^{-9} \text{ M}$ at $N = 105$ bp.^{16,17} The computed oscillation in the J factor between 95 and 105 bp, however, is still greater than that observed experimentally ($J_{\text{obs}} = 4.57 \times 10^{-11} \text{ M}$ at $N = 100$ bp).

Phased Bending Anisotropy. Compared to the ideal isotropic rod model, in which the double helix is assumed to be equally likely to bend in all directions, the bending of real DNA is anisotropic, with deformability in the direction of the major and minor grooves typically exceeding that across the grooves. In other words, the variation in roll is generally greater than that in tilt.^{22,33}

To assess the effects of anisotropic bending on ring closure, we consider a naturally straight, inextensible DNA

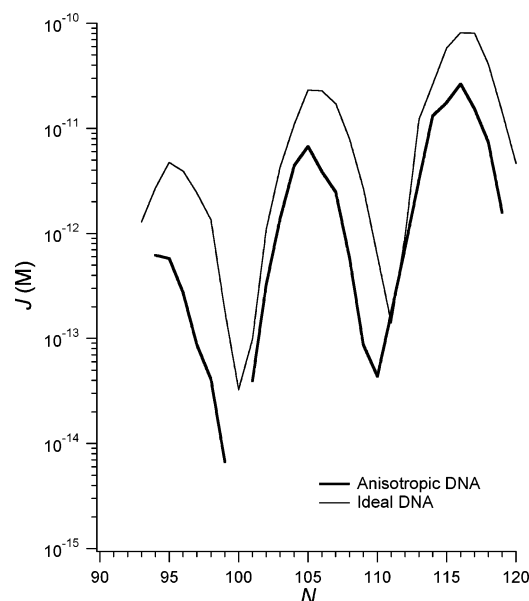


Figure 4. Influence of phased bending anisotropy on the computed likelihood of DNA ring closure in naturally straight chains of 94–119 bp (thick solid curve). Here the DNA is constituted from two types of base-pair steps: five successive steps are subject to ideal isotropic bending and the next five or six steps in the $(X_5Z_5X_5Z_6)_n$ conformational repeating pattern bend preferentially via roll in a hingelike manner. The data point at $N = 100$ bp is subject to large sampling error and thus omitted. The corresponding variation of $\log J$ vs N for an ideal DNA with the same assumed helical repeat (10.5 bp per turn) is shown as the thin solid line.

model in which roughly half of the base-pair steps are assigned the bending properties of an ideal, isotropic rod, and the remainder are subject to preferential bending via roll. The force constants of the former steps are the same as those employed above ($f_{11} = f_{22} = 4.84^{-2}$), while those of the latter steps ($f_{11} \approx 0^{-2}$ and $f_{22} = 6.84^{-2}$) are assigned so that the flexibility in roll exceeds that in tilt to such an extent, that fluctuations via tilt are prevented. This deformational pattern, which preserves the $\sim 500 \text{ \AA}$ persistence length of DNA,⁴⁶ corresponds to Schellman's classic "hinge" model of DNA bending.⁴⁷ The fluctuations in twist are governed at all steps by the same elastic constant used above, i.e., $f_{33} = 4.09^{-2}$. A 21-bp repeating sequence is employed to preserve the assumed 10.5 bp helical repeat, i.e., $(X_5Z_5X_5Z_6)_n$, where the XX and XZ steps bend isotropically and the ZZ and ZX steps exhibit hinge bending. Other step parameters are assigned the intrinsic values listed above.

Surprisingly, the DNA modeled with phased anisotropic bending is stiffer than an ideal DNA elastic rod in terms of

Table 2. Effect of Intrinsic Curvature on the J Factors of Short DNA Duplexes of Chain Length N

r_ϵ	intrinsic curvature ^a				
	210 bp	420 bp	630 bp	840 bp	straight
95	$8.42 \times 10^{-6} \text{ M}$	$3.83 \times 10^{-9} \text{ M}$	$1.52 \times 10^{-10} \text{ M}$	$4.48 \times 10^{-11} \text{ M}$	$4.75 \times 10^{-12} \text{ M}$
100	$3.21 \times 10^{-10} \text{ M}$	$1.62 \times 10^{-12} \text{ M}$	$3.94 \times 10^{-14} \text{ M}$	$2.74 \times 10^{-14} \text{ M}$	$3.25 \times 10^{-14} \text{ M}$
105	$3.77 \times 10^{-5} \text{ M}$	$2.37 \times 10^{-8} \text{ M}$	$1.18 \times 10^{-9} \text{ M}$	$2.18 \times 10^{-10} \text{ M}$	$2.31 \times 10^{-11} \text{ M}$

^a Intrinsic curvature is measured in terms of the number of base pairs of an idealized $X_5Z_5X_5Z_6$ repeating sequence designed to form an O-ring of the specified chain length. Intrinsic roll (θ_2^0) is fixed at zero and intrinsic twist (θ_3^0) at $\sim 34.3^\circ$ (10.5 bp per helical turn) at XX and XZ steps and at the following values at ZZ and ZX steps: 210 bp (5.4°, 34.03°); 420 bp (2.7°, 34.21°); 630 bp (1.8°, 34.25°); 840 bp (1.35°, 34.26°); straight (0°, $\sim 34.3^\circ$).

its ring-closure propensities, even though its closed configuration has lower elastic energy. The computed J factors are as much as an order of magnitude lower than those of an ideal rod with the same helical repeat for particular chain lengths in the range 94–119 bp (Figure 4). While the bending anisotropy may enhance deformations into the grooves, the local restrictions on bending across the grooves seemingly contribute to more effective cyclization. A possible reason for these observations may be that, whereas cyclized, ideal DNA is free to rotate about its helical axis, the corresponding global torsional motion of a covalently closed DNA subject to anisotropic bending is hindered by an energy barrier. That is to say, the ideal DNA molecule can “turn inside-out” and thereby cyclize with a bend in any direction with respect to the first base pair, whereas the hinged DNA chain is restricted to global bending in a single plane.

Roll-Twist Coupling. The localized twisting of DNA is strongly correlated with its bending.⁴⁸ The unwinding of the double helix (via decrease in twist) is generally accompanied by increased roll, i.e., local bending into the major groove, and the overwinding of the duplex (via increase in twist) is usually accompanied by decreased roll, i.e., local bending into the minor groove. Furthermore, the degree of twist-roll coupling is sequence-dependent, being generally most pronounced for pyrimidine-purine dimers and weakest for purine-pyrimidine steps.^{22,48}

Here we explore the effects of roll-twist coupling in an otherwise ideal model of B DNA, i.e., the chain is inextensible and naturally straight with a 10.5 bp helical repeat and the bending of successive base pairs is isotropic in the sense that $f_{11} = f_{22}$. The base-pair steps are subdivided into two categories and grouped, as above, into repeating $(X_5Z_5X_5Z_6)_n$ conformational blocks. The bending is independent of twisting in five successive (XX and XZ) base-pair steps, but roll is coupled to twist in the next five or six (ZZ and ZX) steps. The sign and magnitude of the twist-roll coupling modulus, $f_{23} = f_{32} = 5.41^{-2}$, are compatible with the interdependence of base-pair step parameters seen in high-resolution structures,²² i.e., the increase in roll will tend to lower twist and vice versa. Other elastic constants are held at canonical B-DNA values ($f_{11} = f_{22} = 4.84^{-2}$, $f_{33} = 4.09^{-2}$). The translational parameters and tilt are assigned the standard intrinsic values listed above.

Remarkably the introduction of roll-twist coupling into an otherwise ideal DNA chain increases the J factor and concomitantly reduces the amplitude of oscillations in $\log J$ with N (Figure 5). The computed variation in J with chain length closely matches the measured J factors of the “random” E8 sequences used as a control in the ring-closure studies of Cloutier and Widom^{16,17} (open symbols in Figure 5). The overall base-pair composition of the latter sequences is similar to that of the more easily cyclized “601” nucleosome-binding sequences (filled symbols in Figure 5), i.e., both sequence families are slightly GC-rich with 2–3 more pyrimidine-purine (YR) steps than random expectation, but there are distinct differences in the chemical content and distribution of the YR steps. Specifically, there is a greater proportion of TA dimers than CA·TG steps in the nucleosome-positioning sequences compared to the control se-

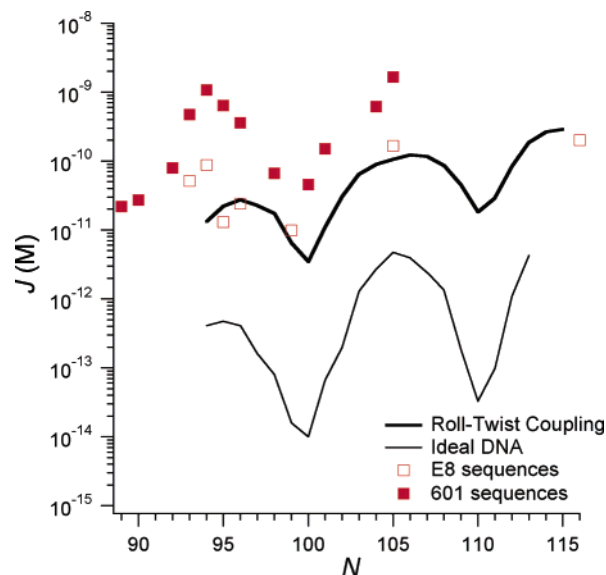


Figure 5. Predicted effects of phased roll-twist coupling on the J factors of naturally straight DNA chains of 94–115 bp (thick solid curve). Here the DNA is made up of two types of base-pair steps: five successive steps exhibit ideal elastic behavior, and the next five or six steps in the $(X_5Z_5X_5Z_6)_n$ conformational repeating pattern are subject to roll-twist coupling. The corresponding variation of $\log J$ vs N for an ideal DNA with the same assumed helical repeat (10.5 bp per turn) is shown as the thin solid line. The observed J factors of “601” nucleosome-positioning sequences^{16,17} are denoted by filled symbols and those of E8 control sequences^{16,17} by open symbols.

quences and the TA steps in the positioning sequence are spaced ~ 10 bp apart, in nearly perfect phase with the helical repeat. Notably, the energy required for melting the TA step, i.e., the TA·TA duplex fragment, is less than that of the CA·TG step,^{49,50} suggesting that conformational deformations of the TA are more facile and could dominate the motions of the double helix.

The natural changes in local twist, which accompany fluctuations in roll, clearly make it unnecessary to postulate¹⁷ that the DNA which forms small rings is more easily twisted than the classic ideal elastic rod. The build-up of roll, which is required to bring the ends of the DNA in contact, automatically enhances the degree of twisting in the model. See ref 51 for a related discussion showing that the observable, effective twisting resistance of a closed DNA ring with the same pattern of roll-twist coupling is close to zero.⁵¹

Phased Deformability. Large protein-induced bends of DNA occur predominantly at pyrimidine-purine steps,^{25,26} the dimers expected to be most easily bent on the basis of steric clash arguments⁵² and potential energy calculations.^{53,54} The enhanced deformability of such dimers contributes to the global plasticity of DNA without significant disruption of the double helical structure.

To test whether the regular spacing of flexible TA steps may enhance the ring closure of short nucleosome-positioning sequences, we constructed a hypothetical, naturally straight DNA model (with the same 10.5 bp helical repeat and elastic properties as ideal B-DNA) interspersed every 10 bp by a dimer with twice the RMS fluctuation of tilt and

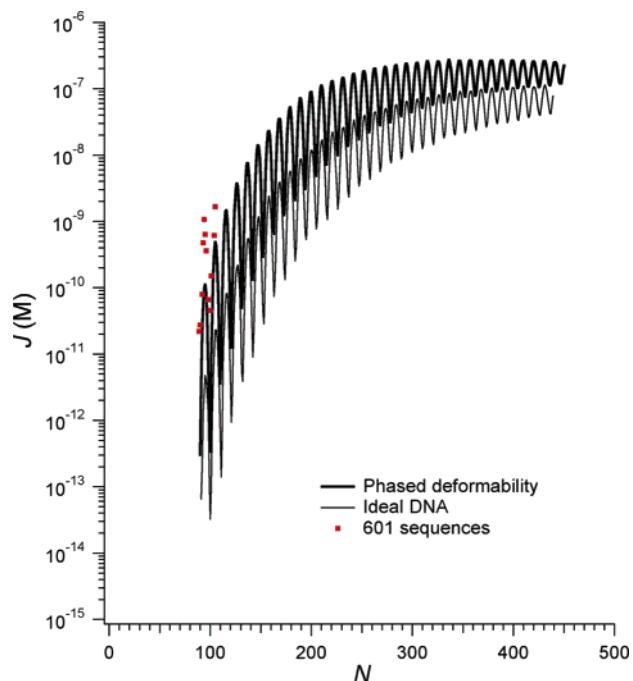


Figure 6. Influence of phased deformability on the computed likelihood of DNA ring closure in chains of 90–430 bp (thick solid curve). The ideal B-DNA model is interspersed every 10 bp by a deformable step subject to isotropic bending fluctuations twice those of the other steps. The corresponding behavior of an ideal DNA with the same assumed helical repeat (10.5 bp per turn) is shown as the thin solid line. The observed J factors of “601” nucleosome-positioning sequences^{16,17} are denoted by filled symbols.

roll, i.e., $f_{11} = f_{22} = (2 \times 4.84)^{-2} = 9.68^{-2}$. The latter range of bending is comparable to the composite variation of tilt and roll in the TA steps found in well resolved X-ray structures. For example, the sum of the fluctuations of tilt and roll ($\langle \Delta\theta_1^2 \rangle + \langle \Delta\theta_2^2 \rangle$) for TA steps in a dataset of 239 nonredundant protein-bound DNA structures of 2.5 Å or better resolution (Y. Li and W. K. Olson, unpublished data) is 15.26^2 if all extant data are considered and 7.02^2 if the most extremely deformed TA steps are omitted, i.e., values that bound the $(2 \times 9.68^2) = 13.69^2$ range of deformations assumed here. Values of intrinsic tilt, shift, slide, and rise are identical to those used in the analysis of naturally straight DNA.

As seen in Figure 6, the softening of DNA with force constants that are consistent with the deformability of TA steps and their placement in nucleosome-positioning sequences increases the J factor by over 2 orders of magnitude to the range of values found^{16,17} for short chains (filled symbols in the figure). The added fluctuations, however, do not dampen the oscillations of $\log J$ with N to the extent found experimentally. The amplitude of the variation in $\log J$ with N for the DNA chains with phased deformability is roughly 90% of that for the ideal DNA elastic rod. Although the imposed pattern of localized flexibility increases the J factor at all chain lengths, the effects are less pronounced at longer chain lengths. The enhancement in J levels off with increase in chain length to a factor ~ 2.5 times the value predicted for the ideal B-DNA duplex, the dif-

ference reflecting the lower persistence length of the locally softened DNA. The regularly spaced deformations also lower the chain length of maximum ring-closure probability, from a calculated value of 483 bp for the naturally straight DNA to 336 bp for the modified duplex. The cyclization propensities of long DNA multimers with regularly spaced, structural features such as those considered here are thus expected to differ from those of the mixed sequences traditionally used in the physical characterization of DNA.

Concluding Remarks. The present calculations clearly show that, in contrast to earlier suggestions,^{17–20} it is not necessary to invoke significant distortions in the DNA double helix, e.g., melting, kinking, “free” twisting, to bring the computed J factors in line with the enhanced cyclization propensities of molecules much shorter than the persistence length.^{16,17} A number of well-known features of DNA double helical structure—including the presence of intrinsic curvature, roll-twist coupling, or enhanced pyrimidine-purine deformability—can account for the unexpectedly large J factors of short chains. Moreover, chains with roll-twist coupling dampen the sawtooth oscillations in J with increase in chain length to the extent found experimentally. The internal coupling of dimeric variables also avoids the known detrimental effects of enhanced DNA deformability on nucleosome formation, i.e., the entropic penalty associated with the binding of a flexible sequence.⁵⁰

These conclusions could not have been drawn without the improvements in chain sampling introduced herein. Many of the rare “closed” arrangements of DNA captured in the course of these calculations would be missed with conventional Metropolis-Monte Carlo configurational sampling. We enhance the efficiency of finding closed states of DNA through the selective addition of half-chain segments, combining only pairs of half-segments that are known in advance to match the end-to-end spacing of the full-length molecule. The Gaussian sampling technique ensures complete coverage of phase space and does not suffer from the correlations between sample points associated with Metropolis-Monte Carlo approaches. The generation of spatial configurations from random moves along the eigenvectors of the base-pair step parameters allows us to consider the coupling of conformational variables that accounts so well for the mesoscale looping of DNA.

The improved sampling of DNA configurations also provides a valuable guide for the development of theories that use expansion methods to estimate the configurational partition function of a closed ring. The reliable fitting of the highly irregular spatial density distributions of short polymers, including DNA, has long been known to require spatial moments higher than second order.^{1,2,32,55} The extent to which approximate partition functions based on series expansions of the Boltzmann factor about the minimum energy configuration match the computed partition function will be reported elsewhere.

Finally, the present calculations omit consideration of the polyanionic character of DNA. Calculations of ring-closure probabilities that include long-range electrostatic interactions between anionic phosphate groups are expected to increase the chain length of maximum ring-closure probability.

Acknowledgment. We thank Drs. Jonathan Widom and Alexander Vologodskii for generously sharing experimental data, Dr. A. R. Srinivasan for assistance in modeling naturally curved DNA, and Dr. Edward W. Castner, Jr. for access to plotting software. Support of this work through U.S.P.H.S. Grant GM34809 is gratefully acknowledged. Computations were carried out at the Rutgers University Center for Computational Chemistry.

References

- (1) Yoon, D. Y.; Flory, P. J. Moments and distribution functions for polymer chains of finite length. II. Polymethylene chains. *J. Chem. Phys.* **1974**, *61*, 5366–5380.
- (2) Suter, U. W.; Mutter, M.; Flory, P. J. Macrocyclization equilibria. 2. Poly(dimethylsiloxane). *J. Am. Chem. Soc.* **1976**, *98*, 5740–5745.
- (3) Hagerman, P. J. Analysis of the ring-closure probabilities of isotropic wormlike chains: application to duplex DNA. *Biopolymers* **1985**, *24*, 1881–1897.
- (4) Levene, S. D.; Crothers, D. M. Ring closure probabilities for DNA fragments by Monte Carlo simulation. *J. Mol. Biol.* **1986**, *189*, 61–72.
- (5) Kahn, J. D.; Crothers, D. M. Protein-induced bending and DNA cyclization. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6343–6347.
- (6) Podtelezchnikov, A. A.; Mao, C. D.; Seeman, N. C.; Vologodskii, A. Multimerization-cyclization of DNA fragments as a method of conformational analysis. *Biophys. J.* **2000**, *79*, 2692–2704.
- (7) Marky, N. L.; Olson, W. K. Loop formation in polynucleotide chains. I. Theory of hairpin loop closure. *Biopolymers* **1982**, *21*, 2329–2344.
- (8) Vologodskii, A. V.; Levene, S. D.; Frank-Kamenetskii, M. D.; Cozzarelli, N. R. Conformational and thermodynamic properties of supercoiled DNA. *J. Mol. Biol.* **1992**, *227*, 1224–1243.
- (9) Rippe, K.; von Hippel, P. H.; Langowski, J. Action at a distance: DNA-looping and initiation of transcription. *TIBS* **1995**, *20*, 500–506.
- (10) Jacobson, H.; Stockmayer, W. H. Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.* **1950**, *18*, 1600–1606.
- (11) Yamakawa, H.; Stockmayer, W. H. Statistical mechanics of wormlike chains. II. Excluded volume effects. *J. Chem. Phys.* **1972**, *57*, 2843–2854.
- (12) Livshits, M. A. Calculating the probability for cyclization of a short intrinsically curved and twisted DNA fragment. *Mol. Biol.* **1996**, *30*, 85–90.
- (13) Zhang, Y. L.; Crothers, D. M. Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.* **2003**, *84*, 136–153.
- (14) Kratky, O.; Porod, G. Röntgenuntersuchung Gelöster Fadenmoleküle. *Rec. Trav. Chim.* **1949**, *68*, 1106–1122.
- (15) Shimada, J.; Yamakawa, H. Ring-closure probabilities for twisted wormlike chains. Application to DNA. *Macromolecules* **1984**, *17*, 689.
- (16) Cloutier, T. E.; Widom, J. Spontaneous sharp bending of double-stranded DNA. *Mol. Cell.* **2004**, *14*, 355–362.
- (17) Cloutier, T. E.; Widom, J. DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3645–3650.
- (18) Yan, J.; Marko, J. F. Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys. Rev. Lett.* **2004**, *93*, 108108.
- (19) Wiggins, P. A.; Phillips, R.; Nelson, P. C. Exact theory of kinkable elastic polymers. *Phys. Rev. E* **2005**, *71*, 021909.
- (20) Du, Q.; Smith, C.; Shiffeldrim, N.; Vologodskii, A. Cyclization of short DNA fragments and bending fluctuations of the double helix. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5397–5402.
- (21) Lowary, P. T.; Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **1998**, *276*, 19–42.
- (22) Olson, W. K.; Gorin, A. A.; Lu, X.-J.; Hock, L. M.; Zhurkin, V. B. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11163–11168.
- (23) Zhurkin, V. B.; Tolstorukov, M. Y.; Xu, F.; Colasanti, A. V.; Olson, W. K. Sequence-dependent variability of B-DNA: an update on bending and curvature. In *DNA Conformation and Transcription*; Ohyama, T., Ed.; Landes Bioscience/Eurekah.com, Georgetown, TX, 2005; pp 18–34.
- (24) Davey, C. A.; Sargent, D. F.; Luger, K.; Mäder, A. W.; Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **2002**, *319*, 1087–1113.
- (25) Suzuki, M.; Yagi, N. Stereochemical basis of DNA bending by transcription factors. *Nucleic Acids Res.* **1995**, *23*, 2083–2091.
- (26) Werner, M. H.; Gronenborn, A. M.; Clore, G. M. Intercalation, DNA kinking, and the control of transcription. *Science* **1996**, *271*, 778–784.
- (27) Crothers, D. M.; Drak, J.; Kahn, J. D.; Levene, S. D. DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.* **1992**, *212*, 3–29.
- (28) Metropolis, N. A.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
- (29) Alexandrowicz, Z. Monte Carlo of chains with excluded volume: a way to evade sample attrition. *J. Chem. Phys.* **1969**, *51*, 561–565.
- (30) Dickerson, R. E.; Bansal, M.; Calladine, C. R.; Diekmann, S.; Hunter, W. N.; Kennard, O.; von Kitzing, E.; Lavery, R.; Nelson, H. C. M.; Olson, W. K.; Saenger, W.; Shakked, Z.; Sklenar, H.; Soumpasis, D. M.; Tung, C.-S.; Wang, A. H.-J.; Zhurkin, V. B. Definitions and nomenclature of nucleic acid structure parameters. *J. Mol. Biol.* **1989**, *208*, 787–791.
- (31) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C*; Cambridge University Press: New York, 1986.
- (32) Flory, P. J.; Suter, U. W.; Mutter, M. Macrocyclization equilibria. 1. Theory. *J. Am. Chem. Soc.* **1976**, *98*, 5733–5739.

- (33) Zhurkin, V. B.; Lysov, Y. P.; Ivanov, V. I. Anisotropic flexibility of DNA and the nucleosomal structure. *Nucleic Acids Res.* **1979**, *6*, 1081–1096.
- (34) El Hassan, M. A.; Calladine, C. R. The assessment of the geometry of dinucleotide steps in double-helical DNA: a new local calculation scheme. *J. Mol. Biol.* **1995**, *251*, 648–664.
- (35) Coleman, B. D.; Olson, W. K.; Swigon, D. Theory of sequence-dependent DNA elasticity. *J. Chem. Phys.* **2003**, *118*, 7127–7140.
- (36) Merlitz, J.; Rippe, K.; Klenin, K. V.; Langowski, J. Looping dynamics of linear DNA molecules and the effect of DNA curvature: a study of Brownian dynamics simulation. *Biophys. J.* **1998**, *74*, 773–779.
- (37) Horowitz, D. S.; Wang, J. C. Torsional rigidity of DNA and length dependence of the free energy of DNA supercoiling. *J. Mol. Biol.* **1984**, *173*, 75–91.
- (38) Heath, P. J.; Clendenning, J. B.; Fujimoto, B. S.; Schurr, J. M. Effect of bending strain on the torsion elastic constant of DNA. *J. Mol. Biol.* **1996**, *260*, 718–730.
- (39) Shore, D.; Langowski, J.; Baldwin, R. L. DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4833–4837.
- (40) Shore, D.; Baldwin, R. L. Energetics of DNA twisting. I. Relation between twist and cyclization probability. *J. Mol. Biol.* **1983**, *170*, 957–981.
- (41) Shore, D.; Baldwin, R. L. Energetics of DNA twisting. II. Topoisomer analysis. *J. Mol. Biol.* **1983**, *170*, 983–1007.
- (42) Taylor, W. H.; Hagerman, P. J. Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. II. NaCl-dependence of DNA flexibility and helical repeat. *J. Mol. Biol.* **1990**, *212*, 363–376.
- (43) Trifonov, E. N. DNA in profile. *Trends Biochem. Sci.* **1991**, *16*, 467–470.
- (44) Hagerman, P. J. Straightening out the bends in curved DNA. *Biochim. Biophys. Acta* **1992**, *1131*, 125–132.
- (45) Anderson, J. N. Detection, sequence patterns and function of unusual DNA structures. *Nucleic Acids Res.* **1986**, *14*, 8513–8533.
- (46) Olson, W. K.; Marky, N. L.; Jernigan, R. L.; Zhurkin, V. B. Influence of fluctuations on DNA curvature. A comparison of flexible and static wedge models of intrinsically bent DNA. *J. Mol. Biol.* **1993**, *232*, 530–554.
- (47) Schellman, J. A. The flexibility of DNA. *Biopolymers* **1974**, *13*, 217–226.
- (48) Gorin, A. A.; Zhurkin, V. B.; Olson, W. K. B-DNA twisting correlates with base pair morphology. *J. Mol. Biol.* **1995**, *247*, 34–48.
- (49) Protozanova, E.; Yakovchuk, P.; Frank-Kamenetskii, M. D. Stacked-unstacked equilibrium at the nick site of DNA. *J. Mol. Biol.* **2005**, *342*, 775–785.
- (50) Travers, A. DNA dynamics: bubble 'n flip for DNA cyclization. *Curr. Biol.* **2005**, *15*, R377–379.
- (51) Olson, W. K.; Swigon, D.; Coleman, B. D. Implications of the dependence of the elastic properties of DNA on nucleotide sequence. *Philos. Trans. R. Soc.* **2004**, *362*, 1403–1422.
- (52) Calladine, C. R. Mechanics of sequence-dependent stacking of bases in B-DNA. *J. Mol. Biol.* **1982**, *161*, 343–352.
- (53) Ulyanov, N. B.; Zhurkin, V. B. Sequence-dependent anisotropic flexibility of B-DNA. A conformational study. *J. Biomol. Struct. Dynam.* **1984**, *2*, 361–385.
- (54) Sarai, A.; Mazur, J.; Nussinov, R.; Jernigan, R. L. Sequence dependence of DNA conformational flexibility. *Biochemistry* **1989**, *28*, 7842–7849.
- (55) Olson, W. K. The flexible DNA double helix. In *Stereodynamics of Molecular Systems*; Sarma, R. H., Ed.; Pergamon Press: New York, 1979; pp 297–314.

CT060025+