Sequence-Dependent Motions of DNA: A Normal Mode Analysis at the Base-Pair Level

Atsushi Matsumoto and Wilma K. Olson
Department of Chemistry, Rutgers, the State University of New Jersey, Wright-Rieman Laboratories, Piscataway, New Jersey 08854-8087 USA

ABSTRACT Computer simulation of the dynamic structure of DNA can be carried out at various levels of resolution. Detailed high resolution information about the motions of DNA is typically collected for the atoms in a few turns of double helix. At low resolution, by contrast, the sequence-dependence features of DNA are usually neglected and molecules with thousands of base pairs are treated as ideal elastic rods. The present normal mode analysis of DNA in terms of six base-pair "step" parameters per chain residue addresses the dynamic structure of the double helix at intermediate resolution, i.e., the mesoscopic level of a few hundred base pairs. Sequence-dependent effects are incorporated into the calculations by taking advantage of "knowledge-based" harmonic energy functions deduced from the mean values and dispersion of the base-pair "step" parameters in high-resolution DNA crystal structures. Spatial arrangements sampled along the dominant low frequency modes have end-to-end distances comparable to those of exact polymer models which incorporate all possible chain configurations. The normal mode analysis accounts for the overall bending, i.e., persistence length, of the double helix and shows how known discrepancies in the measured twisting constants of long DNA molecules could originate in the deformability of neighboring base-pair steps. The calculations also reveal how the natural coupling of local conformational variables affects the global motions of DNA. Successful correspondence of the computed stretching modulus with experimental data requires that the DNA base pairs be inclined with respect to the direction of stretching, with chain extension effected by low energy transverse motions that preserve the strong van der Waals' attractions of neighboring base-pair planes. The calculations further show how one can "engineer" the macroscopic properties of DNA in terms of dimer deformability so that polymers which are intrinsically straight in the equilibrium state exhibit the mesoscopic bending anisotropy essential to DNA curvature and loop formation.

INTRODUCTION

The large-scale fluctuations of DNA are key to understanding kinetically complicated events, such as the ease of the long, threadlike molecule snaking though the pores of a gel or closing into a loop between separately bound regulatory proteins. DNA loop formation is implicated, in turn, in a number of important biological processes, including the regulation of transcription (Bellomy et al., 1988; Schleif, 1992) and the organization of chromatin (Dillon et al., 1997; Bazett-Jones et al., 1999; Ringrose et al., 1999).

The motions of polymeric DNA are generally modeled in terms of a spatially homogeneous, naturally straight, elastic rod which ignores realistic features of chemical structure (reviewed in Olson, 1996). The fluctuations and correlations of structural parameters in crystals of pure DNA and of DNA-protein complexes, however, show that the equilibrium rest states and elastic constants of neighboring base pairs are dependent on sequence and further reveal the presence of strong couplings between modes of deformation—e.g., between bending, twisting, and stretching—that are the direct result of the chiral nature of DNA chemical structure (Olson et al., 1998). These local features become important as the deformations of sequential base-pair steps accumulate with the increase of chain length, introducing notable structural variability at the mesoscopic level, i.e., in chains of a few hundred base pairs.

Electron micrographs and other low resolution images of mesoscopic DNA fragments reveal broad distributions of molecular shapes and end-to-end distances (Muzard et al., 1990; Thresher and Griffith, 1992; Bednar et al., 1995; Lyubchenko et al., 1995; Hansma et al., 1996; Bustamante et al., 1997). The sequential features responsible for these spatial arrangements are often deduced by comparison of the experimentally observed images of a given DNA with the distribution of configurations generated by Metropolis-Monte Carlo sampling of the likely structural fluctuations of the constituent dimers (Hagerman, 1985; Levene and Crothers, 1986; Kahn and Crothers, 1998). Limitations on computer resources make routine comparison of simulations and experiment impractical, because hundreds of thousands of coordinate sets must be generated to characterize an individual polymer molecule. The study of the large-scale, collective motions of long DNA is consequently impeded.

The normal modes of DNA provide an alternative, computationally less demanding way to study the equilibrium properties of DNA. The normal modes are coupled vibrations found by assuming that the molecular potential energy can be approximated by a harmonic function of the configurational variables and by then solving a generalized eigenvector problem to give a closed analytical description of the motion. The eigenvalues give the vibrational time scales

Submitted July 23, 2001, and accepted for publication April 3, 2002.

Address reprint requests to Wilma K. Olson, Department of Chemistry, Rutgers, the State University of New Jersey, Wright-Rieman Laboratories, 610 Taylor Road, Piscataway, NJ 08854-8087; Tel.: 732-445-3993; Fax: 732-445-5958; E-mail: olson@rutchem.rutgers.edu.

0006-3495/02/07/22/20 $2.00

(frequencies) and the eigenvectors the details of the corresponding motion. The motion can be described and visualized at each frequency or as a superposition of independent modes. Time-averaged equilibrium and kinetic properties can be calculated accurately and efficiently as weighted sums, e.g., thermodynamic functions can be determined from the normal mode frequencies. The price paid for these benefits is the limited accuracy of the harmonic energy approximation. The method, nevertheless, provides a useful first impression of the flexibility of a molecule and shows how the motions may change when the chemical structure is modified or the chain is perturbed by the binding of proteins and other ligands.

Normal mode or vibrational analysis is a well established and much used technique for the study of both small molecules and proteins: the technique is successful in reproducing the vibrational spectra of small molecules (Wilson et al., 1955) and in analyzing the collective motions involved in the folding of protein fragments and in the binding and release of substrates and products to and from enzymes (Levy and Karplus, 1979; Brooks and Karplus, 1983; Kitao and Gō, 1999; Berendsen and Hayward, 2000). Analyses of DNA normal modes to date have focused on the collective motions of very short chain fragments ranging from small oligomers (Tidor et al., 1983; Irikura et al., 1985; Garcia and Soumpasis, 1989; Kottalam and Case, 1990; Ha Duong and Zakrzewska, 1997a,b, 1998; Lin et al., 1997) to a few turns of double helix (Matsumoto and Gō, 1999). The size of the latter systems is limited by the kinds of parameters which have been used to describe three-dimensional structure—typically the Cartesian coordinates of the constituent atoms or the dihedral angles of a nucleic acid fragment with rigid chemical bonds and valence angles. The small, almost imperceptible motions of short DNA oligonucleotides described at this high level of resolution are closely tied to the detailed sequence-dependent fine structure of the double helix and the association of water, drugs, proteins, and other molecules with individual nucleotide residues. The conformational effects which are collectively responsible for the large-scale, sequence-dependent properties of polymeric DNA chains are difficult to discern in such studies. The normal modes of infinitely long DNA, however, have been examined by Prohofsky and associates (Hua and Prohofsky, 1988; Chen and Prohofsky, 1995) by repeating a short fragment of DNA and taking advantage of helical symmetry to reduce the number of independent coordinates. The complexity of chains that can be studied from this perspective is again limited by the length of the repeating unit.

In the present investigation, we examine the motions of long fragments of DNA by taking advantage of "knowledge-based" harmonic energy functions deduced from known high-resolution crystal structures of DNA (Olson et al., 1998). The deformations of individual dimers are described by six independent "step" parameters which specify the spatial arrangements of neighboring base pairs: three angular variables called Tilt, Roll, and Twist and three variables called Shift, Slide, and Rise with dimensions of distance (Dickerson et al., 1989). This rigid-body representation of base pairs significantly reduces the number of independent variables per chain molecule, thereby making it possible to study the normal modes of much longer DNA fragments than could heretofore be treated in either Cartesian or dihedral angle space. In addition, the elastic energy incorporates the sequence-dependent anisotropy of DNA bending as well as the known correlations of base-pair step parameters. The only missing structural information is the detailed conformation of the sugar-phosphate backbone, including the charged phosphate groups. The latter atoms and the surrounding aqueous solvent, such as the water molecules and counterions in the first or second solvation layer around the DNA, are implicitly treated in the energy terms so that their omission introduces no serious errors when duplex deformations are limited to energies of the order of k_BT (where k_B is the Boltzmann constant and T the temperature in Kelvin). It should be noted, however, that the surrounding solvent does not have a viscosity and that we do not consider the damping effect of solvent on large-scale polymeric properties in this study. This low-resolution model, nevertheless, provides a straightforward way to deduce the effects of sequence on global folding, something that is said (Berendsen and Hayward, 2000) still to elude analyses of the slow conformational changes of proteins. In principle, there is no limit on the length of DNA that can be described in this manner so long as the constituent dimers obey the harmonic energy model, i.e., there are no excursions of the molecule from the classical B-form double helical structure. Longer chains are characterized by a greater number of normal modes, which when appropriately weighted and superimposed, give rise to a broader distribution of global molecular configurations. In practice, chain length is restricted by the form of the kinetic energy, which assumes that changes in atomic coordinates brought about by the fluctuations of individual "step" parameters are small. These upper limits preclude the need to treat the long-range electrostatic self-repulsion that influences the folding of long supercoiled DNA molecules (Fenley et al., 1994; Vologodskii and Cozzarelli, 1995; Westcott et al., 1997).

We focus attention here on the normal modes of regularly repeating, linear polymers free of bound ligands to establish a point of reference for studies to be reported elsewhere of arbitrary DNA sequences and of chains with ends held in place by specific anchoring conditions, such as the looped configurations imposed by protein binding. We compare our simplified energy model with previous all-atom treatments of short oligomers and with ideal elastic rod representations of polymeric DNA. We identify the local base-pair fluctuations responsible for the dominant (lowest frequency) normal modes. The derived polymeric properties account satisfactorily for the overall bending, i.e., persistence length, of the double helix and reveal the critical role

of sequence in global twisting. The computations further suggest that the constants impeding the large-scale stretching of single DNA molecules are related to the concerted bending, twisting, and lateral displacements of neighboring base-pair planes as well as to their axial separation. We identify particular sequence contexts under which ideal elastic behavior breaks down, concentrating on conditions that induce the mesoscopic bending anisotropy which underlies DNA loop formation. Finally, by comparing the end-to-end dimensions of the simulated double helices with exact values obtained from standard matrix formulations of polymer configurational statistics (Flory, 1969; Maroun and Olson, 1988; Marky and Olson, 1994), we investigate the chain lengths at which the normal mode analysis of DNA modeled in terms of base-pair steps is valid.

METHODS

Base-pair representation

The atoms of each base pair lie in the plane of a local, orthogonal coordinate frame (**x**, **y**, **z**), which is defined in accordance with recently established guidelines (Olson et al., 2001). (Note: The **x** and **y** axes lie in the plane of the base pair, with **x** pointing in the direction of the major groove along the pseudodyad axis of the base pair and **y** running along the long axis of the base pair in the direction of the leading (sequence) strand, parallel to the C1′ · · · C1′ vector, and displaced so as to pass through the intersection of **x** with the vector connecting the pyrimidine Y(C6) and purine R(C8) atoms. The **z** axis is perpendicular to the base-pair plane, pointing in the 5′-3′ direction of the leading strand.) The coordinate system of the kth base pair is further characterized by the position of its origin $\mathbf{o}_k$ and the rotation matrix $\mathbf{R}_k$, which relates the local base-pair frame to a fixed, orthonormal global reference frame. The rotation matrix is given by $[\boldsymbol{\gamma}_x\boldsymbol{\gamma}_y\boldsymbol{\gamma}_z]$, where the $\boldsymbol{\gamma}_\nu(\nu = x, y, z)$ are unit vectors, expressed in columns, along the positive coordinate axes of the base-pair frame.

The relative position and orientation of the $(k + 1)$th base pair with respect to its predecessor k are given respectively by the difference between coordinate origins, $\mathbf{v}_{k,k+1} = (\mathbf{o}_{k+1} - \mathbf{o}_k)$ and the product of the rotation matrix $\mathbf{R}_{k+1}$ and the inverse of $\mathbf{R}_k$, i.e., the dimer transformation matrix $\mathbf{T}_{k,k+1} = \mathbf{R}_{k+1}(\mathbf{R}_k)^{-1}$. The components of the projection of $\mathbf{v}_{k,k+1}$ on the coordinate axes of base pair k, (τ_x, τ_y, τ_z), are used as translational parameters in the present calculation.

The matrix **T** can also be expressed in terms of the rotation of magnitude φ around the axis **u** which brings the coordinate frames on successive base pairs into coincidence, where the components of the unit vector **u** are (u_x, u_y, u_z) and the angle φ equals $(\varphi_x^2 + \varphi_y^2 + \varphi_z^2)^{1/2}$ (Chasles, 1830). The elements of **T**, so expressed, are given by (Jeffreys and Jeffreys, 1946):

$$t_{\nu\mu} = (1 - \cos\varphi)u_\nu u_\mu - \sin\varphi \sum_\kappa \epsilon_{\nu\mu\kappa}u_\kappa + \cos\varphi\ \delta_{\nu\mu}, \quad (1)$$

where $\delta_{\nu\mu}$ is the Kronecker delta, i.e., $\delta_{\nu\mu} = 1$ when $\nu = \mu$, $\delta_{\nu\mu} = 0$ when $\nu \neq \mu$, and $\epsilon_{\nu\mu\kappa} = \pm 1$ when ν, μ, κ is an even or an odd permutation of 1, 2, 3, respectively, and vanishes otherwise. The rotational components $(\varphi_x, \varphi_y, \varphi_z)$ are equated to (Tilt, Roll, Twist) after the definition of Babcock et al. (1994).

The translational parameters between base-pair planes are generally expressed in terms of a "middle" frame so that numerical values are independent of the direction from which a DNA structure is analyzed (Lu and Olson, 1998), i.e., either the sequence or the complementary strand. The translational parameters defined in this work in terms of the kth coordinate frame are related to Shift, Slide, Rise values in the literature, e.g., (Olson et al., 1998), through the matrix $\mathbf{T}^{1/2}$ that effects the halfway rotation $(\varphi/2)$ between neighboring base-pair planes:

$$\begin{bmatrix}\tau_x\\ \tau_y\\ \tau_z\end{bmatrix} = \mathbf{T}^{1/2}\begin{bmatrix}\text{Shift}\\ \text{Slide}\\ \text{Rise}\end{bmatrix}. \quad (2)$$

The dependence of the $\tau_\nu(\nu = x, y, z)$ on Tilt, Roll, Twist (through their incorporation in **T**) as well as on Shift, Slide, Rise precludes description of the local conformational energy in terms of a quadratic expression in the six base-pair step parameters. To overcome this limitation, we express the rotation matrix in Eq. 2 in terms of the equilibrium "rest" values of Tilt, Roll, and Twist, i.e., $\mathbf{T}^{1/2} \approx \mathbf{T}_o{}^{1/2}$. The latter approximation is valid so long as the dimer structure does not depart significantly from this minimum energy state.

Normal mode analysis

Our treatment of the normal modes of DNA at the level of "step" parameters builds upon general formulations (Noguti and Gō, 1983; Braun et al., 1984; Higo et al., 1985; Levitt et al., 1985) previously developed to describe the collective motions of proteins in terms of internal chemical coordinates, i.e., dihedral angles. The conformational potential energy V of a double helix of N base pairs is thus approximated by the multi-dimensional parabola,

$$V = \frac{1}{2}\sum_{ij} f_{ij}\Delta\theta_i\Delta\theta_j, \quad (3)$$

where $\Delta\theta_i$ is the instantaneous fluctuation of the ith "step" parameter from its equilibrium value and f_{ij} is an element of a $6(N - 1) \times 6(N - 1)$ matrix of elastic force constants, **F**, i.e., the terms describing the potential deformability of the six parameters in each of the $N - 1$ base-pair steps.

The kinetic energy K is similarly expressed in quadratic form in terms of $\dot{\theta}_i$, the first derivative of θ_i with respect to

time, and the weighted "mass" coefficients h_{ij} of the $6(N-1) \times 6(N-1)$ matrix $\mathbf{H}$ (defined below):

$$K = \frac{1}{2} \sum_{ij} h_{ij} \Delta\dot{\theta}_i \Delta\dot{\theta}_j. \quad (4)$$

If the $\Delta\theta_i$ and $\Delta\ddot{\theta}_i$ are collected respectively in the vectors $\mathbf{\Theta}$ and $\ddot{\mathbf{\Theta}}$, the equations of motion simplify to:

$$\mathbf{H}\Delta\ddot{\mathbf{\Theta}} + \mathbf{F}\Delta\mathbf{\Theta} = \mathbf{0}, \quad (5)$$

with a periodic solution of the form:

$$\Delta\theta_i = \sum_{n=1}^{6(N-1)} A_{in}\alpha_n \cos(\omega_n t + \delta_n). \quad (6)$$

Here ω_n is the frequency, δ_n the phase angle, α_n the amplitude of the nth normal mode, and A_{in} the fractional contribution to the ith "step" parameter from the nth normal mode. (In practice the components of $\Delta\mathbf{\Theta}$ and the frequencies ω_n are obtained by introducing the vector of normal coordinates $\mathbf{Q}$ defined by $\Delta\mathbf{\Theta} = \mathbf{AQ}$, where the elements of $\mathbf{Q}$ are $\alpha_n \cos(\omega_n t + \delta_n)$ and $\mathbf{A}$ is a $6(N-1) \times 6(N-1)$ transformation matrix. Eq. 5 can then be rewritten as the eigenvalue expression, $\mathbf{HA\Lambda} = \mathbf{FA}$, with $\mathbf{A}^{\mathrm{T}}\mathbf{HA}$ equated to the identity matrix of order $6(N-1)$ and the normal mode frequencies and directions found from the diagonalization of $\mathbf{F}$, *i.e.*, $\mathbf{\Lambda} = \mathbf{A}^{\mathrm{T}}\mathbf{FA}$, where $\mathbf{A}^{\mathrm{T}}$ denotes the transpose of $\mathbf{A}$ and $\mathbf{\Lambda}$ is a diagonal matrix with elements $\lambda_{nn} = \omega_n^2$. The phase angles, however, cannot be determined with this procedure.) The fluctuations of each "step" parameter $\Delta\theta_i$ at time t are thereby expressed as a linear combination of harmonic oscillators, the energies of which are proportional to ω_n^2. The contribution of each mode to the variation of individual "step" parameters decreases rapidly with increase in ω_n, so that relatively few (low frequency) modes are responsible for the large-scale deformations of the molecule.

Kinetic energy

The kinetic energy coefficients in Eq. 4 incorporate the mass m_a and the Cartesian coordinates $\mathbf{r}_a$ of individual atoms in the DNA through the relationship:

$$h_{ij} = \sum_a m_a \left(\frac{\partial \mathbf{r}_a}{\partial \theta_i}\right)\left(\frac{\partial \mathbf{r}_a}{\partial \theta_j}\right). \quad (7)$$

To evaluate the h_{ij}, we take advantage of analytical expressions for the $(\partial\mathbf{r}_a/\partial\theta_i)$ developed to treat the normal mode dynamics of a system of two molecules, each of which moves in dihedral angle space (Braun et al., 1984; Higo et al., 1985). The set of rigid body parameters used here to relate neighboring base-pair planes is identical in form to the variables previously used to describe the relative global positions and orientations of different molecules. The internal atomic motions of DNA are effectively separated from the overall rotations and translations of the molecule by expressing the atomic displacements in terms of an embedded coordinate frame, chosen according to the so-called Eckart condition (Eckart, 1935) to minimize the mass-weighted square atomic displacement of DNA (McLachlan, 1979) in its equilibrium rest state and in a state where one of the "step" parameters is altered from its minimum energy value (Noguti and Gō, 1983).

Because the atomic motions $\Delta\mathbf{r}_a$ brought about by "step" parameter change $\Delta\theta_i$ are assumed to be small in this common frame (Eckart, 1935), the kinetic energy term (Eq. 4) effectively restricts the length of DNA which can be studied by normal mode analysis. This limitation does not apply to the mesoscopic chains lengths considered below, where typical room temperature fluctuations of individual base-pair steps, such as the $\pm 5°$ perturbations of Roll which raise the potential energy by $\sim k_{\mathrm{B}}T/2$, limit atomic movement to 0.02–0.07 persistence lengths. The persistence length of mixed sequence DNA under ambient aqueous salt conditions, by comparison, is the same magnitude as the contour length of a 150 to 200-bp chain (Hagerman, 1988; Smith et al., 1992, 1996; Bustamante et al., 1994; Bednar et al., 1995; Baumann et al., 1997). The limitation on chain length is further discussed below in the analysis of DNA end-to-end dimensions.

Both backbone and base atoms must be included in the evaluation of the h_{ij}. The sugar-phosphate backbone is incorporated in the present calculations by the superposition of a canonical B-form 5′-nucleotide helical fragment (Chandrasekaran and Arnott, 1989) in the reference frame of each base. Because complementary base-pair parameters are fixed at ideal planar values, each complementary nucleotide pair is thereby treated as a rigid body and the small variations in backbone conformation that accompany fluctuations in the geometry of neighboring base pairs are ignored.

Force field

The potential energy coefficients in Eq. 3 are taken from knowledge-based elastic functions extracted from the three-dimensional arrangements of DNA base-pair steps in protein-DNA crystal structures (Olson et al., 1998):

$$V_{\mathrm{XY}} = \frac{1}{2} \sum_{i=1}^{6} \sum_{j=1}^{6} g_{ij} \Delta\phi_i \Delta\phi_j. \quad (8)$$

The g_{ij} in this expression are the force constants impeding deformations of the XpY dimer, and the $\Delta\phi_i$ correspond to the instantaneous fluctuations of each of the six "step" parameters of that dimer from the B-DNA rest state, i.e., $\Delta\phi_1 = \Delta\mathrm{Tilt}_{\mathrm{XY}}$, $\Delta\phi_2 = \Delta\mathrm{Roll}_{\mathrm{XY}}$, $\Delta\phi_3 = \Delta\mathrm{Twist}_{\mathrm{XY}}$, $\Delta\phi_4 = \Delta\mathrm{Shift}_{\mathrm{XY}}$, $\Delta\phi_5 = \Delta\mathrm{Slide}_{\mathrm{XY}}$, $\Delta\phi_6 = \Delta\mathrm{Rise}_{\mathrm{XY}}$ (full details of the potential function are available in Table S-1 in the Supplementary Material). If these elements are collected

respectively in the 6×6 force constant matrix $\mathbf{G}_{XY}$ and the 6×1 vector $\Delta\Phi_{XY}$, the dimer deformation energy V_{XY} reduces to $(1/2)\Delta\mathbf{\Phi}_{XY}{}^{T}\mathbf{G}_{XY}\Delta\mathbf{\Phi}_{XY}$. The choice of translational parameters and the aforementioned approximation of $\mathbf{T}$ in Eq. 2 by the constant matrix $\mathbf{T}_o$ introduce a linear relationship between the $\Delta\mathbf{\Phi}_{XY}$ and the corresponding elements of the longer $6(N-1)$ vector $\Delta\mathbf{\Theta}$ used in Eq. 5. That is, $\Delta\mathbf{\Phi}_{XY} = \mathbf{C}_{XY}\Delta\mathbf{\Theta}_{XY}$, where the molecular deformation vector $\Delta\mathbf{\Theta}$ (Eq. 5) is built up from the $(N-1)$ sets of constituent $\Delta\mathbf{\Theta}_{XY}$ dimer fluctuations and $\mathbf{C}_{XY}$ is a 6×6 constant matrix characteristic of the mean orientation of the XY dimer in B-form DNA. The potential energy of the DNA as a whole can be expressed in the form of Eq. 3 by constructing the $6(N-1)$ pseudodiagonal matrix of force constants $\mathbf{F}$ from the $\mathbf{F}_{XY} = \mathbf{C}_{XY}{}^{T}\mathbf{G}_{XY}\mathbf{C}_{XY}$ corresponding to the DNA base-pair sequence. For example, the f_{ij} of the self-complementary $(CGTACG)_2$ hexamer duplex would be collected in the 30×30 array,

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{CG} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{GT} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_{TA} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{F}_{AC} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{F}_{CG} \end{bmatrix}, \qquad (9)$$

where the boldface $\mathbf{0}$ values in this expression are 6×6 null matrices.

Because the lowest energy conformation of a given DNA sequence is self-evident from the knowledge-based force field, i.e., the minimum energy three-dimensional structure is defined by the set of average "step" parameters, the time-consuming energy minimization step normally carried out before normal mode analysis does not need to be performed in the present calculations.

The potential functions introduced below are first approximations of the sequence-dependent structure of DNA based on the observed conformations of selected dimers in high resolution protein-bound and B-DNA crystal structures. Although mean (equilibrium) sequence-dependent parameters have not significantly changed since first reports (Gorin et al., 1995; Olson et al., 1998), the local force constants are expected to change as new structural data accumulate. Moreover, there are not yet sufficient crystallographic data to address DNA deformability beyond the dimer level despite suggestions (Nadeau and Crothers, 1989) of longer-range organization of DNA structure.

End-to-end dimensions

The unperturbed, mean-square end-to-end distance of DNA, $\langle r^2\rangle_0$, is obtained by three independent approaches. First, exact values of $\langle r^2\rangle_0$ are computed using standard matrix methods for the mean extension of a DNA of N base pairs, with $\langle r^2\rangle_0$ equated to the average scalar product of the end-to-end vector, $\langle \mathbf{r}\cdot\mathbf{r}\rangle_0$, and $\mathbf{r}$ defined as

$$\mathbf{r} = \mathbf{v}_1 + \mathbf{T}_{12}\mathbf{v}_2 + \mathbf{T}_{12}\mathbf{T}_{23}\mathbf{v}_3 + \cdots + \mathbf{T}_{12}\mathbf{T}_{23}\cdots\mathbf{T}_{N-2,N-1}\mathbf{v}_{N-1}. \qquad (10)$$

Here $\mathbf{T}_{k,k+1}$ is the transformation matrix that relates coordinate frames in successive base-pair planes (Eq. 1) and $\mathbf{v}_k$ is the virtual bond vector that connects base-pair centers (Eq. 2). Boltzmann averages of the $\mathbf{T}$ and $\mathbf{v}$ at each base-pair step are incorporated in dimeric generator matrices, the serial product of which yields all possible configurations of the selected chain sequence. The reader is referred elsewhere for the precise formulation of these expressions (Flory, 1969; Maroun and Olson, 1988; Marky and Olson, 1994). The zero subscript on $\langle r^2\rangle_0$ and other (unperturbed) averages denotes the omission of chain self-intersection and solvent-polymer interactions in their calculation (Flory, 1969). The DNAs studied below are sufficiently short and stiff so that excluded volume effects are automatically avoided.

Second, we express $\langle r^2\rangle_0$ in terms of the fluctuations $\Delta\mathbf{r}$ of the end-to-end vector with respect to the position $\mathbf{r}^o$ of terminal base pairs in the (minimum energy) rest state,

$$\begin{aligned}\langle r^2\rangle_0 &= \langle \mathbf{r}\cdot\mathbf{r}\rangle_0 \\ &= \langle(\mathbf{r}^o + \Delta\mathbf{r})\cdot(\mathbf{r}^o + \Delta\mathbf{r})\rangle_0 \\ &= \mathbf{r}^o\cdot\mathbf{r}^o + 2\mathbf{r}^o\cdot\langle\Delta\mathbf{r}\rangle_0 + \langle\Delta\mathbf{r}\cdot\Delta\mathbf{r}\rangle_0. \end{aligned} \qquad (11)$$

Here the superscript ^{o} denotes (constant) rest state values. The value of $\Delta\mathbf{r}$, which is determined by the fluctuations of "step" parameters, is written as a Taylor series expansion in the $\Delta\theta_i$,

$$\Delta\mathbf{r} = \sum_{i=1}^{6(N-1)}\left(\frac{\partial r}{\partial\theta_i}\right)\Delta\theta_i + \frac{1}{2}\sum_{i=1}^{6(N-1)}\sum_{j=1}^{6(N-1)}\left(\frac{\partial^2\mathbf{r}}{\partial\theta_i\partial\theta_j}\right)\Delta\theta_i\Delta\theta_j + \cdots. \qquad (12)$$

If the series is approximated by the first term, the average displacement vector $\langle\Delta\mathbf{r}\rangle_0$ is zero because $\langle\Delta\theta_i\rangle = 0$ and the mean-square end-to-end distance is greater than the separation of chain ends in the rest state (because the self product, $\langle\Delta\mathbf{r}\cdot\Delta\mathbf{r}\rangle_0$, in Eq. 11 must be positive). Because such a result is inconsistent with the end-to-end displacement of a naturally straight, inextensible DNA, we consider up to second-order terms of the Taylor series (Eq. 12) in the calculation with the result,

$$\begin{aligned}\langle\Delta\mathbf{r}\rangle_0 &= \frac{1}{2}\sum_{i=1}^{6(N-1)}\sum_{j=1}^{6(N-1)}\left(\frac{\partial^2\mathbf{r}}{\partial\theta_i\partial\theta_j}\right)\langle\Delta\theta_i\Delta\theta_j\rangle, \\ \langle\Delta\mathbf{r}\cdot\Delta\mathbf{r}\rangle_0 &= \sum_{i=1}^{6(N-1)}\sum_{j=1}^{6(N-1)}\left(\frac{\partial\mathbf{r}}{\partial\theta_i}\right)\cdot\left(\frac{\partial\mathbf{r}}{\partial\theta_j}\right)\langle\Delta\theta_i\Delta\theta_j\rangle. \end{aligned} \qquad (13)$$

The first and second derivatives in these expressions are evaluated numerically. The values of $\langle\Delta\theta_i\Delta\theta_j\rangle$ are taken

directly from the experimental data used to derive the knowledge-based force field (Olson et al., 1998). Such averages would have to be determined numerically over all normal modes if an all-atom force field were used. It should also be noted that the first derivatives that appear in Eq. 13, $(\partial \mathbf{r}/\partial \theta_i)$, differ from similar terms employed in Eq. 7 to compute the kinetic energy. As explained above, the latter derivatives are computed in a carefully chosen coordinate frame that essentially eliminates rotational and translation motions of the molecule. The variation of chain ends required to evaluate Eq. 13, by contrast, is performed in a common global coordinate frame embedded in the plane of the first base pair.

Finally, we compute the root-mean-square end-to-end distance of DNA molecules which are confined to the set of spatial configurations associated with one or more normal modes. That is, we use the set of "step" parameters associated with mode n at time t (Eq. 6) to generate the end-to-end vector $\mathbf{r}(n, t)$ and the Boltzmann factor $\sigma(n, t)$ of each of the polymeric states that comprise certain global motions and then evaluate $\langle r^2 \rangle_0$ as an energy weighted sum over representative configurations:

$$\langle r^2 \rangle_0 = \frac{\sum_{n\,t} \mathbf{r}(n, t) \cdot \mathbf{r}(n, t)\sigma(n, t)}{\sum_{n\,t} \sigma(n, t)}. \qquad (14)$$

In practice, we set an upper limit of 7.5 $k_B T$ on the potential energy, $V(\Delta\theta_i(n, t))$ (Eq. 3), thereby excluding configurations where $\sigma(n, t) = \exp[-V/k_B T] \leq \exp[-7.5]$.

RESULTS AND DISCUSSION

Elastic features of DNA repeating polymers

We start by determining the normal modes of four simple repeating polymers—the poly dA · poly dT and poly dG · poly dC homopolymers with a monomeric repeating unit (the AA · TT or GG · CC dimer) and the poly d(AT) and poly d(GC) alternating copolymers made up of sequential purine-pyrimidine (AT · AT or GC · GC) and pyrimidine-purine (TA · TA or CG · CG) base-pair steps. We classify the motions, after Matsumoto and Gō (1999), on the basis of the global structural changes revealed through computer visualization of the eigenvectors associated with each mode (i.e., Eq. 6). The color-coded spectrum of lowest frequency bending, twisting, and stretching modes of a 200 bp chain is presented in Fig. 1 for each of the four polymers and schematic illustrations of some of these motions are given in Fig. 2. Numerical values of the frequencies are tabulated in the Supplementary Material (Table S-2). We focus on the lowest frequency modes because the amplitudes of the normal modes and their consequent importance to the overall motions of DNA diminish at higher frequency.

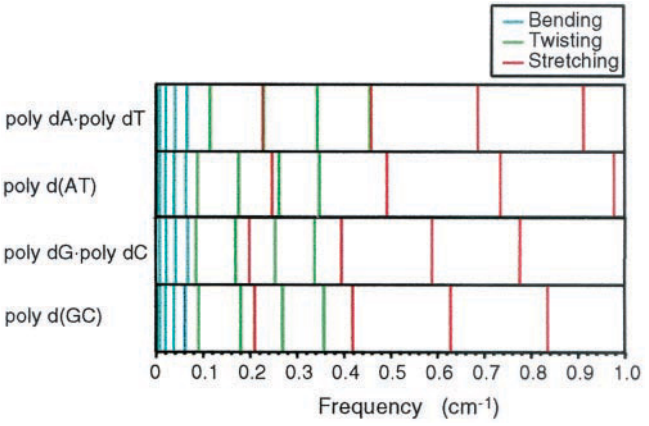

FIGURE 1 Color-coded spectrum of lowest frequency bending (*blue*), twisting (*green*), and stretching (*red*) modes of 200-bp segments of the poly dA · poly dT and poly dG · poly dC homopolymers and the poly d(AT) and poly d(GC) alternating copolymers. As the frequency increases, the amplitude of each normal mode and its consequent importance to the overall motions of the DNA becomes lower.

It should be noted that these low frequency modes would not be observed in a real system due to viscous effects, which we do not consider here. While our method does not necessarily describe the time-development of DNA chains in a real system, we show below that these low normal mode frequencies are closely related to the global dynamic properties of DNA, such as the bending rigidity, twisting rigidity, and stretching rigidity. Thus, the low normal mode frequencies presented here are useful indicators of the dynamic properties of DNA.

Each of the pure (planar) bending modes (blue lines at the left of Fig. 1) is doubly degenerate, describing mutually perpendicular motions with equivalent (bending) energy. The superposition of these modes leads to the global isotropic bending characteristic of an ideal elastic rod. In contrast to the bending motions, the twisting modes (green lines) are higher in frequency, more widely spaced, and more sensitive to polymer composition. Specifically, the frequencies of the twisting modes of the simulated alternat-

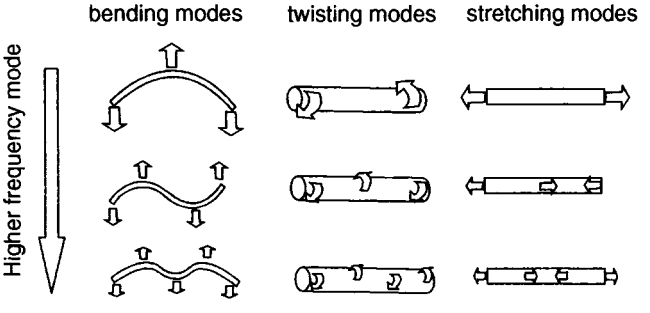

FIGURE 2 Schematic illustration of representative low frequency normal modes of an elastic rod. The arrows point in the directions of bending, twisting, and stretching motions in each mode.

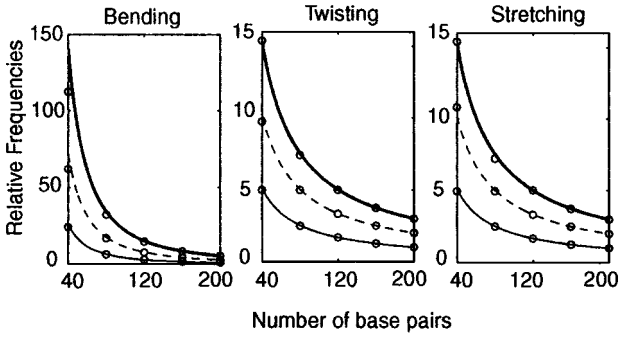

FIGURE 3 Superposition of the computed dependence of representative low frequency bending, twisting, stretching modes of poly dA · poly dT (scatter points) on chain length and the predicted variation (*thin solid lines* for the lowest frequencies, *dashed lines* for the second lowest frequencies, and *thick solid lines* for the third lowest frequencies) for an elastic rod with the same contour length. Frequencies are normalized with respect to the lowest frequency of each type of motion in a 200 bp-chain.

ing copolymers and the poly dG · poly dC homopolymer are lower in value than those of the poly dA · poly dT model. The stretching modes (red lines), which are also sensitive to base-pair sequence, tend to be even more widely spaced and to occur at still higher frequencies than the twisting modes.

The rod-like behavior of poly dA · poly dT is further illustrated in Fig. 3 where the chain length dependence of the computed low frequency modes is superimposed on the predicted variation (Strutt, 1894) of the bending, twisting, and stretching frequencies of a homogeneous elastic rod. (The frequency of the nth bending mode of an elastic rod of length L with uniform circular cross-section is given by $\nu_n^b = C_b p_n^2$, where C_b is a material constant and p_n satisfies the condition $p_n L = (4.730, 7.853, 10.9965, \cdots)$ for $(n = 1, 2, 3, \ldots)$ (Strutt, 1894). The twisting and stretching frequencies follow a similar form, varying directly with n and inversely with L, *i.e.*, $\nu_n^t = (n/2L)C_t$ and $\nu_n^s = (n/2L)C_s$.) The frequencies of the poly dA · poly dT motions (scatter points) are normalized with respect to the lowest frequency bending, twisting, and stretching modes of a 200-bp chain. The normal mode frequencies of the elastic rod (smooth curves) are similarly expressed as multiples of the lowest frequency motions of an ideal rod with the same contour length (199 base-pair steps × 3.27 Å per step = 650 Å). The agreement between computed and theoretical values is remarkably close, with only minor discrepancies in the relative bending frequencies at short chain lengths where the theory is known to break down. Similar agreement between the base-pair level representation of DNA motions and elastic rod behavior is obtained for the low frequencies modes of poly dG · poly dC and the two alternating copolymers (data not shown).

The mechanical constants associated with the overall elastic behavior of the four repeating polymers are collected in Table 1. These values are obtained by substituting the computed normal mode frequencies in the classical expressions for the normal modes of an ideal elastic rod (Strutt, 1894). Specifically, the bending constant A, the twisting constant C, and the stretching constant Y are given by

$$A = \frac{(2\pi\nu_n^b)^2 M}{p_n^4 L}, \quad C = \frac{(2L\nu_n^t)^2 I_M}{n^2}, \quad Y = \frac{(2\nu_n^s)^2 ML}{n^2}, \tag{15}$$

where the ν_n are the computed frequencies of the nth bending (b), twisting (t), or stretching (s) mode, M is the total mass of the molecule, L is the length of the DNA helical axis (see Kosikov et al., 1999, and Matsumoto and Gō, 1999 for computational details) and I_M is the moment of inertia per unit length around the twisting axis. As evident from the table, the computed values of the bending constant A and the persistence length a derived from A on the basis of elastic rod theory ($a = A/k_B T$ with $T = 298$ K) are approximately independent of sequence and in excellent agreement with values extracted from experimental studies of mixed sequence DNA (Hagerman, 1988; Smith et al., 1992, 1996; Bustamante et al., 1994; Bednar et al., 1995; Baumann et al., 1997). The elastic constants, however, may change as the local force constants are refined.

The sensitivity of the calculated twisting constants to base-pair sequence may account in part for the wide range of experimentally derived values of C (Barkley and Zimm, 1979; Thomas et al., 1980; Hurley et al., 1982; Millar et al., 1982; Horowitz and Wang, 1984). The differences in mea-

TABLE 1 Mechanical constants describing the simulated elastic properties of four regularly repeating DNA polymers

Molecule	A (10^{-19} erg-cm) Bending	a (Å) Persistence length*	C (10^{-19} erg-cm) Twisting	Y (10^3 pN) Stretching
poly dA · poly dT	2.3	547	2.9	2.5
poly d(AT)	2.2	536	1.8	3.0
poly dG · poly dC	2.7	647	1.7	2.0
poly d(GC)	2.2	531	1.8	2.2
DNA†	2.0	500	1.5–4	1.0–1.4

*Persistence length of an ideal elastic rod ($a = A/k_B T$ with $T = 298$ K).

†Experimental values for mixed sequence DNA obtained from the following sources: bending constant (Barkley and Zimm, 1979; Millar et al., 1982); persistence length (Hagerman, 1988); twisting constant (Barkley and Zimm, 1979; Thomas et al., 1980; Hurley et al., 1982; Millar et al., 1982; Horowitz and Wang, 1984; Hagerman, 1988); stretching constant (Smith et al., 1996; Baumann et al., 1997; Wang et al., 1997; Bouchiat and Mezard, 1998).

sured twisting constants, usually attributed to experimental effects, such as the bending strain that accompanies loop closure (Heath et al., 1996), may also reflect the intrinsic deformability of the constituent base pairs. The greater overall twisting stiffness of poly dA · poly dT compared to the other polymers in Table 1 stems in part from the substantially higher cost of distorting AA dimers compared to other base-pair steps in the knowledge-based potential (see Table S-1 in the Supplementary Material and discussion below).

The stretching constants Y extracted here, however, are two- to threefold greater than values deduced from recent single-molecule manipulations of DNA (Smith et al., 1996; Baumann et al., 1997; Wang et al., 1997; Bouchiat and Mezard, 1998). As discussed below, the relative ease of global stretching reflects a number of local factors, in addition to the force constants that govern the axial, i.e., van der Waals', separation of neighboring base-pair planes. In other words, the global stretching of DNA is not simply a function of Rise.

Matsumoto and Gō (1999) previously derived DNA elastic constants on the basis of a normal mode analysis in the dihedral angle space of short (24–36 bp) double helical molecules, with poorer correspondence between computation and experiment than the present work, i.e., the previously computed bending and twisting constants are respectively three to ten times greater than the corresponding values in Table 1. The improved agreement found here seemingly reflects our use of knowledge-based potential energies rather than an atomic force field. Optimization of DNA conformation on the basis of the detailed interactions of all atoms typically yields a large number of closely spaced minimum energy substates of similar structural character (Poncin et al., 1992). The normal mode analysis of DNA in dihedral angle (and Cartesian) space is performed with respect to one of these minimum energy states and the transitions between closely related, competing minima are not considered. The statistical approach used to generate our elastic force field, by contrast, reflects a large number of known crystal structures, each of which is analogous to one of the minimum energy substates found through all-atom energy minimization. Therefore, any fluctuation on our elastic energy surface is comparable to a transition between different minimum energy points derived with an atomic force field. The correspondence between the mechanical constants in Table 1 and values extracted from experiment suggests that transitions between conformational substates have a significant influence on the global properties of DNA.

A similar argument is made by Song and Schurr (1990) in accounting for the unusually large values of DNA persistence length (~2100 Å) deduced from measurements of transient electric dichroism. They offer three possible reasons for the large discrepancy between their results and the persistence lengths for mixed sequence DNA (~500 Å) obtained with other experimental approaches. One explanation, which they favor, is that the potential energy of DNA bending is not a smooth quadratic function and that the energy exhibits several discrete minima separated by barriers. In the short-time scale of their experiment the DNA is trapped in one of the minima and the persistence length thereby becomes larger.

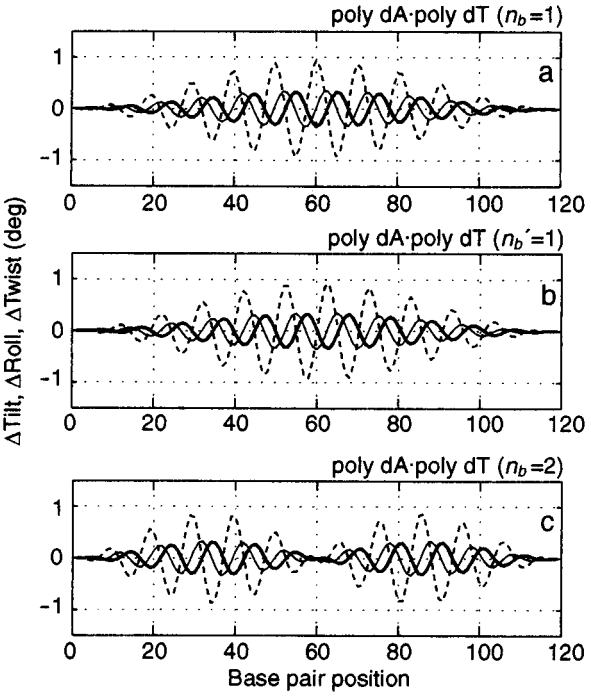

FIGURE 4 Deformations of local angular "step" parameters, which are collectively responsible for the lowest frequency in-plane bending motions of a 120-bp poly dA · poly dT duplex: (*top, middle*) degenerate modes of lowest frequency, $n_b = n_b' = 1$; (*bottom*) one of the two second lowest frequency modes, $n_b = 2$. Plots illustrate the sequential fluctuations of Tilt (*thin solid line*), Roll (*dashed line*), and Twist (*thick solid line*) at the moment when the potential energy of the molecule is raised by $k_BT/2$; fluctuations are reversed a half cycle later of the mode.

It should be noted that there is an open controversy over the ionic strength dependence of the persistence length and the lower limit of the persistence length of mixed sequence DNA at ambient salt conditions. We cite a DNA persistence length of 500 Å in Table 1 consistent with a variety of independent measurements. As noted above, larger values of the persistence length are deduced in some work, e.g., (Song and Schurr, 1990). There is similar disagreement among elastic constants obtained for specific DNA sequences (see Table S-3 in the Supplementary Material) which may be related in part to the discussion above, *i.e.*, chains may sample only a small part of the energy surface in some experiments but may explore vast areas of conformation space in others.

Sequence effects on global motion

Figs. 4–6 illustrate the deformations of local "step" parameters, which are collectively responsible for selected low

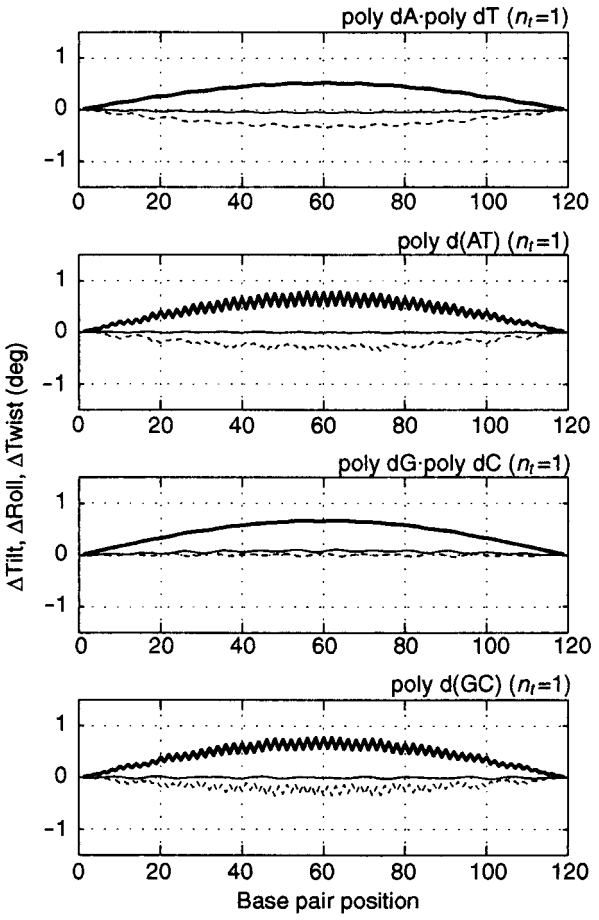

FIGURE 5 Fluctuations of local angular "step" parameters, which are collectively responsible for the lowest frequency global twisting modes (n_t = 1) of 120-bp fragments of poly dA · poly dT, poly d(AT), poly dG · poly dC, and poly d(GC). See legend to Fig. 4.

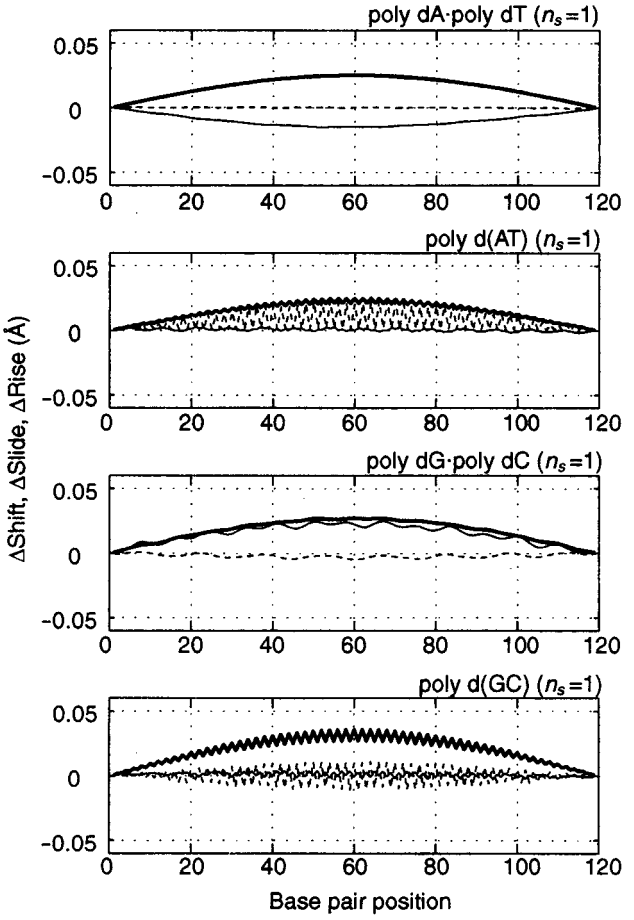

FIGURE 6 Fluctuations of local translational "step" parameters associated with the lowest frequency global stretching (n_s = 1) of 120-bp fragments of poly dA · poly dT, poly d(AT), poly dG · poly dC, and poly d(GC). Plots illustrate the sequential fluctuations of Shift (*thin solid line*), Slide (*dashed line*), and Rise (*thick solid line*) at the moment when the potential energy of the molecule is raised by $k_BT/2$. See legend to Fig. 4.

frequency normal modes of the aforementioned repeating polymers. These plots, for DNA chains of 120 bp, capture the sequential fluctuations of each conformational variable at the moment when the potential energy of the molecule is raised by $k_BT/2$; the fluctuations of all parameters are reversed a half cycle later of the mode. As detailed below, the mechanisms used to effect overall chain motions reflect the conformational properties of the constituent dimer units.

Bending

The in-plane bending of poly dA · poly dT, detailed in Fig. 4, takes advantage of the local bending anisotropy of the double helix, i.e., the tendency of DNA to deform through rolling rather than tilting motions and an intrinsic feature of the knowledge-based AA · TT potential. The same patterns of angular distortions are found in the corresponding modes of the other three polymers and thus are not reported here. The regular patterns of Roll deformations in the illustrated examples (dashed lines) resemble early "mini-kinked" models of global DNA bending (Olson, 1979; Yathindra, 1979; Zhurkin et al., 1979) with marked angular distortions of opposite sense at ~5 bp intervals, corresponding to sites on the surface of the double helix where the cumulative effects of local structural change combine constructively. The lesser, out-of-phase variations of Tilt (thin solid lines) contribute in the same fashion to global bending, much like the sequentially shifted sinusoidal changes in Roll and Tilt characteristic of classical models of "smooth" bending (Namoradze et al., 1977; Levitt, 1978; Sussman and Trifonov, 1978). In contrast to the uniformity of the ideal models, the local bending deformations revealed by the normal mode analysis are variable with Roll and Tilt angles of greater or lesser magnitude in different parts of the polymeric fragment, e.g., the distortions responsible for the lowest frequency bending mode are concentrated around the midpoint of the chain. The small accompanying fluctuations

of Twist in the present study (thick solid lines) reflect the well known correlation between Roll and Twist built into the energy model.

The global bending isotropy brought about by degenerate bending modes arises at the local level from a simple 2- to 3-bp phase shift of conformational distortions (compare Figs. 4*a*, *b*). A comparable shift of "step" parameters also underlies the two normal modes that dominate the global bending of short RNA duplexes (Zacharias and Sklenar, 2000). These subtle changes follow from the ideas behind DNA "mini-kinking," where the sites of largest Roll (and concomitantly zero Tilt) deformation determine the direction of overall curvature (Olson, 1979; Yathindra, 1979; Zhurkin et al., 1979). A 90° change in the direction of global bending comes about by shifting the pattern of local structural perturbations so that Roll is null and Tilt becomes a local maximum or minimum at corresponding sites in the degenerate mode.

The DNA makes use of the same conformational mechanism to change the direction of global bending in the second lowest frequency bending mode. The point of inflection in the variation of "step" parameters versus poly dA · poly dT base-pair position in Fig. 4 *c* gives rise to a 180° shift of bending directions in the two halves of the molecule (see schematic in Fig. 2). The change in direction stems from the asymmetric pattern of Roll variation in the 120-bp chain, e.g., the distortions are local maxima near base pairs 10, 20, 30, 40, 50 in the first half of the molecule and local minima near base pairs 70, 80, 90, 100, 110 in the second half. Similar representations of the higher frequency bending modes (data not shown) contain additional points of inflection where all three angular variables are undistorted. If the inflection points become too closely spaced, the ~10 bp pattern of conformational fluctuations in phase with the helical repeating unit disappears.

The critical frequency at which the distance between adjacent points of inflection drops to 10 bp can be estimated as follows. Roughly speaking, there are $(2n + 1)/4$ standing waves in the nth bending mode (Fig. 2). The wavelength of the mode is thus given by $\lambda = 4L/(2n + 1)$, where L is the contour length of the DNA. The upper limit on the number of bending modes n is then obtained by setting the half wavelength corresponding to the distance between two adjacent points of inflection to 10 bp, i.e., $\lambda/2 = 10$ bp. The maximum n for a 120-bp stretch of poly dA · poly dT is thus 12, corresponding to a bending frequency of ~1.4 cm^{-1} in an elastic rod with equivalent material properties and contour length. In the case of the poly dA · poly dT 120-mer, we find a pair of bending modes with close frequencies below the theoretical frequency limit at 1.07 and 1.08 cm^{-1} and with 11 points of inflection, separated by ~10 bp, in each of these modes. As anticipated from elastic theory, there are no other such pairs of degenerate bending modes at higher frequency.

The coupling of local "step" parameters in the DNA force field gives rise to changes in neighboring base-pair displacement as the chain bends globally (see Fig. S-1 in the Supplementary Material for plots of the local translational fluctuations associated with the bending of individual polymers). The deformations in Roll which dominate the overall molecular motion are accompanied by changes in Slide, whereas the lesser changes in Tilt introduce concomitant fluctuations of Shift. The extremes of deformability found in pyrimidine-purine (YR) versus purine-pyrimidine (RY) steps, i.e., the enhanced flexibility of YR steps versus the natural stiffness of RY steps (Olson et al., 1998), introduces a zig-zag pattern in the variation of "step" parameters in poly d(AT) and poly d(GC) models compared to the smooth parametric responses of the simulated homopolymers in Fig. 4 (see plots of translational parameters for the alternating copolymers in Fig. S-1 in the Supplementary Material). For example, the increments in Roll that lead to global bending are consistently greater at the YR steps of alternating copolymers (data not shown), and the coupled sequence-dependent variation of base-pair displacement in these chains is even more pronounced, with distortions in Slide restricted almost entirely to the YR steps.

Twisting

As evident from Fig. 5, the lowest frequency global twisting motions of regular DNA polymers are dominated by fluctuations in local Twist (thick solid lines). The dimer deformations build up from the ends of the molecule with the largest deformations of neighboring base-pair steps found in the center of the chain. The alternating copolymers show the zig-zag pattern of local conformational change noted above with greater distortions occurring at YR steps. The degree of local Twist deformation of poly dA · poly dT is smaller than that in the other chains as expected from its higher global twisting constant (Table 1). The increase in Twist angles in Fig. 5 is accompanied by a concomitant decrease in Roll values (dashed lines), the decrease in local bending depending on the strength of the energy coupling term, e.g., small for GG compared to AA and YR steps. The translational parameters show similar sequence-dependent, coupled variations that reflect the knowledge-based energy function. For example, fluctuations of Shift are negligible in the alternating copolymers where the Twist-Shift energy coefficient is by definition zero. (The directional dependence of Shift (Dickerson et al., 1989) leads naturally to null force constants at self-complementary dimer steps. Indeed, the coefficients of all mixed energy terms involving Shift or Tilt (Eq. 8), except for the Tilt-Shift contribution, are zero at such steps.) Similarly, Slide varies in opposing directions at YR versus RY steps in poly d(GC) in accordance with the positive Twist-Slide coefficient at CG steps and the negative value at GC steps. Graphs of the sequential variation of translational parameters in the lowest frequency twisting

modes of all four polymers are available in Fig. S-2 in the Supplementary Material.

Stretching

The global stretching of DNA homopolymers and alternating copolymers stems primarily from changes in Rise that grow to a maximum in the middle of the chain (thick solid lines in Fig. 6). The computed, sequence-dependent variation in Rise in the lowest frequency stretching mode reflects the local force constants, with CG steps much more easily deformed than any other dimer. The changes in Rise are coupled, via the elastic energy potential, to fluctuations in Shift in the homopolymers (thin solid lines in Fig. 6) and to Slide in the copolymers (dashed lines in the figure). Indeed, the variation of Shift is comparable to that of Rise at GG steps as is the change in Slide relative to that of Rise at TA steps. Poly d(GC) is unusual in stretching with an alternation in the direction of Slide at YR and RY steps. As discussed below, the coupling of Rise to Shift or Slide contributes to the overall stretching force constants. Concerted movements of base pairs which mimic the conformational behavior of stretched DNA fibers (Chandrasekaran and Arnott, 1989), e.g., same sign variation of Shift and Rise, enhance the global stretching motion. Notably, deformations of angular parameters in the lowest frequency global stretching mode (shown in Fig. S-3 in the Supplementary Material) are an order of magnitude smaller than the changes in Tilt, Roll, and Twist which effect global bending and twisting.

Influence of conformational coupling on large-scale motions

DNA polymers with the "realistic" conformational features of the knowledge-based potential show very different patterns of deformation from ideal elastic rod models. As outlined above, the sequence-dependent fluctuations of base-pair step parameters in the present calculations stem from built-in conformational correlations, which are expressed by the g_{ij} cross terms in the knowledge-based force field (Eq. 8). To make this point clear, we repeated the preceding normal mode analyses of DNA homopolymers and alternating copolymers with all cross terms of the energy function set to zero. The results confirm expectations.

Coupling effects in regular DNA polymers

First of all, the sequence-dependent fluctuations of secondary "step" parameters involved in global bending, e.g., Twist (Fig. 4), become smaller when the cross terms are set to zero. That is, the fluctuations of "step" parameters other than Tilt and Roll become negligible in all four repeating polymers when subjected to these computational conditions. Similarly, the fluctuations of base-pair step parameters other than those of Twist become small and irregular in the global twisting modes (Fig. 5) if cross terms are omitted. Sequence-dependent effects on global twisting are thereby diminished.

In the case of stretching, the situation is slightly different. The neglect of cross terms has only a limited effect on local conformational fluctuations in the lowest frequency stretching modes of the four repeating polymers. The variation of "step" parameters in the "modified" polymers without cross terms roughly resembles the plots reported for the intact force field in Fig. 6. For example, sizable fluctuations of Shift persist in the homopolymer chains in the absence of conformational coupling and Slide assumes nonzero values in the alternating copolymers under the same conditions.

To gain further insight into this behavior, we performed the normal mode analysis of an even more primitive DNA polymer, using the knowledge-based force constants but setting all cross terms (g_{ij}) in the potential function to zero and choosing an ideal, naturally straight B-DNA equilibrium rest state with $\text{Tilt}_o = \text{Roll}_o = 0°$, $\text{Twist}_o = 36°$, $\text{Shift}_o = \text{Slide}_o = 0$ Å, $\text{Rise}_o = 3.4$ Å. In this case, the fluctuations of all parameters other than Rise are very small, suggesting that the variation of "step" parameters which gives rise to the overall stretching of DNA may be influenced by intrinsic molecular conformation (see analysis and additional discussion below).

Coupling effects in ideal mixed sequence DNA

We have also considered the influence of cross terms on the normal modes of a simplified computational model introduced some years ago to account at the base-pair level for the observed persistence length of mixed sequence DNA (Olson et al., 1993). In this scheme, there is no coupling of "step" parameters, i.e., $g_{ij} = 0$, and the elastic force constants of individual conformational variables are chosen so that the potential energy increases by $k_BT/2$ if any one variable deviates from its equilibrium rest state by the root-mean-square fluctuation (standard deviation), i.e., $g_{ii} = k_BT/\langle\Delta\theta_i^2\rangle$. We used this model in combination with the crystallographically observed mean parameters and dispersion of a generic DNA base-pair step (Olson et al., 1998) as the basis for the series of normal mode analyses described below, i.e., $\langle\text{Tilt}\rangle = 0 \pm 3.6°$, $\langle\text{Roll}\rangle = 2.7 \pm 5.2°$, $\langle\text{Twist}\rangle = 34.2 \pm 5.5°$, $\langle\text{Shift}\rangle = 0 \pm 0.64$ Å, $\langle\text{Slide}\rangle = -0.1 \pm 0.69$ Å, $\langle\text{Rise}\rangle = 3.4 \pm 0.25$ Å. Elastic constants derived from the normal mode analysis of this generic homopolymer ($A = 2.3 \times 10^{-19}$ erg-cm, $C = 1.5 \times 10^{-19}$ erg-cm, $Y = 1.9 \times 10^3$ pN) are comparable to experimental observations (see data for mixed sequence DNA in Table 1). This correspondence between computation and experiment is the basis for the choice of temperature in all other normal mode calculations in this paper.

TABLE 2 Effect of local conformational coupling on the lowest normal mode frequencies of global bending, twisting, and stretching of a random sequence 120-bp DNA homopolymer*

Coupled variables	$\langle \Delta\theta_i \Delta\theta_j \rangle$	Bending	Twisting	Stretching
—	—	0.02030	0.13582	0.33929
Tilt-Roll	−9.36	0.02030	0.13582	0.33788
Tilt-Twist	9.90	0.02030	0.13572	0.33923
Tilt-Shift	1.15	0.02030	0.13583	0.33935
Tilt-Slide	−1.24	0.02030	0.13582	0.34084
Tilt-Rise	0.45	0.02030	0.13582	0.33255
Roll-Twist	−14.30	0.01975	0.14085	0.34109
Roll-Shift	−1.66	0.02029	0.13582	0.33935
Roll-Slide	−1.79	0.02029	0.13583	0.33218
Roll-Rise	−0.65	0.02030	0.13581	0.31055
Twist-Shift	1.76	0.02030	0.13575	0.33930
Twist-Slide	1.90	0.02030	0.13547	0.34113
Twist-Rise	0.69	0.02030	0.13300	0.39067
Shift-Slide	−0.22	0.02030	0.13582	0.33928
Shift-Rise	−0.08	0.02030	0.13582	0.33891
Slide-Rise	−0.09	0.02030	0.13582	0.38148

*Frequencies reported in cm^{-1}. Force field derived from covariance matrix with diagonal elements given by the root-mean square fluctuation ($\Delta\theta_i^2$) of "step" parameters and off-diagonal elements $\langle \Delta\theta_i \Delta\theta_j \rangle = \langle \Delta\theta_j \Delta\theta_i \rangle$ with magnitude $(1/2)(\langle \Delta\theta_i^2 \rangle \langle \Delta\theta_j^2 \rangle)^{1/2}$ and sign consistent with crystallographically observed correlations. Mean values and root-mean-square fluctuations of dimers based on equal weighting of the "step" parameters in the 16 common dimers from protein-bound DNA crystal structures (Olson et al., 1998): ⟨Tilt⟩ = 0 ± 3.6°, ⟨Roll⟩ = 2.7 ± 5.2°, ⟨Twist⟩ = 34.2 ± 5.5°, ⟨Shift⟩ = 0 ± 0.64 Å, ⟨Slide⟩ = −0.1 ± 0.69 Å, ⟨Rise⟩ = 3.4 ± 0.25 Å.

To test of the influence of conformational coupling on DNA polymeric properties, we examined the effects of nonzero mixed potential energy terms, taken one at a time, on the normal modes of the generic homopolymer. We estimate the magnitude of the covariance $\langle \Delta\theta_i \Delta\theta_j \rangle$ of an arbitrary pair of "step" parameters from the observed dispersion, i.e., $\langle \Delta\theta_i \Delta\theta_j \rangle = (1/2)(\langle \Delta\theta_i^2 \rangle \langle \Delta\theta_j^2 \rangle)^{1/2}$, and base the sign of the correlation on the trends in DNA crystal data. We find the g_{ij} of Eq. 8 by inversion of the covariance matrix which contains these two off-diagonal elements, i.e., $\langle \Delta\theta_i \Delta\theta_j \rangle = \langle \Delta\theta_j \Delta\theta_i \rangle \neq 0$, and diagonal terms related to be $\langle \Delta\theta_i^2 \rangle$.

Values of the lowest frequency bending, twisting, and stretching motions derived from the normal mode analysis of a 120-bp generic homopolymer with all possible sets of local conformational coupling are reported in Table 2, along with the corresponding frequencies of the ideal, mixed-sequence chain. According to these data, the normal mode frequencies of the ideal DNA are relatively unperturbed by the coupling of "step" parameters. However, even a small change in normal mode frequency reveals dimeric features important to overall polymer motion. For example, the frequency of in-plane bending changes if Roll is coupled to other "step" parameters. Thus, Roll is important to global bending. Likewise, Twist is critical to large-scale twisting. The situation with stretching is more complicated. Parameters other than Rise are critical to overall chain extension and compression, notably Roll and Slide, which respectively determine the global inclination and displacement of base pairs (Saenger, 1984; Kosikov et al., 1999) (see additional discussion below).

The greatest change of normal mode frequencies in Table 2 occurs when Twist and Rise are correlated. The introduction of this local conformational interaction lowers the overall twisting frequency and raises the global stretching frequency with parallel changes in the magnitudes of the associated elastic constants. That is, it is easier to twist the homopolymer if Rise varies in concert with Twist but more difficult to stretch the molecule as a whole under the same conditions. As noted above, fluctuations of base-pair Twist dominate the overall twisting of regular DNA polymers and deformations of Rise influence global stretching. It is therefore reasonable that coupled changes in Twist and Rise introduce correlations between molecular twisting and stretching as well as alter the frequencies and elastic constants of these global motions. Such local features may underlie the apparent coupling of global stretching and twisting deduced from single molecule elasticity experiments (Kamien et al., 1997; Marko, 1997).

The data in Table 2 similarly show that the coupling of Tilt, Roll, Shift, or Slide with Rise alters the stretching frequency of the generic homopolymer, with more significant changes associated with Roll and Slide. Recent atomic-level models of DNA extension-compression find comparable concerted changes in Twist, Roll, and Slide along simulated low energy stretching pathways (Kosikov et al., 1999). As discussed in the following section, these locally coupled distortions facilitate overall DNA stretching.

Although the introduction of cross-terms has relatively limited effects on the magnitudes of normal mode frequencies (Table 2), the local pathways that bring about these global changes are sensitive to the assigned conformational correlations. In the absence of cross-terms in the potential function, the bending motions of the generic homopolymer arise exclusively from fluctuations in Tilt and Roll. Other step parameters contribute to global bending only if they are coupled to these two primary variables. For example, the correlation between Twist and Roll characteristic of most base-pair steps induces changes in Twist during global bending of the generic homopolymer with some regions of the molecule under-twisted and other parts concomitantly over-twisted. Analogous conformational effects accompany the global twisting and stretching motions.

Implications for single molecule stretching

The sensitivity of the global stretching frequencies to the local dimeric potential provides new insights into the conformational response of DNA to direct micromanipulation. Global stretching of DNA exclusively through changes in Rise is only possible in an ideal, naturally straight molecule

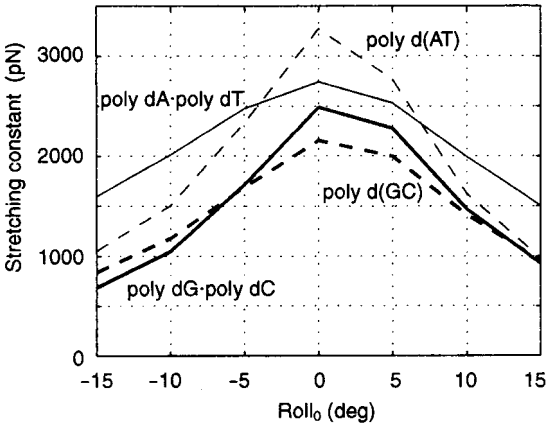

FIGURE 7 Variation of the mechanical stretching constant as a function of the equilibrium Roll angle for an ideal, naturally straight DNA chain with the force constants of different polymer sequences: poly dA · poly dT (*thin solid line*), poly dG · poly dC (*thick solid line*), poly d(AT) (*thin dashed line*), poly d(GC) (*thick dashed line*). The remaining equilibrium "step" parameters are held fixed at the following values in all cases: Tilt$_o$ = 0°, Twist$_o$ = 36°, Shift$_o$ = Slide$_o$ = 0 Å, Rise$_o$ = 3.4 Å.

where the base-pair planes remain perpendicular to the global helical axis. Inasmuch as Rise is the energetically stiffest "step" parameter, other lower energy dimer deformations will come into play in the extension and compression of a real DNA molecule, even in the absence of cross-terms in the potential energy function. The stretching constant is expected to be greatest for an ideal, naturally straight DNA and to become smaller as the equilibrium structure deviates from this reference.

To confirm this expectation, we carried out normal mode analyses on a series of 120-bp DNA chains with perturbed equilibrium rest states. Starting from the ideal homopolymer with Tilt$_o$ = Roll$_o$ = 0°, Twist$_o$ = 36°, Shift$_o$ = Slide$_o$ = 0 Å, Rise$_o$ = 3.4 Å at all base-pair steps, we introduced identical nonzero equilibrium Roll angles in every dimer and computed the stretching constant Y from the normal modes of that DNA (Eq. 15). Four different kinds of polymers were considered with "step" parameters assigned the force constants of poly dA · poly dT, poly dG · poly dC, poly d(AT), and poly d(GC) sequences. Because the choice of Roll$_o$ alters helical contour length as well as normal mode frequencies, the changes of mechanical constants in Fig. 7 reflect the slight compression of the double helical axis brought about by the assumed nonzero equilibrium Roll, e.g., ~4% reduction in contour length for Roll$_o$ = ±10°, as well as the natural global stretching motions (see Kosikov et al., 1999 for analytical expressions that relate regular DNA helical structure to the constituent "step" geometry). The major contributions to Y in Fig. 7, however, arise from the change in normal mode frequencies.

It is worth noting that the base pairs of DNA structures with nonzero equilibrium Roll are inclined with respect to

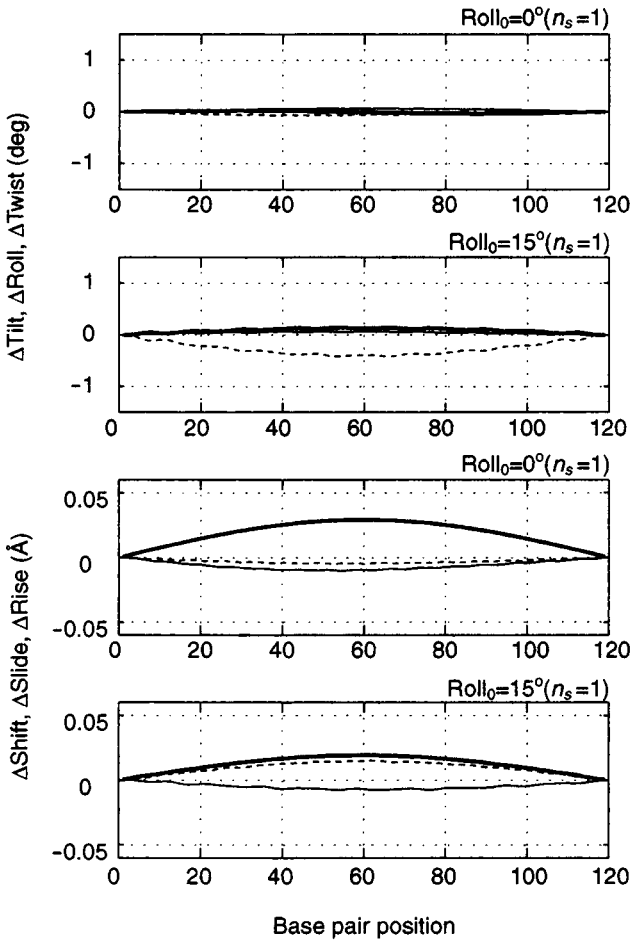

FIGURE 8 Comparative plots of the sequential fluctuations of local "step" parameters associated with the lowest frequency global stretching modes (n_s = 1) of a naturally straight 120-bp "poly dA · poly dT" chain with Roll$_o$ = 0° or 15°. Other chain parameters are identical to those listed in the legend to Fig. 7. Shift (*thin solid line*); Slide (*dashed line*); Rise (*thick solid line*).

the helical axis. The displacement of neighboring base pairs in the direction of inclination, i.e., variation of Slide, adds naturally to the global stretching. Because Slide is a much "softer" conformational variable than Rise, global stretching is energetically enhanced with increase (or decrease) of Roll$_o$. The mechanical stretching constant in turn decreases, attaining values comparable to those extracted from experiment for Roll$_o$ = ±10–15° (Fig. 7). In other words, when the base-pair planes are perpendicular to the helical axis, along which the stretching motion takes place, energetically costly deformations via Rise result in a large elastic constant. When the equilibrium Roll angle deviates from zero, stretching takes place by lower energy conformational routes with a concomitant decrease in the mechanical constant. For example, the change in Roll$_o$ from 0° to 15° in a polymer with poly dA · poly dT force constants allows both Slide and Roll to contribute to overall chain extension and compression (Fig. 8). A similar enhancement of global

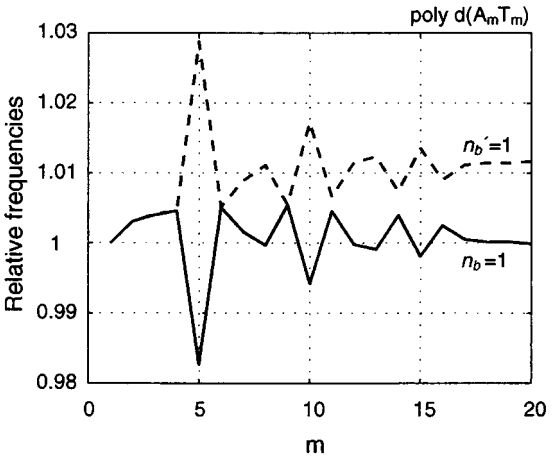

FIGURE 9 The first (n_b = 1) and second (n_b' = 1) lowest global bending frequencies (*solid* and *dashed lines*, respectively) of a series of intrinsically straight poly d(A_mT_m) chains of 120-bp plotted as a function of m. Frequencies are normalized with respect to the lowest bending frequency of the poly d(AT) alternating copolymer (m = 1).

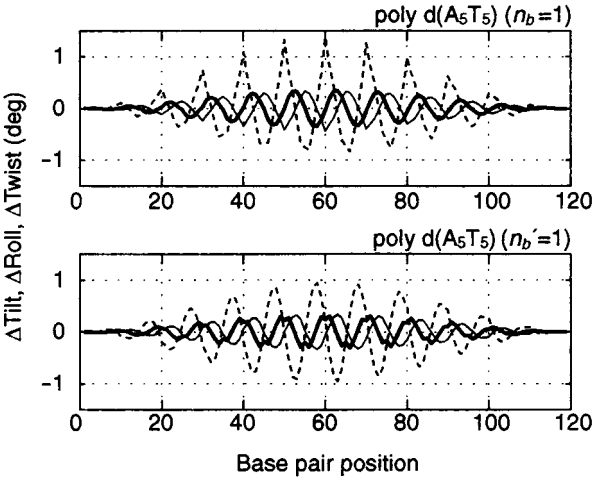

FIGURE 10 Fluctuations of local angular "step" parameters which underlie the global bending anisotropy of the poly d(A_5T_5) block copolymer. (*Top*) lowest frequency mode (n_b = 1); (*bottom*) second lowest mode (n_b' = 1). See legend to Fig. 4.

stretching and a comparable local conformational response are obtained if the equilibrium Slide is assigned nonequilibrium values in a corresponding series of normal mode analyses (data not shown).

We rationalize the lower stretching constants of DNA extracted from recent experiments (Smith et al., 1996; Baumann et al., 1997; Wang et al., 1997; Bouchiat and Mezard, 1998) compared to values obtained from normal mode analysis (Table 1) along the same lines. The regular polymers studied here are very similar to a homopolymer with $Roll_o = 0°$. On the other hand, the base pairs of the mixed sequence DNA used in single-molecule stretching experiments may not lie perpendicular to the direction of stretching, or may reorient upon forced extension (via concerted changes in Roll, Twist, and Slide) as shown in model systems (Kosikov et al., 1999). It therefore seems natural that the experiments yield smaller stretching constants than the present computations.

Mesoscopic anisotropy

The interplay between local and large-scale conformational motions revealed in the preceding sections suggests new ways in which one might "engineer" the macroscopic properties of DNA in terms of dimer deformability. Such sequence-dependent modeling differs from conventional static representations of the folding of specific DNA sequences at 0 K (see Olson and Zhurkin, 1996 for an overview on DNA bending). For example, one can take advantage of the enhanced deformability of pyrimidine-purine steps (Olson et al., 1998) in combination with the periodic fluctuations of Roll angles which underlie the low frequency bending modes of DNA to design regular polymers which are intrinsically straight in the equilibrium state (at 0 K) yet which bend in a preferred direction at ambient temperatures.

We illustrate this mesoscopic anisotropy in Fig. 9 with the two lowest frequency bending modes of a series of intrinsically straight poly d(A_mT_m) sequences where m is half the length of the repeating unit, i.e., m = 1 corresponds to the alternating copolymer discussed above with two base pairs per repeating unit and m = 5 to the poly d(A_5T_5) block copolymer with a 10 base-pair sequential repeat. To emphasize the influence of sequence-dependent deformability on global properties, we assign an ideal B-DNA rest state to all dimers, i.e., $Tilt_o = Roll_o = 0°$, $Twist_o = 36°$, $Shift_o = Slide_o = 0$ Å, $Rise_o = 3.4$ Å, in place of the crystallographically based AA, AT, and TA averages (which generate a very slightly curved equilibrium structure). We consider chain fragments of 120 bp and sequences where m = 1–20. Because m is not necessarily an even multiple of the assumed chain length, most of the chains in the analysis contain a nonintegral number of sequential repeats.

As evident from the figure, the degeneracy of bending frequencies that is characteristic of large-scale isotropic bending breaks down for particular sequential repeats, namely at m = 5, 7, 8, 10–20 where the frequencies are no longer equivalent. The DNA in these cases bends more easily in the direction of the lower frequency mode than in the roughly perpendicular direction of the higher frequency mode. The effects are expected to be even more pronounced if the DNA is intrinsically bent at 0 K or if it is modified by the regular binding of drugs or other ligands. The apparent isotropic rod behavior of sequences other than the alternating copolymer, i.e., chains with m = 2–4, 6, and 9, may also be affected by such factors.

The conformational mechanism which underlies the bending anisotropy of poly d(A_5T_5), the sequence in Fig. 9

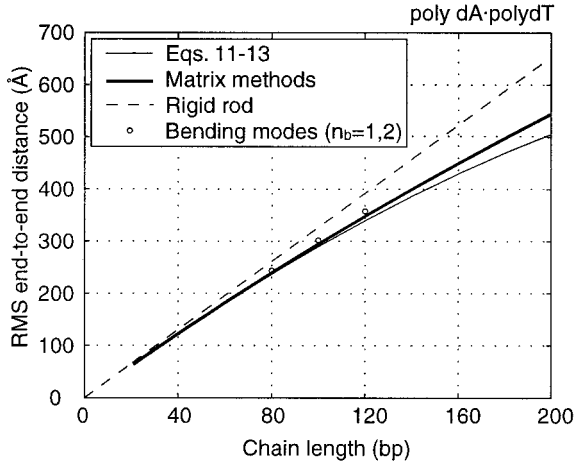

FIGURE 11 Computed variation of the unperturbed root-mean-square end-to-end distance of poly dA · poly dT, $\sqrt{\langle r^2\rangle_0}$, with chain length N. Values obtained with the Taylor series approximation of Eqs. 11–13 (*thin solid line*) are compared against exact averages obtained with direct matrix generator approaches (Flory, 1969; Maroun and Olson, 1988; Marky and Olson, 1994) (*thick solid line*) and the contour length L of the rigid-rod equilibrium state (*dashed line*). The values obtained by Eq. 14, in which two pairs of the degenerate bending modes are considered, are also plotted by open circles.

with greatest propensity for directional bending, is revealed in Fig. 10, where the local angular distortions associated with a total energy increase of $k_BT/2$ are reported along the chain contour for the two lowest frequency (energy) bending modes. The lower energy mode takes advantage of the repetition of the highly deformable TA dimer in phase with the 10-fold double helical repeat. The cost of global bending via large (positive) changes in Roll is reduced compared to that in the higher energy mode where the corresponding extremes of Roll occur at TT steps. The slightly higher cost of concomitantly bending AT versus AA steps (via negative Roll) into the minor groove is not sufficient to compensate for the energetic advantage of TA over TT bending in the two modes.

End-to-end dimensions

The normal mode analysis of global DNA motions in terms of elastic "step" parameters is comparable, although not equivalent, to the approximation of DNA configuration by a Taylor series expansion of the displacement of the end-to-end vector about the (minimum energy) rest state with terms up to second order in $\Delta\theta_i$ considered (Eq. 12). The extent to which this representation of molecular shape matches the exact dimensions computed by enumeration of all possible chain configurations (Flory, 1969; Maroun and Olson, 1988; Marky and Olson, 1994) provides insight into the range of chain lengths over which the present normal mode approximation is valid. The end-to-end distances of poly dA · poly dT computed by these two approaches match closely in chains of 20–200 bp (Fig. 11), the discrepancy at the upper limit of chain length being less than 7%. This very close agreement suggests that the normal mode analysis of DNA at the level of base-pair "step" parameters is valid in chains of this length.

Surprisingly, the DNA is more extended, with values closer to the contour length of a rigid rod, if all chain configurations are enumerated in the evaluation of $\langle r^2\rangle_0$. For example, the root-mean-square end-to-end separation of a 100-bp chain determined with the exact matrix generator approach is 298 Å versus a value of 291 Å obtained with Eqs. 11–13 and an equilibrium contour length of 324 Å. Interestingly, if the motions of the polymer are limited to the degenerate in-plane global bending modes of lowest frequency ($n_b = 1$), $\sqrt{\langle r^2\rangle_0}$ is 306 Å, and if additional pairs of low frequency bending modes are considered in Eq. 14, the computed value drops to 302 Å ($n_b = 1, 2$) and 300 Å ($n_b = 1$–3). The close correspondence of the latter values with the average extension of all possible chain configurations (i.e., ~1% overestimation of the root-mean-square end-to-end distance compared to the exact value) illustrates the dominance of the low frequency bending modes in determining the overall shape of DNA. These modes produce the largest overall motions at the lowest cost in energy.

The incorporation of global twisting or stretching in the enumeration of $\langle r^2\rangle_0$ adds only slightly to polymeric extension with the end-to-end distance of the 100 bp poly dA · poly dT chain increasing to 301 Å (from 300 Å) when the molecule samples, in addition to the first three pairs of global bending modes, the lowest frequency twisting ($n_t = 1$) and stretching ($n_s = 1$) modes. As anticipated from the approximation of the kinetic energy at the level of "step" parameters (Eq. 7), the correspondence between normal mode estimates of $\sqrt{\langle r^2\rangle_0}$ and exactly enumerated values improves at shorter chain lengths and grows worse in longer chains. The estimated end-to-end distances of 80 and 120 bp poly dA · poly dT chains which sample the two lowest frequency in-plane bending modes ($n_b = 1, 2$) overestimate the exact averages of 242 and 351 Å by 0.8% and 2.0%, respectively.

Finally, direct enumeration of the persistence length of poly dA · poly dT with matrix generating methods yields a limiting persistence length of 554 Å (170 bp), in excellent agreement with the persistence length of 547 Å extracted with the classical elastic expression ($a = A/k_BT$ with $T =$ 298 K) from the independently computed global bending constant of the same polymer (Table 1). The correspondence between these values lends further support to the representation of DNA equilibrium properties through a base-pair level normal mode analysis.

CONCLUSIONS

Polymer motions from a base-pair level perspective

Although normal mode analysis has been used to study DNA dynamics for a number of years, it is only now

possible, with the base-pair level perspective introduced in this work, to examine the vibrations of mesoscopic pieces of DNA. The simplification of double helical structure to six base-pair "step" parameters per chain residue reduces the number of independent variables compared to the number considered in traditional, higher resolution normal mode studies (Tidor et al., 1983; Irikura et al., 1985; Garcia and Soumpasis, 1989; Kottalam and Case, 1990; Ha Duong and Zakrzewska, 1997a,b, 1998; Lin et al., 1997; Matsumoto and Gō, 1999). The use of an elastic energy function further simplifies the calculations by eliminating the costly energy minimization step that usually precedes normal mode analysis. The treatment of DNA chains of ~100 bp can consequently be performed in a fraction of the time that it takes to analyze the normal modes of 2–3 turns of double helix in dihedral angle (or Cartesian coordinate) space. Although many fine structural details are lost in the present simulations, such as the correspondence of calculated values with experimentally observed Raman frequencies of bond stretching and bending, the present computations reproduce conformational features observed in earlier studies at the level of neighboring base pairs. For example, the global bending isotropy of regular DNA polymers arises at the local level from the same 2 to 3-bp phase shift of Roll distortions (Fig. 4) found in simulations of short duplexes in a combined helicoidal and dihedral angle space (Zacharias and Sklenar, 2000).

Notably, the current low resolution description of DNA motions shows better correspondence with measured global properties of DNA than previous normal mode studies at higher resolution, e.g., (Matsumoto and Gō, 1999). The normal mode analysis of DNA in dihedral angle (or Cartesian coordinate) space is performed with respect to one of a large number of closely spaced minimum energy substates of similar structural character (Poncin et al., 1992) and the transitions between closely related, competing minima are not considered. As a result, the elastic force constants based on earlier calculations are several times higher than those reported in this work. The knowledge-based, elastic force field used here (Eq. 8) reflects a large number of base-pair steps in known crystal structures, each of which is analogous to one of the minimum energy substates used as the equilibrium reference point in higher resolution normal mode analyses. The fluctuations on the dimer energy surface are thus comparable to transitions between different minimum energy states on an atomic-scale map. The successful correspondence between the bending and twisting constants computed at the base-pair level with values obtained from experiment (Table 1) reveals the significant influence of transitions between conformational substates on the global properties of DNA.

Sequence-dependent insights

At the polymeric level, the local conformational features in the dimeric DNA model bring structural insights not possible with conventional elastic treatments. The calculations reported here show how well known discrepancies in the measured twisting constants of long DNA molecules could originate in the deformability of neighboring base-pair steps. The present set of force constants restricts the twisting of AA · TT steps much more so than any other dimer, resulting in a substantially higher twisting modulus for poly dA · poly dT compared to other regularly repeating polymers. The computed ratio A/C of the ease of global bending versus twisting is a composite of the conformational variability of successive dimers. Computational estimates of these molecular constants are expected to improve as the database of high resolution DNA structures grows and the knowledge-based force field of dimer distortions becomes stabilized, i.e., independent of the information in new structural examples.

The present study further illustrates how the natural coupling of local conformational variables affects the global motions of DNA. The synchronous variation of Roll and/or Slide with Rise lowers the global stretching constant substantially, thereby providing a new structural perspective on the micromanipulation of individual molecules. Successful correspondence of the computed stretching modulus with experimental data requires that the DNA base pairs be inclined with respect to the direction of stretching or reoriented upon forced extension. Chain extension is thereby effected by low energy transverse motions which vary the overlap but preserve the strong van der Waals' stacking of neighboring base-pair planes. The computed stretching constant significantly exceeds observed values if the DNA is pulled in a direction perpendicular to the base-pair planes. It should be noted that the end-to-end distance of DNA is far smaller than its contour length at the beginning of a single-molecule stretching experiments, and that the tensile force is simply an entropic effect (Smith et al., 1992, 1996). Clearly, the base-pair planes are not perpendicular to the stretching direction at low imposed extension, and this is not the state with which we make comparison. Instead we consider more extended chains where the end-to-end distance is comparable to the contour length, the entropic effect is less important, and the influence of local structural change become dominant. The overall structure of DNA may be nearly straight in this state, but the base-pair planes may not be perfectly perpendicular to the stretching direction.

The calculations also reveal how appropriate phasing of base-pair deformability with the double helical repeat induces a mesoscopic bending anisotropy conducive to DNA loop formation. Even a naturally straight chain, such as the poly $d(A_5T_5)$ model in Fig. 9, will bend in a preferred direction if intrinsically flexible and stiff dimer steps are spaced at half-turn increments, i.e., ~5 bp apart, on opposite sides of the double helix. Chains with such features are more likely to associate with strategically placed, upstream or downstream proteins in a hairpin loop structure, and are expected to experience more difficulty in snaking through

the pores of a gel. This picture of intrinsic global mobility differs from conventional "static" interpretations of the anomalously slow gel mobilities of certain DNA sequences on the basis of the intrinsic equilibrium structure of the constituent dimer or trimer steps. Moreover, basic conformational principles, which are sometimes broken to account for experimental observation from the latter perspective, remain intact when the intrinsic, sequence-dependent flexibility of DNA is also considered. For example, there is no need to introduce a trimer or tetramer model of DNA deformation (Brukner et al., 1995; Dlakic and Harrington, 1996), or to adjust the equilibrium structure of individual dimers from their crystallographically observed mean values, as recently reported in Liu and Beveridge (2001), to account for the anomalous retardation of multimers of $d(GA_4T_4C)$ on polyacrylamide gels and the more normal behavior of $d(GT_4A_4C)$ chains under the same conditions (Hagerman, 1986). The intrinsic stiffness of the AT steps in the former sequence combines with the deformability of CG dimers a half helical turn away so that the polymer bends as a whole, in spite of its relatively straight "static" equilibrium structure (0.2°/bp), preferentially in a single plane, i.e., with non-degenerate in-plane bending frequencies. The alternation of flexible CG and TA dimers on either side of the double helix, however, compensates for the natural static curvature (~1°/bp) of $d(GT_4A_4C)$ polymers, resulting in less directed global bending (with roughly equivalent in-plane bending frequencies) that facilitates passage of the long chain through a gel. The same overall picture is seen in detailed molecular dynamics (Sprous et al., 1999) and Monte Carlo (Zhurkin et al., 1991; Olson and Zhurkin, 2000) simulations: the bending in short A_4T_4-containing duplexes occurs in an unchanging direction over time, whereas the direction of bending for short T_4A_4-containing duplexes is highly variable.

Future directions

The correspondence between the root-mean-square end-to-end distances of DNA chains restricted to configurations along the lowest frequency bending modes with exact values obtained by averaging over all possible polymer configurations confirms the validity of the sequence-dependent harmonic treatment of long, relatively stiff double helical molecules and points to the potential utility of the method to characterize the distribution of spatial configurations. The configurations sampled along the dominant normal modes provide a more rapid way to estimate cyclization and loop formation probabilities than distributions of three-dimensional structures enumerated by conventional Metropolis-Monte Carlo procedures.

We are grateful to Dr. Xiang-Jun Lu for providing force field data and to Dr. A. R. Srinivasan for assistance with the matrix generator calculations of chain dimensions. Support of this work through U.S.P.H.S. grant GM34809 and the New Jersey Commission on Science and Technology (Center for Biomolecular Applications of Nanoscale Structures) is gratefully acknowledged. Computations were carried out at the Rutgers University Center for Computational Chemistry.

GLOSSARY

A	bending constant of an ideal elastic rod
$\mathbf{A}$	$6(N-1)\times6(N-1)$ eigenvector matrix with components A_{in} that describe the fractional contribution of the ith "step" parameter to the nth normal mode
A_{in}	component of $\mathbf{A}$ (see preceding entry)
a	persistence length
C	twisting constant of an ideal elastic rod
C_b	material constant used in the evaluation of the bending of an elastic rod
C_s	material constant used in the evaluation of the stretching of an elastic rod
C_t	material constant used in the evaluation of the twisting of an elastic rod
$\mathrm{C_{XY}}$	6×6 constant matrix characteristic of the intrinsic orientation of the XY dimer in B-form DNA
$\mathbf{F}$	$6(N-1)\times6(N-1)$ matrix of elastic force constants
$\mathbf{F}_{\mathrm{XY}}$	6×6 matrix of force constants which incorporate the intrinsic orientation of the XY dimer; $\mathbf{C}_{\mathrm{XY}}{}^{\mathrm{T}}\mathbf{G}_{\mathrm{XY}}\mathbf{C}_{\mathrm{XY}}$
f_{ij}	element of $\mathbf{F}$ describing the potential deformability of the ith and jth "step" parameters
$\mathbf{G}_{\mathrm{XY}}$	6×6 matrix of force constants associated with the XpY dimer
g_{ij}	element of $\mathbf{G}_{\mathrm{XY}}$ describing the deformation of the XpY dimer
$\mathbf{H}$	$6(N-1)\times6(N-1)$ kinetic energy matrix
h_{ij}	weighted "mass" coefficient element of $\mathbf{H}$
I_M	moment of inertia per unit length around the twisting axis of an elastic rod
i, j	"step" parameter indices
K	kinetic energy
k	base-pair index
k_B	Boltzmann constant
L	contour length of DNA measured by the displacement along the helical axis
M	total mass of a molecule
m	repeat length index of poly $d(A_mT_m)$ sequences
m_a	mass of atom a
N	number of base pairs
n	normal mode index
n_b, n_t, n_s	bending, twisting, and stretching mode indices

n_b'	index of degenerate bending mode
$\mathbf{o}_k$	Cartesian coordinates of the origin of the kth base pair
p_n	parameter describing the frequency of the nth normal mode of an ideal elastic rod
$\mathbf{Q}$	$6(N-1)$ vector of normal coordinates equal to $\mathbf{A}^{-1}\Delta\Theta$
$\mathbf{R}_k$	rotation matrix which relates the kth base-pair frame to the global reference frame
$\mathbf{r}$	end-to-end vector
$\mathbf{r}^o$	end-to-end vector of the equilibrium (minimum energy) configuration
$\langle r^2\rangle_0$	unperturbed mean-square end-to-end distance
$\mathbf{r}_a$	Cartesian coordinates of atom a
T	transpose of a matrix
T	temperature in Kelvin (always premultiplied by k_B to denote the thermal energy)
$\mathbf{T}_o$	constant transformation matrix obtained by evaluating $\mathbf{T}$ at the (Tilt, Roll, Twist) equilibrium rest state
$\mathbf{T}_{k,k+1}$	dimer transformation matrix which relates the kth and $(k+1)$th base-pair frames
$t_{\nu\mu}$	components of $\mathbf{T}$
t	time
V	total conformational potential energy
V_{XY}	conformational potential energy of the XpY dimer
$\mathbf{u}$	axis of the single rotation which brings different coordinate frames into coincidence
(u_x, u_y, u_z)	components of $\mathbf{u}$
$\mathbf{v}_{k,k+1}$	displacement (virtual bond) vector between the coordinate origins of base pairs k and $k+1$
$(\mathbf{x}, \mathbf{y}, \mathbf{z})$	base-pair coordinate axes
Y	stretching constant of an ideal elastic rod
α_n	amplitude of the nth normal mode
$[\boldsymbol{\gamma}_x\ \boldsymbol{\gamma}_y\ \boldsymbol{\gamma}_z]$	unit vector components of the rotation matrix $\mathbf{R}$
δ_n	phase angle of the nth normal mode
$\delta_{\nu\mu}$	Kronecker delta
$\varepsilon_{\nu\mu\kappa}$	permutation symbol
$\boldsymbol{\Theta}$	collective vector of deformations of all $6(N-1)$ "step" parameters
θ_i	One of $6(N-1)$ "step" parameters in a DNA with N base pairs
$\dot{\theta}_i$	time derivative of θ_i
$\Delta\theta_i$	instantaneous fluctuation of the ith "step" parameter from its equilibrium value
$\boldsymbol{\Lambda}$	$6(N-1)$ diagonal matrix with nonzero elements equal to ω_n^2
λ	wavelength of a bending mode
λ_{nm}	element of $\boldsymbol{\Lambda}$
ν_n^b	frequency of the nth bending mode of an ideal elastic rod
ν_n^s	frequency of the nth stretching mode of an ideal elastic rod
ν_n^t	frequency of the nth twisting mode of an ideal elastic rod
$\sigma(n, t)$	statistical weight of the global configuration of DNA in normal mode n at time t
(τ_x, τ_y, τ_z)	components of the $k,k+1$ "step" displacement (Shift, Slide, Rise) projected on the coordinate axes of base-pair k
$\Delta\boldsymbol{\Phi}_{XY}$	collective vector of the deformations of the six XpY "step" parameters
ϕ_i	one of the six "step" parameters of the XpY dimer
φ	magnitude of the single rotation which brings different coordinate frames into coincidence
$(\varphi_x, \varphi_y, \varphi_z)$	components of φ equated respectively to the (Tilt, Roll, Twist) angular "step" parameters
ω_n	frequency of the nth normal mode

REFERENCES

Babcock, M. S., E. P. D. Pednault, and W. K. Olson. 1994. Nucleic acid structure analysis: mathematics for local Cartesian and helical structure parameters that are truly comparable between structures. *J. Mol. Biol.* 237:125–156.

Barkley, M. D., and B. H. Zimm. 1979. Theory of twisting and bending of chain macromolecules: analysis of the fluorescence depolarization of DNA. *J. Chem. Phys.* 70:2991–3007.

Baumann, C. G., S. B. Smith, V. A. Bloomfield, and C. Bustamante. 1997. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* 94:6185–6190.

Bazett-Jones, D. P., J. Côté, C. C. Landel, C. L. Peterson, and J. L. Workman. 1999. The SWI/SNF complex creates loop domains in DNA and polynucleosome arrays and can disrupt DNA-histone contacts within these domains. *Mol. Cell. Biol.* 19:1470–1478.

Bednar, J., P. Furrer, V. Katritch, A. Z. Stasiak, J. Dubochet, and A. Stasiak. 1995. Determination of DNA persistence length by cryo-electron microscopy: separation of the static and dynamic contributions to the apparent persistence length of DNA. *J. Mol. Biol.* 254:579–594.

Bellomy, G. R., M. C. Mossing, and M. T. Record, Jr. 1988. Physical properties of DNA in vivo as probed by the length dependence of the *lac* operator looping process. *Biochemistry.* 27:3900–3906.

Berendsen, H. J., and S. Hayward. 2000. Collective protein dynamics in relation to function. *Curr. Opin. Struct. Biol.* 10:165–169.

Berkoff, B., M. Hogan, J. LeGrange, and R. Austin. 1986. Dependence of oxygen quenching of intercalated methylene blue triplet lifetime on DNA base-pair composition. *Biopolymers.* 25:307–316.

Bouchiat, C., and M. Mezard. 1998. Elasticity model of a supercoiled DNA molecule. *Phys. Rev. Lett.* 80:1556–1559.

Braun, W., S. Yoshioki, and N. Gō. 1984. Formulation of static and dynamic conformational energy analysis of biopolymer systems consisting of two or more molecules. *J. Phys. Soc. Japan.* 53:3269–3275.

Brooks, B., and M. Karplus. 1983. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* 80:6571–6575.

Brukner, I., R. Sanchez, D. Suck, and S. Pongor. 1995. Trinucleotide models for DNA bending propensity: comparison of models based on DNaseI digestion and nucleosome packaging data. *J. Biomol. Struct. Dynam.* 13:309–317.

Bustamante, C., J. F. Marko, E. D. Siggia, and S. Smith. 1994. Entropic elasticity of λ-phage DNA. *Science.* 265:1599–1600.

Bustamante, C., C. Rivetti, and D. J. Keller. 1997. Scanning force microscopy under aqueous solutions. *Curr. Opin. Struct. Biol.* 7:709–716.

Chandrasekaran, R., and S. Arnott. 1989. The structures of DNA and RNA helices in oriented fibers. *In* Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology, Group VII/1b, Nucleic Acids. W. Saenger, editor. Springer-Verlag, Berlin. 31–170.

Chasles, M. 1830. Note sur les propriétés générales du système de deux corps semblables entre'eus et placés d'une manière quelconque dans l'espace; et sur la déplacement fini ou infiniment petit d'un corps solide libre. *Férussac. Bull. Sci. Math.* 14:321–326.

Chen, H. H., D. C. Rau, and E. Charney. 1985. The flexibility of alternating dA-dT sequences. *J. Biomol. Struct. Dynam.* 2:709–719.

Chen, Y. Z., and E. W. Prohofsky. 1995. Normal mode calculation of a netropsin-DNA complex: effect of structural deformation on vibrational spectrum. *Biopolymers.* 35:657–666.

Dickerson, R. E., M. Bansal, C. R. Calladine, S. Diekmann, W. N. Hunter, O. Kennard, E. von Kitzing, R. Lavery, H. C. M. Nelson, W. K. Olson, W. Saenger, Z. Shakked, H. Sklenar, D. M. Soumpasis, C.-S. Tung, A. H.-J. Wang, and V. B. Zhurkin. 1989. Definitions and nomenclature of nucleic acid structure parameters. *J. Mol. Biol.* 208:787–791.

Diekmann, S., W. Hillen, B. Morgeneyer, R. D. Wells, and D. Pörschke. 1982. Orientation relaxation of DNA restriction fragments and the internal mobility of the double helix. *Biophy. Chem.* 15:263–270.

Dillon, N., T. Trimborn, J. Strouboulis, P. Fraser, and F. Grosveld. 1997. The effect of distance on long-range chromatin interactions. *Mol. Cell.* 1:131–139.

Dlakic, M., and R. E. Harrington. 1996. The effects of sequence context on DNA curvature. *Proc. Natl. Acad. Sci. U. S. A.* 93:3847–3852.

Eckart, C. 1935. Some studies concerning rotating axes and polyatomic molecules. *Phys. Rev.* 47:552–558.

Fenley, M. O., W. K. Olson, I. Tobias, and G. S. Manning. 1994. Electrostatic effects in short superhelical DNA. *Biophys. Chem.* 50:255–271.

Flory, P. J. 1969. Statistical Mechanics of Chain Molecules. Interscience, New York. Chapter 4.

Fujimoto, B. S., J. H. Shibata, R. L. Schurr, and J. M. Schurr. 1985. Torsional dynamics and rigidity of fractionated poly(dGdC). *Biopolymers.* 24:1009–1022.

Garcia, A. E., and D. M. Soumpasis. 1989. Harmonic vibrations and thermodynamic stability of a DNA oligomer in monovalent salt solution. *Proc. Natl. Acad. Sci. U.S.A.* 86:3160–3164.

Gorin, A. A., V. B. Zhurkin, and W. K. Olson. 1995. B-DNA twisting correlates with base pair morphology. *J. Mol. Biol.* 247:34–48.

Ha Duong, T., and K. Zakrzewska. 1997a. Calculation and analysis of low frequency normal modes for DNA. *J. Comp. Chem.* 18:796–811.

Ha Duong, T., and K. Zakrzewska. 1997b. Influence of drug binding on DNA flexibility: a normal mode analysis. *J. Biomol. Struct. Dynam.* 14:691–701.

Ha Duong, T., and K. Zakrzewska. 1998. Sequence specificity of bacteriophage 434 repressor-operator complexation. *J. Mol. Biol.* 280:31–39.

Hagerman, P. J. 1981. Investigation of the flexibility of DNA using transient electric birefringence. *Biopolymers.* 20:1503–1535.

Hagerman, P. J. 1985. Analysis of the ring-closure probabilities of isotropic wormlike chains: application to duplex DNA. *Biopolymers.* 24: 1881–1897.

Hagerman, P. J. 1986. Sequence-directed curvature of DNA. *Nature.* 321:449–450.

Hagerman, P. J. 1988. Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.* 17:265–286.

Hansma, H. G., I. Revenko, K. Kim, and D. E. Laney. 1996. Atomic force microscopy of long and short double-stranded, single-stranded and triple-stranded nucleic acids. *Nucleic Acids Res.* 24:713–720.

Heath, P. J., J. B. Clendenning, B. S. Fujimoto, and J. M. Schurr. 1996. Effect of bending strain on the torsion elastic constant of DNA. *J. Mol. Biol.* 260:718–730.

Higo, J., Y. Seno, and N. Gō. 1985. Formulation of static and dynamic conformational energy analysis of biopolymer systems consisting of two or more molecules—avoiding a singularity in the previous method. *J. Phys. Soc. Jpn.* 54:4053–4058.

Hogan, M., J. LeGrange, and B. Austin. 1983. Dependence of DNA helix flexibility on base composition. *Nature.* 304:752–754.

Hogan, M. E., and R. H. Austin. 1987. Importance of DNA stiffness in protein-DNA binding specificity. *Nature.* 329:263–266.

Horowitz, D. S., and J. C. Wang. 1984. Torsional rigidity of DNA and length dependence of the free energy of DNA supercoiling. *J. Mol. Biol.* 173:75–91.

Hua, X. M., and E. W. Prohofsky. 1988. Normal-mode calculation for methylated Z-DNA poly(dG-m5dC)·(dG-m5dC). *Biopolymers.* 27: 645–655.

Hurley, I., P. Osei-Gyimah, S. Archer, C. P. Scholes, and L. S. Lerman. 1982. Torsional motion and elasticity of the deoxyribonucleic acid double helix and its nucleosomal complexes. *Biochemistry.* 21: 4999–5009.

Irikura, K. K., B. Tidor, B. R. Brooks, and M. Karplus. 1985. Transition from B to Z DNA: contribution of internal fluctuations to the configurational entropy difference. *Science.* 229:571–572.

Jeffreys, H., and B. S. Jeffreys. 1946. Methods of Mathematical Physics. Cambridge University Press, Cambridge, UK.

Joanicot, M., and B. Révet. 1987. DNA conformational studies from electron microscopy: I. Excluded volume effect and structure dimensionality. *Biopolymers.* 26:315–326.

Kahn, J. D., and D. M. Crothers. 1998. Measurement of the DNA bend angle induced by the catabolite activator protein using Monte Carlo simulation of cyclization kinetics. *J. Mol. Biol.* 276:287–309.

Kamien, R. D., T. C. Lubensky, P. Nelson, and C. S. O'Hern. 1997. Direct determination of DNA twist-stretch coupling. *Europhys. Lett.* 38: 237–242.

Kitao, A., and N. Gō. 1999. Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Biol.* 9:164–169.

Kosikov, K. M., A. A. Gorin, V. B. Zhurkin, and W. K. Olson. 1999. DNA stretching and compression: large-scale simulations of double helical structures. *J. Mol. Biol.* 289:1301–1326.

Kottalam, J., and D. A. Case. 1990. Langevin modes of macromolecules: applications to crambin and DNA hexamers. *Biopolymers.* 29: 1409–1421.

Levene, S. D., and D. M. Crothers. 1986. Ring closure probabilities for small DNA fragments by Monte Carlo simulations. *J. Mol. Biol.* 189: 61–72.

Levitt, M. 1978. How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations. *Proc. Natl. Acad. Sci. U.S.A.* 75:640–644.

Levitt, M., C. Sander, and P. S. Stern. 1985. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J. Mol. Biol.* 181:423–447.

Levy, R. M., and M. Karplus. 1979. Vibrational approach to the dynamics of an α-helix. *Biopolymers.* 18:2465–2495.

Lin, D., A. Matsumoto, and N. Gō. 1997. Normal mode analysis of a double-stranded DNA dodecamer d(CGCGAATTCGCG). *J. Chem. Phys.* 107:3684–3690.

Liu, Y., and D. L. Beveridge. 2001. A refined prediction method for gel retardation of DNA oligonucleotides from dinucleotide step parameters: reconciliation of DNA bending models with crystal structure data. *J. Biomol. Struct. Dynam.* 18:505–526.

Lu, X. J., and W. K. Olson. 1998. Resolving the discrepancies among nucleic acid conformational analyses. *J. Mol. Biol.* 285:1563–1575.

Lyubchenko, Y., B. L. Jacobs, S. M. Lindsay, and A. Stasiak. 1995. Atomic force microscopy of nucleoprotein complexes. *Scan. Microsc.* 9:705–727.

Marko, J. F. 1997. Stretching must twist DNA. *Europhys. Lett.* 38: 183–188.

Marky, N. L., and W. K. Olson. 1994. Configurational statistics of the DNA duplex: extended matrix generator matrices to treat the rotations and translations of adjacent residues. *Biopolymers.* 34:109–120.

Maroun, R. C., and W. K. Olson. 1988. Base sequence effects in double helical DNA: II. Configurational statistics of rodlike chains. *Biopolymers.* 27:561–584.

Matsumoto, A., and N. Gō. 1999. Dynamic properties of double-stranded DNA by normal mode analysis. *J. Chem. Phys.* 110:11070–11075.

McLachlan, A. D. 1979. Least squares fitting of two structures. *J. Mol. Biol.* 128:74–79.

Millar, D. P., R. J. Robbins, and A. H. Zewail. 1982. Torsion and bending of nucleic acids studied by subnanosecond time-resolved fluorescence depolarization of intercalated dyes. *J. Chem. Phys.* 76:2080–2094.

Muzard, G., B. Theveny, and B. Révet. 1990. Electron microscopy mapping of pBR322 DNA curvature: comparison with theoretical models. *EMBO J.* 9:1289–1298.

Nadeau, J. G., and D. M. Crothers. 1989. Structural basis for DNA bending. *Proc. Natl. Acad. Sci. U. S. A.* 86:2622–2626.

Namoradze, N. Z., A. N. Goryunov, and T. M. Birshtein. 1977. On conformations of the superhelix structure. *Biophys. Chem.* 7:59–70.

Noguti, T., and N. Gō. 1983. Dynamics of native globular proteins in terms of dihedral angles. *J. Phys. Soc. Jpn.* 52:3283–3288.

Olson, W. K. 1979. The flexible DNA helix: II. Superhelix formation. *Biopolymers.* 18:1235–1260.

Olson, W. K. 1996. Simulating DNA at low resolution. *Curr. Opin. Struct. Biol.* 6:242–256.

Olson, W. K., A. A. Gorin, X.-J. Lu, L. M. Hock, and V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. U.S.A.* 95:11163–11168.

Olson, W. K., X.-J. Lu, J. Westbrook, H. M. Berman, M. Bansal, S. K. Burley, R. E. Dickerson, S. C. Harvey, U. Heinemann, S. Neidle, Z. Shakked, M. Suzuki, C.-S. Tung, H. Sklenar, E. Westhof, and C. Wolberger. 2001. A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* 313:229–237.

Olson, W. K., N. L. Marky, R. L. Jernigan, and V. B. Zhurkin. 1993. Influence of fluctuations on DNA curvature: a comparison of flexible and static wedge models of intrinsically bent DNA. *J. Mol. Biol.* 232: 530–554.

Olson, W. K., and V. B. Zhurkin. 1996. Twenty years of DNA bending. *In* Biological Structure and Dynamics, Vol. 2. R. H. Sarma and M. H. Sarma, editors. Adenine Press, Schenectady, NY. 341–370.

Olson, W. K., and V. B. Zhurkin. 2000. Modeling DNA deformations. *Curr. Opin. Struct. Biol.* 10:286–297.

Pedone, F., F. Mazzei, M. Matzeu, and F. Barone. 2001. Torsional constants of 27-mer DNA oligomers of different sequences. *Biophys. Chem.* 94:175–184.

Poncin, M., B. Hartmann, and R. Lavery. 1992. Conformational sub-states in B-DNA. *J. Mol. Biol.* 226:775–794.

Pörschke, D. 1986. Structure and dynamics of double helices in solution modes of DNA bending. *J. Biomol. Struct. Dynam.* 4:373–389.

Révet, B., J. M. Malinge, E. Delain, M. Le Bret, and M. Leng. 1984. Electron microscopic measurement of chain flexibility of poly(dG-dC)·poly(dG-dC) modified by cis-diamminedichloroplatinum(II). *Nucleic Acids Res.* 12:8349–8362.

Ringrose, L., S. Chabanis, P.-O. Angrand, C. Woodroofe, and A. F. Stewart. 1999. Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances. *EMBO J.* 18:6630–6641.

Saenger, W. 1984. Principles of Nucleic Acid Structure. Springer-Verlag, New York. Chapter 11.

Schleif, R. 1992. DNA looping. *Annu. Rev. Biochem.* 61:199–223.

Smith, S. B., Y. Cui, and C. Bustamante. 1996. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science.* 271:795–799.

Smith, S. B., L. Finzi, and C. Bustamante. 1992. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science.* 258:1122–1126.

Song, L. U., and J. M. Schurr. 1990. Dynamic bending rigidity of DNA. *Biopolymers.* 30:229–237.

Sprous, D., M. A. Young, and D. L. Beveridge. 1999. Molecular dynamics studies of axis bending in $d(G_5\text{-}(GA_4T_4C)_2\text{-}C_5)$ and $d(G_5\text{-}(GT_4A_4C)_2C_5)$: effects of sequence polarity on DNA curvature. *J. Mol. Biol.* 285:1623–1632.

Strutt, J. W., Baron Rayleigh. 1894. The Theory of Sound. Cambridge University Press, London.

Sussman, J. L., and E. N. Trifonov. 1978. Possibility of nonkinked packing of DNA in chromatin. *Proc. Natl. Acad. Sci. U.S.A.* 75:103–107.

Théveny, B., D. Coulaud, M. Le Bret, and B. Révet. 1988. Local variations of curvature and flexibility along DNA molecules analyzed from electron micrographs. *In* DNA Bending and Curvature, Vol. 3. W. K. Olson, M. H. Sarma, R. H. Sarma, and M. Sundaralingam, editors. Adenine Press, Schenectady, NY. 39–55.

Thomas, J. C., S. A. Allison, C. J. Apellof, and J. M. Schurr. 1980. Torsion dynamics and depolarization of fluorescence of linear macromolecules: II. Fluorescence polarization anisotropy measurements on a clean viral ϕ29 DNA. *Biophys. Chem.* 12:177–188.

Thomas, J. C., and J. M. Schurr. 1983. Fluorescence depolarization and temperature dependence of the torsion elastic constant of linear ϕ 29 deoxyribonucleic acid. *Biochemistry.* 22:6194–6198.

Thomas, T. J., and V. A. Bloomfield. 1983. Chain flexibility and hydrodynamics of the B and Z forms of poly(dG-dC)·poly(dG-dC). *Nucleic Acids Res.* 11:1919–1930.

Thresher, R., and J. Griffith. 1992. Electron microscopic visualization of DNA and DNA-protein complexes as adjunct to biochemical studies. *Methods Enzymol.* 211:481–490.

Tidor, B., K. K. Irikura, B. R. Brooks, and M. Karplus. 1983. Dynamics of DNA oligomers. *J. Biomol. Struct. Dynam.* 1:231–252.

Vologodskii, A. V., and N. R. Cozzarelli. 1995. Modeling of long-range electrostatic interactions in DNA. *Biopolymers.* 35:289–296.

Wang, M. D., H. Yin, R. Landick, J. Gelles, and S. M. Block. 1997. Stretching DNA with optical tweezers. *Biophys. J.* 72:1335–1346.

Westcott, T. P., I. Tobias, and W. K. Olson. 1997. Modeling self-contact forces in the elastic theory of DNA supercoiling. *J. Chem. Phys.* 107: 3967–3980.

Wilson, Jr., E. B., J. C. Decius, and P. C. Cross. 1955. Molecular Vibrations. McGraw-Hill Book Company, Inc., New York. Chapter 10.

Yathindra, N. 1979. DNA flexibility and tertiary folding: a model for superfolding of DNA in chromatin. *Curr. Sci.* 48:753–756.

Zacharias, M., and H. Sklenar. 2000. Conformational deformability of RNA: a harmonic mode analysis. *Biophys. J.* 78:2528–2542.

Zhurkin, V. B., Y. P. Lysov, and V. Ivanov. 1979. Anisotropic flexibility of DNA and the nucleosomal structure. *Nucleic Acids Res.* 6:1081–1096.

Zhurkin, V. B., N. B. Ulyanov, A. A. Gorin, and R. L. Jernigan. 1991. Static and statistical bending of DNA evaluated by Monte Carlo simulations. *Proc. Natl. Acad. Sci. U. S. A.* 88:7046–7050.