JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

Comparative Study of the Bending Free Energies of C- and G-Based DNA: A-, B-, and Z-DNA and Associated Mismatched Trinucleotide Repeats

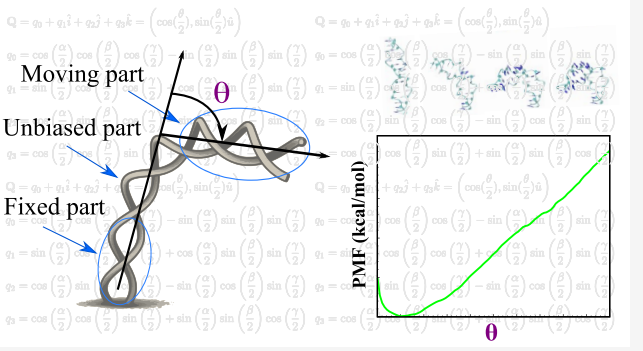
Published as part of Journal of Chemical Information and Modeling special issue "Applications of Free-Energy Calculations to Biomolecular Processes".

Ashkan Fakharzadeh, Mahmoud Moradi, Celeste Sagui, and Christopher Roland*

Cite This: J. Chem. Inf. Model. 2025, 65, 5672−5689 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: DNA's structural flexibility plays a crucial role in various biological functions such as gene replication, repair, and regulation as well as DNA−protein recognition. We investigate the bending free energy of short DNA helices, including $d(5'\text{-}(CG)_7C\text{-}3')_2$ in A-, B-, and Z-forms, and C- and G-rich trinucleotide repeat helices, using orientation quaternions with enhanced sampling methods. The orientation quaternion technique provides an effective method to induce rotational transformations or to restrain the orientation of certain domains of biomolecular systems. This methodology was implemented in the AMBER simulation package and used to induce DNA bending in two separate ways: free bending and directional bending. We found that the bending free energy varies quadratically for moderate bending and then becomes almost linear for larger bending angles. The left-handed Z-DNA helix was found to exhibit the highest rigidity among the canonical DNA forms studied. The mechanisms associated with bending were also investigated with evidence for type I and type II kinks depending on the sequence and the helical form considered. The duplexes exhibit high flexibility in the presence of CC and GG mismatches, particularly CGG and GGC trinucleotide repeats in the Z-form, which have the lowest bending free energies. These calculations provide new insight into the mechanics of the global conformational flexibility of DNA molecules by quantifying the energetic cost and preferred directions of bending.



INTRODUCTION

The structural flexibility of DNA plays an important role in many biological functions, such as gene replication, expression, regulation, DNA repair, protein−DNA recognition, and simply fitting DNA inside the cell.[1] For instance, protein−DNA recognition often involves substantial DNA bending which has important consequences for the binding thermodynamics.[2] Likewise, DNA bending is believed to be a key feature facilitating the mismatch recognition and/or repair processes in nucleic acids through a reduction of the energetic barriers for base flipping.[3] Clearly, a characterization of the structural aspects of bent or deformed nucleic acids is crucial for the understanding of many of the biological functions that these molecules undertake.

Unfortunately, bent and deformed DNA is associated with transient conformations that are difficult to probe. Thus, controlling the direction of bending by means of a pulling experiment is hard. In turn, statistical analysis of DNA and protein−DNA structures indicates that the majority of bending is accommodated by changes in roll angles, with contradictory experimental results for the favorability of bending toward the minor or major grooves.[4−6] In terms of the microscopics of bending, DNA stiffness arises from a multiterm energetic balance, including electrostatic screening of the backbone's phosphate groups, base-pair stacking (van der Waals and $\pi-\pi$ interactions), backbone torsional strain, and solvent effects, whose details are strongly influenced by the specifics of the nucleic acid sequence and the salt of the environment.[7]

In terms of our theoretical understanding of bent DNA, the worm-like chain (WLC) model[8] is probably the most successful model for describing the flexibility of DNA on a coarse-grained or large-length scale (150 base pairs or more).[9] However, the model's ability to describe DNA bending on short-length scales

Received: March 11, 2025
Revised: April 29, 2025
Accepted: May 7, 2025
Published: May 16, 2025

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

is open to questions,[10−13] with considerable debate over the conformational changes, sequence-dependence, and physical characteristics for short scale DNA bending. Different types of deformations have been observed in bent DNA structures, including flipped-out bases,[14] local bubbles,[15] and kinked structures.[16,17] Sharp kinking of bent DNA was first proposed by Crick and Klug in 1975,[18] and two types of kinks have been put forward: type I kinks are distinguished by significant bending at a single base pair (with a very high roll angle) and with little disturbance of the neighboring pairs; while type II kinks are characterized by a localized melting of DNA. Three successive base pairs are involved in this type of kink, in which the central base pair is broken and its bases are stacked onto the 5′ direction of the corresponding strand.[18] Sample structures of DNA illustrating these type I and type II kinks, taken from crystal structures of protein−DNA complexes, are shown in the Supporting Information (SI) in Figure S1.

In principle, molecular dynamics (MD) simulations are well suited to studying DNA deformations on short-length scales. However, such simulations may be limited by the so-called sampling problem: given the considerable free-energy barriers associated with structural changes in DNA, the simulation times are often insufficient to correctly probe bending conformations. The sampling problem may, however, be overcome by means of special biasing techniques that enable the simulations to probe the otherwise inaccessible regions of phase space. For instance, Curkusu et al.[19] used a screw axis restraint to induce continuous bending on several DNA structures, including CG B-DNA, A(adenine)-tract, and DNAs with mismatches or abasic sites, and discovered that while moderate bending occurs via coupled base-pair step roll, strong bending results in local type II kinks. Furthermore, they suggested a greater variety of bent structures and directions for C:C and G:A mismatches in comparison to regular B-DNA. Sharma et al.[20] also investigated the bending of DNA with different single mismatches, where the angle between the centers of the heavy atoms in three regions of DNA was biased using the umbrella sampling (US)[21] method. Their analysis revealed various bending free energies that correlated with base flipping. In a more recent study, Ma et al.[6] examined the bending preference of regular B-DNA toward the minor/major grooves and the effects of salt on the bending of orientation using roll, tilt, and twist angles.

Given the importance of bent DNA, especially in the context of promoting mismatch repair, we have undertaken a large-scale MD study investigation of such helical structures based on the CG sequence and associated mismatched trinucleotide repeats. Specifically, we have studied the free energies and structural characteristics of CG-based A-, B- and Z-DNA, as well as of CCG and GGC trinucleotide repeat sequences with either CC or GG mismatches. These trinucleotide repeat sequences are in B- or Z-form, or both. The considered CG-based A-DNA is included here simply for the sake of comparison and completeness.

Nucleotides naturally create right-handed A-DNA or B-DNA forms, with B-DNA being the most common form at neutral pH and physiological salt concentrations, but alternative forms become accessible under defined physicochemical triggers. A-DNA emerges in low-water environments (e.g., ≥75% ethanol/TFE), within protein-bound pockets, and in DNA/RNA hybrid duplexes.[22,23] The A-DNA form is similar to B-DNA but is wider and denser along its axis.[24] On the other hand, Z-DNA is a left-handed double helix with Watson−Crick (WC) base pairs. Z-DNA is favored by high-salt buffers *in vitro*, yet *in vivo* it can be induced by CpG methylation, negative superhelical stress, or molecular-crowding organic cosolvents. This left-handed helix is thinner and more rigid than B-DNA, with a zigzag backbone pattern that plays an important role in gene expression regulation and recombination.[25−27] Despite these specialized conditions, we considered the A- and Z-form to benchmark against B-DNA and to provide a consistent comparison set for future computational and experimental studies. For comparison purposes, SI Figure S2 illustrates the different helical DNA forms, while Table S1 lists some common structural characteristics.

To investigate the effect of sequence and repeat expansion on DNA bendability, we examined C- and G-rich Trinucleotide Repeats (TRs) in the form of (CCG), (GCC), (GGC), and (CGG) as these are behind some of the Trinucleotide Repeat Expansion Diseases (TREDs).[28−34] TREDs represent genetic disorders caused when the number of trinucleotide repeats associated with a gene exceeds a gene-specific threshold. These TRs display various secondary structures including hairpins, e-motifs, and Z-DNA depending on the environment.[35−37] In TREDs, CGGs occur in the 5′-UTR of the FMR1 gene,[38] while CCGs are present in the 5′-UTR and coding regions of multiple genes. Pathological expansions of CGG TRs cause conditions such as FXTAS, premature ovarian failure,[39] and fragile X mental retardation syndrome.[40] CGG expansions and methylation in the XYLT1 gene's exon 1 are associated with Baratela−Scott syndrome.[41] CCG TRs relate to three TREDs; with the largest expansion in the FRM2 gene causes FRAXE[42] and influences diseases like Huntington's and myotonic dystrophy type 1.[43]

In prior investigations,[35−37] we characterized the structure and dynamics of these repeats: CCG forms stable B-DNA with CC mismatches inside the helical core, surrounded by two WC-paired GpC steps; the GCC reading frame exhibits CpG steps surrounding mismatches where mismatched C-bases flip out symmetrically from the minor groove, resulting in an "e-motif".[35,44,45] GGC and CGG with GG mismatches exhibit CpG and GpC steps, respectively, and can adopt both B- and Z-forms. In the B-form, GG mismatches form hydrogen bonds inside the helix for both GGC and CGG reading frames,[36] while in the Z-form, GG mismatches are extruded in the GGC reading frame but remain stable in syn−syn conformation in the CGG reading frame.[37,46]

In order to induce DNA bending in our MD simulations, we used enhanced sampling methods based on quaternion-based collective variables (CVs). Orientation quaternion CVs are an effective way of inducing rotational transformations or restricting the orientation of a specific domain in a biomolecular system.[47] The orientation quaternion formalism has previously been used in the context of large-scale conformational changes of proteins.[48−50] Here, for the first time, we use them in the context of small nucleic acids. Our investigations make extensive use of these CVs, which we recently implemented in the AMBER simulation package.[51] This methodology allows for an efficient and accurate estimation of the free energy of the different helical structures in terms of the bending angle.

MATERIALS AND METHODS

In terms of methodology, there are two aspects to our study: the use of quaternion CVs in order to define DNA bending and the quantification of the bending process by mapping the free energy. Here, we briefly discuss each component.

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

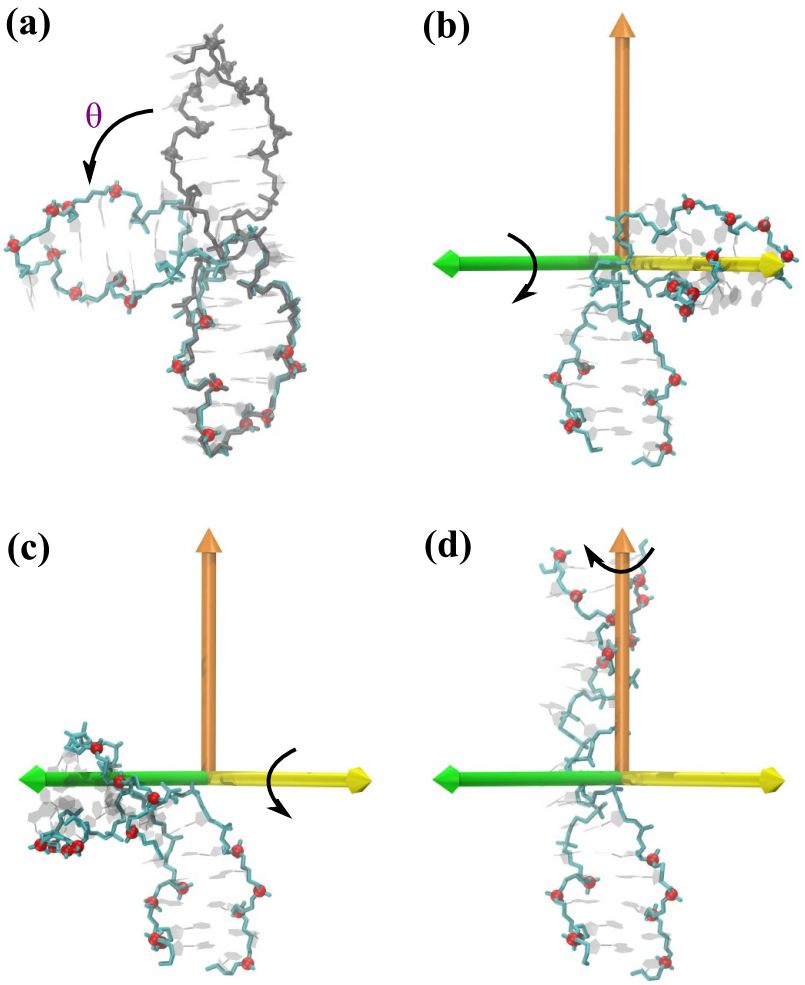


Figure 1. Schematic view of CVs used in this study. The bending angle θ, as defined in Figure S3, represents the angle between the mass center of P atoms in the top segment of DNA (residues 2−5, and 27−30) in the reference frame and the rotated frame. The mass center of P atoms in the bottom segment of DNA (residues 12−15, and 17−20) is restrained. The P atoms are shown with spheres (red and gray). The drawn axes represent the principal axes: roll (orange), yaw (yellow), and pitch (green). (a) Equilibrated initial DNA structure (gray), and a bent one (cyan), (b) rotation of the DNA's top section around the positive direction (clockwise) of the pitch axis, (c) rotation of the DNA's top section around the negative direction (counterclockwise) of yaw axis, and (d) rotation of the DNA's top section around the positive direction (clockwise) of the roll.

Quaternion CVs. A number of ways have been proposed as a measure of bending angle.[19,20,52−54] In principle, the bending angle of a duplex can be defined as a simple geometric angle formed by the mass centers of three groups of atoms, which we will refer to as bottom, middle, and top domains for convenience. Similarly, the arc cosine of the dot product of the principal axes (i.e., the long axis) of two groups of atoms, bottom and top domains, can be used to define the bending angle. Unfortunately, using a simple three-domain definition of an angle to induce bending requires biasing several atoms and often results in undesired deformations of the duplex. Using the principal axes has its own set of technological challenges; for example, the three principal axes may interchange or be ill-defined during the simulation due to conformational changes. We define the bending angle of a duplex as the orientational changes of the top segment of the helix from an initial (equilibrated) one, while the bottom segment is fixed and restrained (see Figure S3). We will employ the orientation quaternion technique which has been shown[48−50] to be a well-behaved, versatile way for constructing system-specific CVs, specifically aimed at inducing orientational changes or restraining the orientation of certain domains. To explain the orientation quaternion procedure, we first need to briefly review quaternions and their application to domain rotation. Quaternions are an extension of complex numbers and are defined as four-component vectors via $q_0 + q_1\hat{i} + q_2\hat{j} + q_3\hat{k}$ where q_0, q_1, q_2, and q_3 are real numbers; and i, j, and k are the basic quaternions ($i^2 = j^2 = k^2 = ijk = -1$). Unit quaternions, where $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$, provide a convenient notation for representing three-dimensional rotations. A rotation can be specified by a unit vector defining the axis of rotation and an angle of θ for the rotation magnitude. Thus, a unit quaternion, $\hat{q}$, is expressed as $(\cos(\theta/2), \sin(\theta/2)\hat{u})$, where $\hat{u}$ is the unit vector associated with the axis of rotation. Using quaternion to represent three degrees of freedom of a rotation is equivalent to using the popular three Euler-angle rotations α, β, and γ around the coordinate axes (See Supporting Note 3 for conversion between quaternions and Euler angles). However, unlike the Euler angles, quaternion parametrization of rotations does not suffer from coordinate singularities.[55] Fiorin et al.[47] have previously shown that the orientation of macromolecules and flexible structures can be expressed through a quaternion. In this approach, the quaternion, $\hat{q}$, is found as the leading eigenvector of a 4 × 4

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

matrix constructed using both the current and the reference coordinates.[47,55]

By using quaternion-based CVs, one can control and induce a certain orientation in a simulation. This is achieved by defining a biasing potential that depends on the quaternion CVs. We have implemented several orientation quaternion-based CVs in the Non-Equilibrium Free-Energy suite of the AMBER simulation package (ver.20),[51] such as Orientation (QUATERNION0,⋯,-QUATERNION3), ORIENTATION_ANGLE, ORIENTATION_PROJ, SPINANGLE, and TILT. A detailed description of implementation, along with examples of usage, is given in SI (Supporting Information, Notes 1 and 2). In this work, we primarily make use of the ORIENTATION_ANGLE and the SPINANGLE CVs. The ORIENTATION_ANGLE, $\theta = 2\cos^{-1}(q_0)$, and ORIENTATION_PROJ (the cosine of ORIENTATION_ANGLE) use only the first component of $\hat{q}$, *i.e.*, q_0, thus they represent the amplitude of the full rotation, regardless of the direction of its axis. A slightly more complete parametrization is achieved by decomposing the full rotation into the two parameters, TILT and SPINANGLE. These quantify the amplitudes of two independent subrotations away from a certain axis $\vec{e}$, and around the same axis $\vec{e}$, respectively. The axis $\vec{e}$ is chosen by the user. The SPINANGLE explicitly is given by $\phi = 2\tan^{-1}(\vec{q}\cdot\vec{e}/q_0)$, where $\vec{q}$ is the vector part of quaternion $\hat{q}_0$.

In comparison to Euler angles, we note that the orientation angle θ does not directly correspond to any single Euler angle but rather to a combination of them as $\cos(\theta/2) = \cos(\alpha/2)\cos(\beta/2)\cos(\gamma/2) + \sin(\alpha/2)\sin(\beta/2)\sin(\gamma/2)$. On the other hand, the spin angle, ϕ, can be related to a single Euler angle if the axis of rotation $\vec{e}$ is aligned with the principal axes. For instance, when $\vec{e}$ is the *Z*-axis and there is no rotation around the other axes (*X* or *Y*), the spin angle ϕ is directly equal to the γ Euler angle (See Supporting Note 3 for a further discussion on the relation between the orientation and spin angles with Euler angles).

We specifically constructed two CVs based on the quaternion formalism as an effective method to induce DNA bending:

(i) Free bending (Orientation Angle): This CV measures the relative orientation from a reference frame using the mass centers of two groups of atoms. For the atom groups, the P atoms from four base pairs in the terminal pieces were employed. After reaching equilibrium, we restricted the mass center of P atoms in the bottom terminus, residues 12−15 and 17−20 (in 15 base-pair duplexes), to the reference frame while allowing the mass center of P atoms in the other terminus, residues 2−5 and 27−30 (in 15 base-pair duplexes), to rotate with respect to the initial frame, i.e., equilibrium. Note that 5′ residues, 1 and 16, were excluded as these do not include any P atoms. A sample AMBER colvar file used in this type of bending is given in Table S2. The pulling force was applied to all P atoms in the top section, while the bottom section was restrained. In addition, the RMSD of C5′ atoms of top and bottom sections were restrained to avoid unnecessary twisting of the whole structure during the long simulations. Given that the bottom segment is fixed, the angle formed between the rotated and the reference frames of the top segment defines the bend angle of the whole DNA. The bending direction is not constrained in this procedure and the molecule is free to adopt any direction spontaneously. A schematic representation of this CV is shown in Figure 1 (as well as SI Figure S3), where the orientation of the top section of DNA is altered by $\theta°$ relative to the initial frame (gray color). Note that our definition of bending angle, θ, differs from several other definitions used in the literature, but we anticipate our bending angle to be similar to the bending angle based on the average screw rotation axes as defined by Curkusu et al.[19] as these two definitions are conceptually similar. A direct comparison of methods was not possible due to the unavailability of codes for other methods. However, we compared our induced bending with the bending angle defined by the program 'Curves+'[56] and found a strong correlation. See SI (Supporting Note 4 and Figure S4) for a detailed comparison of the two methods.

(ii) Directional bending (Spin Angle): A precise directed bending may be achieved by confining the rotation to a specific axis. We employed principal axes of the whole duplex for the axis and P atoms from four termini base pairs for the atom groups. In a manner similar to the free bending, the mass center of P atoms in bottom terminus, residues 12−15 and 17−20 (in 15 base-pair duplexes), was kept in its equilibrium configuration, while the mass center of P atoms in the top terminus, residues 2−5 and 27−30 (in 15 base-pair duplexes), was rotated with respect to the initial frame (equilibrium) using an axis of rotation. In this procedure, bending around axes orthogonal to the bending axis is not allowed by restraining bending around those axes. A sample colvar file used in the directed bending is given in Table S3. The angle around the rotation axis between the rotated and the reference frames of the top segment defines the directed bending angle. It is important to note that the initial orientation of DNA can influence the principal axes as the initial equilibrated structures may be somewhat bent.

The principal axes for a sample DNA are shown in Figure 1 These are the roll axis (orange), parallel to the helical axis; the pitch axis (green), and the yaw axis (yellow). By rotation around the roll axis and restraining bending around the yaw and pitch axes, the helix can either unwind or further tighten. Positive or negative rotations around the yaw axis, while restraining bending around the roll and pitch axes, result in bending toward the minor or major groove in the middle section of the DNA, respectively. This direction of bending is confirmed by measuring helical parameters, specifically, the width of the grooves; contracting one groove width while expanding the other indicates bending toward the respective groove. Conversely, rotations around the pitch axis bend the helix toward the terminal phosphates of the bottom terminus. See SI Supporting Note 2 and Figure S3 for further discussion on our pulling protocol, Supporting Note 5 for details of the directed bending, and definition of grooves.

The CVs used here have three major benefits over earlier CVs given in the literature:[6,19,20,53] (i) They bias just a few atoms (mass center of 16 P atoms at most), allowing the majority of base pairs and steps to adopt their preferred conformations; (ii) CVs definitions are independent of base-pair formation. In other words, in the event of melting, base flipping, and extruded bases, the present CVs are well-defined; (iii) orientation quaternion CVs, by definition, ensure smooth and continuous rotation transitions, which is crucial for accurate simulation dynamics. Moreover, since they return an optimal rotation, they also

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

optimize the force, leading to a more accurate estimation of free energy.

MD Simulations. We carried out nonequilibrium MD for oligonucleotide duplexes and trinucleotide repeats in A-, B-, and Z-form as shown in Table 1. The initial DNA configurations were either taken from our previous studies[35−37] or built via 3DNA parameters.[57] The simulations were carried out using the PMEMD module of the AMBER v.20 software package with OL15[58] force field as this force field is known to significantly improve the overall description of the sugar−phosphate backbone equilibria, particularly in Z-DNA.[37,59,60] The TIP3P water model[61] was used for the explicit waters. The simulations were performed with periodic boundary conditions in a truncated octahedral water box. To neutralize the nucleic acid charges, an appropriate number of Na^+ ions were added with standard ion parameters.[62] A 150 mM extra NaCl salt was also added to mimic the physiological conditions. An additional amount of salt was considered for some simulations according to Table 1. Electrostatic interactions were computed with the PME method[63] with a 9 Å cutoff. The cutoff for van der Waals interactions was set as 9 Å as well. Production runs used Langevin dynamics, with a coupling parameter 1.0 ps^{-1} was used. The SHAKE algorithm[64] was applied to all bonds involving hydrogen atoms. In order to obtain starting conformations for nonequilibrium MD calculations, we used our previously established protocol[36,37,65] to minimize and equilibrate helices, as described in the SI (Supporting Note 6).

Table 1. List of the DNA Models Studied

model	form	salt (NaCl) concentration (M)
$(CG)_7C$	A, B, Z	0.2
$(CG)_{15}$	B	0.2
$(CG)_7C$	Z	0[a]
$(CCG)_5$	B	0.2
$(GCC)_5$	B	0.2
$(CGG)_5$	B	0.2
$(GGC)_5$	B	0.2
$(CGG)_5$	Z	5.0
$(GGC)_8$	Z	5.0

[a]All models are neutralized by the addition of Na^+ ions. The mismatches are highlighted in bold font.

Free-Energy Calculations. We used three different methods to calculate the free energy of bent DNA: Umbrella Sampling (US),[21] Steered MD (SMD),[66,67] and Adaptively Biased Molecular Dynamics (ABMD).[68,69] As all of these methods are well-known, only the parts relevant to our simulation will be mentioned. For each simulation, the bending angle was changed from $\theta = 0°$ to $\theta = 150°$ at a constant volume with explicit waters with both neutralizing sodium cations and additional salt (NaCl) at 300 K. As will be discussed, the SMD method was used for most production runs as it is faster computationally while retaining sufficient accuracy.

Umbrella Sampling (US). For the US[21] simulation, bending was induced by adding a biasing potential along the CV (Orientation Angle) with a force constant of 1 kcal/mol $degree^2$. Windows were sampled every 2° for 7 ns each (with the first 2 ns ignored in the final calculations). The free-energy profile was constructed via the weighted histogram analysis method[70,71] (WHAM).

Steered MD (SMD). In SMD[66,67] simulations, a time-dependent external force in terms of CV is applied to induce changes. Work is used to reconstruct the free-energy profile in this process, however, many iterations may be required to estimate free energy accurately.[50,72−75] To benchmark the SMD simulations, we carried out three sets of simulations at different pulling velocities: 100 simulations lasting 10 ns (fast), 20 lasting 50 ns (slow), and 10 lasting 100 ns (slowest), with a harmonic constant set to 100 kcal/mol. The free energy was estimated from work using the Jarzynski equality.[76] After benchmarking our free energy calculations, we found it most efficient to use the SMD method in the slow mode for the majority of our production runs.

Adaptively Biased Molecular Dynamics (AMBD). ABMD[68,69] method is a nonequilibrium MD method that belongs to the general category of umbrella sampling methods with a history-dependent biasing potential. We used the multiple walker[77] well-tempered (WT)[78] extension due to its effectiveness. Using this method, we calculate two types of free-energy profiles: one-dimensional (1D) profiles for free bending (Orientation Angle) and two-dimensional (2D) profiles for directional bending (Spin Angle) along the pitch and yaw axes (Figure 1). Multiple walker WT ABMD runs were carried out using eight replicas, with a total simulation time of 600 ns per replica. Details of these simulations are described in SI (Supporting Note 7).

The mean values and the error analysis were calculated from the profiles generated by bootstrapping, as discussed in SI (Supporting Note 8). Conformations and trajectories were analyzed using the CPPTRAJ[79] program of the AMBER simulation package and VMD.[80] Helical parameters were examined using the 3DNA[57] and CPPTRAJ. Averaging of the parameters was done after bias elimination using Boltzmann averaging. However, it should be noted that the unweighted estimator might be more efficient.[49]

RESULTS

Establishing a Protocol for DNA Bending. As discussed in the previous section, we evaluated the free energy of bent DNA by three different methods: US, ABMD, and SMD. However, because our bending investigations are computationally intensive, we first establish the simulation protocols to be used. It turns out that 'slow' SMD simulations are an effective way to study DNA bending because of their simplicity and reduced simulation costs, all while maintaining an accuracy comparable to the other methods.

To benchmark this, Figure S5a shows the free-energy changes obtained for free bending (Orientation Angle) up to 150° of $(CG)_7C$ B-DNA using US, ABMD, and SMD methods. For SMD, results are shown for three different speeds and after the results are averaging for a different number of independent replicas. Convergence of the simulation results is confirmed: the overlap of the US distributions in consecutive windows indicates convergence of US simulation results (SI Figure S5b). We view the free-energy profile obtained from the US as our reference. Several work profiles from the multiple SMD runs are shown in SI Figure S5c as well. The free-energy profile from SMD simulations is dominated by the trajectories corresponding to small dissipative work. Snapshots of free-energy profiles obtained from regular intervals of the last 50 ns of ABMD runs are given in SI Figure S5d as well. All the free-energy curves shown in Figure S5 are similar and the following general characteristics are noted. Similar to the previous studies,[6,19,20,81] the overall shape of the free-energy curves is parabolic up to around 50°. Beyond that, the curve flattens out to an almost

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

linear increase in free energy with the bending angle. The global free-energy minimum occurred at 15°. Bending up to 50° requires the overall free-energy change of 5 kcal/mol, which is consistent with experimental results from Atomic Force Microscopy measurements.[11] For bending up to 150°, an overall US free-energy change is 21 kcal/mol, which is 9 kcal/mol higher than previously obtained values,[81] using a different collective variable and force field. It is important to keep in mind that free bending (Orientation Angle CV) evaluates the bending angle in relation to a reference configuration (similar to the case of an RMSD CV). In this particular case, the reference configuration is the equilibrated configuration with zero bending and has a nonzero free energy of about 5 kcal/mol. In turn, the lowest free energy is obtained at a slight bending of about 15° and is due to its entropic advantage.

Comparing the ABMD and SMD results to the US curve (Figure S4), we find that these curves are quite similar and nearly identical up to a bending angle of about 30°. However, as the bending angle increases, the fast SMD (v = 15 deg/ns) produces more dissipative work, leading to an overestimation of the free-energy difference between the minimum and 150° by about 5.5 kcal/mol. These results improve with slower SMD simulations and an increased number of independent replicas, as summarized in Table 2. The ABMD free-energy results are

Table 2. Overall Change in the Free Energy of Bending as Obtained from Different Nonequilibrium Simulation Methods[a]

method	$\Delta F_{\min\rightarrow 150^\circ}$ (kcal mol^{-1})	type
SMD	27.17 ± 5.8	fast
	23.69 ± 1.9	slow
	20.05 ± 1.3	slowest
ABMD	20.23 ± 0.8	–
US	21.61 ± 0.7	–

[a]The standard deviation was used as a measure of error in the analysis.

similar to US and SMD (slow or slowest) but with a higher computation cost (see Supporting Note 7 for details). Overall, the agreement between all free-energy methods is quite good (assuming 'slow' SMD); the largest deviations occur at very high bending angles due to local melting and bond rupturing effects. Therefore, from now on we focus on results obtained with the slow SMD protocol due to its ease of implementation and computational efficiency.

Bending $(CG)_7C$ DNA in Its A-, B-, and Z-Forms. We now discuss the DNA bending results for the different helical forms. Bending is carried out using the quaternion methodology in two different ways: free bending (FB) in which the duplex is allowed to bend in any direction and thereby change its orientation and directed bending (DB) in which the duplex is bent in a specified direction and remains on a specified plane. It should be noted that in equilibrium, the transition barrier between A-DNA and B-DNA is low,[82] while the transition barrier from B-DNA to Z-DNA is quite high.[83] We performed our simulations under nonequilibrium conditions with restraints applied to the top and bottom sections of the DNA. The quantitative values that follow are valid only within the framework of the OL15 force field[58] used throughout this study. Empirical nucleic acid parameters are a moving target; new refinements that rebalance sugar-pucker and A/B populations appear regularly. OL15 was selected because it markedly improves sugar−phosphate backbone equilibria, particularly in Z-DNA,[37,59,60] yet it is known to overstabilize the B-form and to give an unsatisfactory description of the A/B equilibrium.[84,85] We nevertheless retained the A- and Z-DNA data for completeness and to provide a benchmark set for future force field tests.

Free-Energy Profiles. The potential of mean force (PMF) associated with FB (i.e., the Orientation Angle CV) is shown in Figure 2 while Figure 3 shows snapshots of the different helices

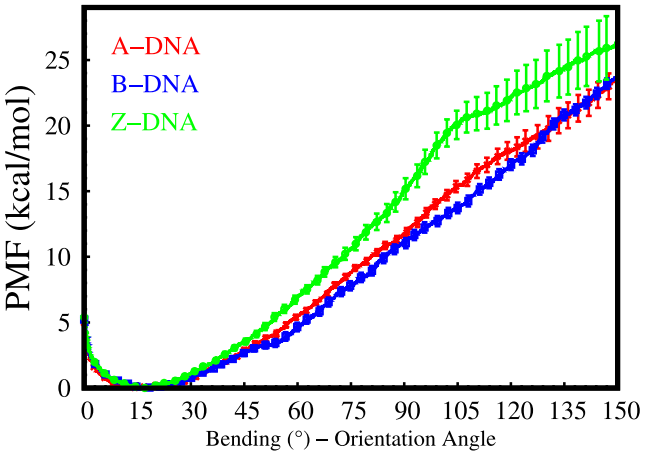


Figure 2. Free energy comparison of free bending of $(CG)_7C$ in A- (red), B- (blue), and Z-form (green) using the 'Orientation Angle' as the CV. The bars show the standard error.

at different bending angles. The free-energy curves display a parabolic behavior between zero and about 45°, and the results are nearly identical for all of the three helical forms. As mentioned previously, all curves have a free energy of about 5 kcal/mol at zero bending, which then decreases to zero at about a 15° bend. The parabolic free energy is followed by an almost linear rise after 45° for A- and B-DNA. The initial decrease in free energy for all three helices is largely due to entropic effects associated with FB, as there is a large number of almost degenerate structures at bending angles around 15° but relatively few such structures at bending angles close to zero. The bending energies for all three helices from the minimum up to about 50° are similar, about 4 kcal/mol. It is also natural that the two right-handed helices, A- and B-DNA, display free energies close to each other. Observations of sample conformations suggest the transition of A-DNA helices to the B-form. In addition, A- or B-DNA helices remain intact for short bending angles, while disruptions become evident for large bending angles of about 120° or more. Left-handed Z-DNA, on the other hand, is stiffer and costs more energy to bend: its free-energy curve rises parabolically until 90−105°, where it displays a shoulder, after which it increases in an approximately linear fashion. This behavior is associated with an unwinding of the helical structure and is evident in the conformations shown in Figure 3.

Turning to DB, by restraining the roll and either pitch and yaw principal axis and rotating along the other axis, it becomes possible to direct the bending toward the major and minor grooves of the DNA helix. The results are shown in Figure 4 and are discussed as follows. First, the bending free energies for both major/minor grooves are relatively close to each other, at least until about 100°, after which deviations in energy costs are evident, especially for Z-DNA. DB also costs more free energy than FB, as DNA is constrained to bend only to a single direction

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

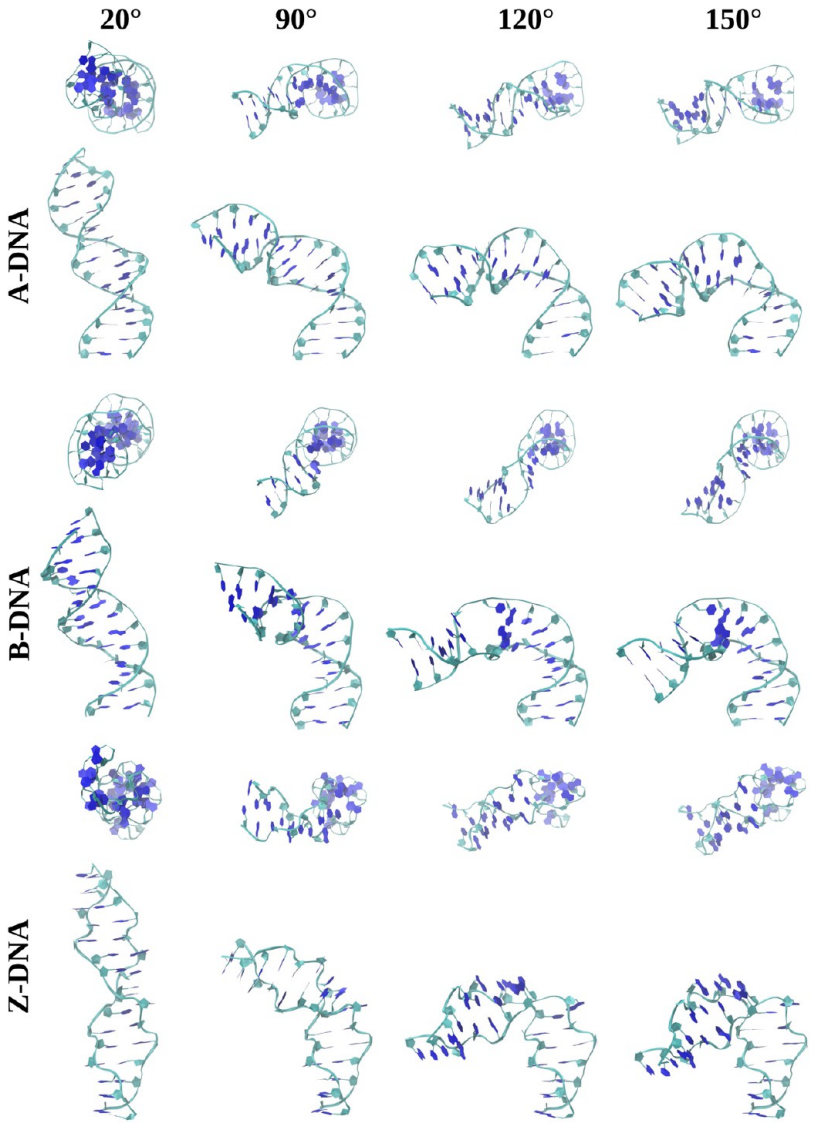


Figure 3. Sample configurations of bent $(CG)_7C$ in A-, B-, and Z-form using the 'Orientation Angle' as the CV. Top and side views are shown.

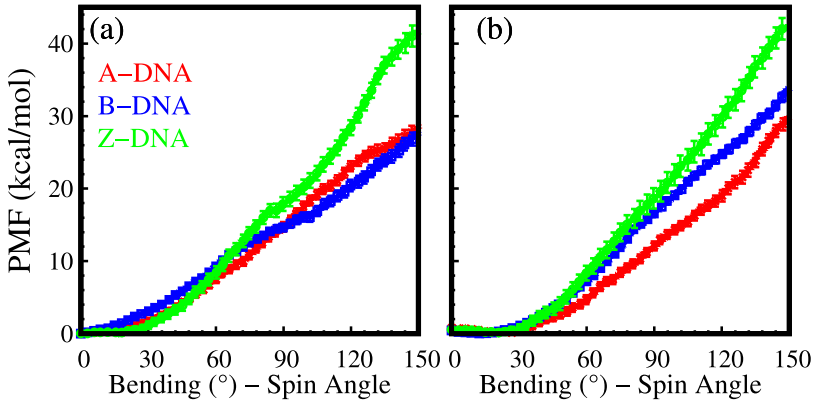


Figure 4. Directed bending of $(CG)_7C$ in A-, B-, and Z-form using the positive and negative directions of the yaw principal axis as the axis of rotation. (a) A clockwise rotation around yaw axis leads to a bending toward the minor groove. (b) Counterclockwise rotation along the yaw axis results in bending toward the major groove. Colors: A-DNA (red), B-DNA (blue), and Z-DNA (green). The bars show the standard error.

without unwinding. While these costs are modest for A- and B-DNA, they are larger (about 10−15 kcal/mol) for Z-DNA. A- and B-DNA exhibit similar free-energy profiles for bending toward minor grooves, but they show deviations when bending toward major grooves. Z-DNA requires more energy to bend, but there appears to be no preference for a given DB direction. This is because the minor groove in Z-DNA is deep and narrow, while the major groove is nearly nonexistent. Finally, we note

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

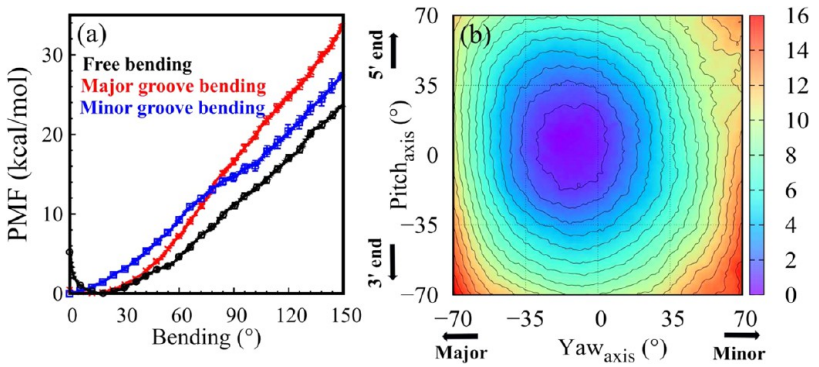


Figure 5. Comparison of free (θ) and directed bending (ϕ) in $(CG)_7C$ B-DNA: (a) free-energy curve for free bending (black), major groove bending (red), and minor groove bending (blue) and (b) two-dimensional free-energy map for yaw and pitch axes of rotation.

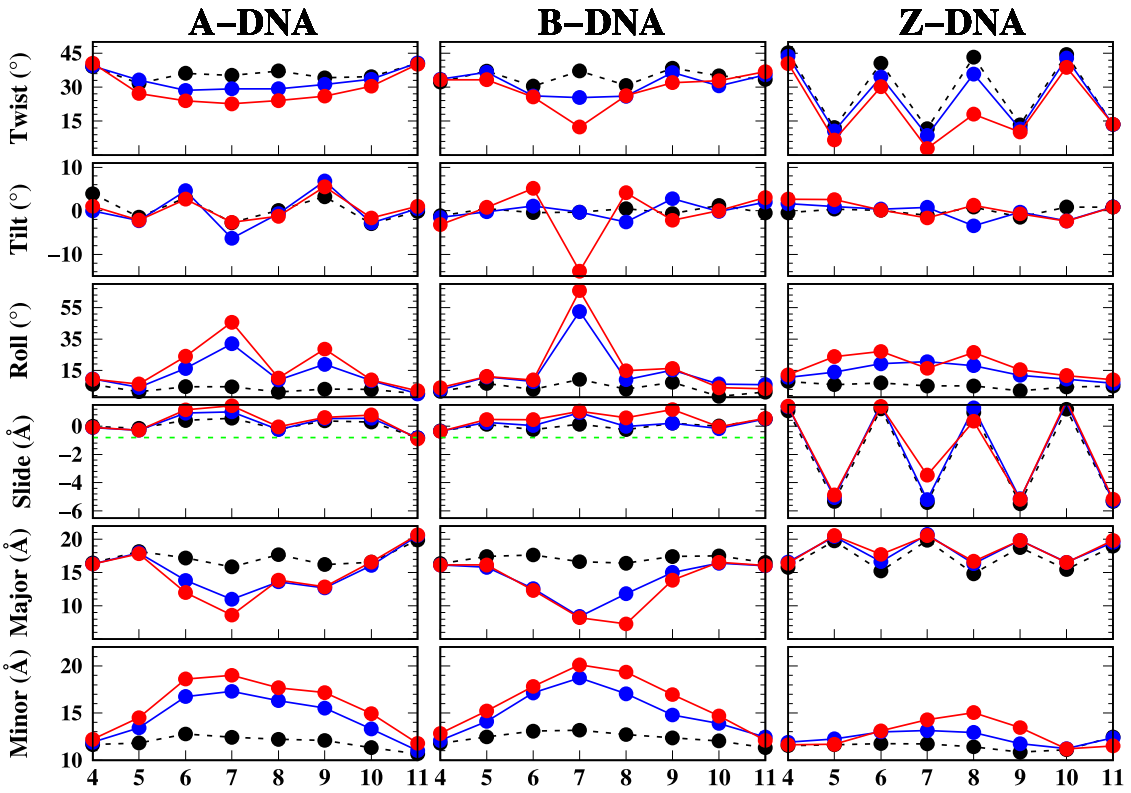


Figure 6. Structural characteristics of double helices: average twist, tilt, roll, slide, major and minor grooves width for the middle base-pair steps of $(CG)_7C$ in A-, B-, and Z-form for three bending angle regions (weak: 10−30° (black color), medium: 80−100° (blue color), and strong: 110−130° (red color)). All bending is restricted toward the major groove. Considered here, are nucleotides 4−12 on one strand and the complementary nucleotides 19−27 on the other, corresponding to steps 4−11 with a pattern of GpC, CpG. The green dashed line in the slide panel points out the value of −0.8 Å. Under equilibrium conditions, slide smaller than −0.8 Å corresponds to A-form helices, while values larger than −0.8 Å are related to B-forms.[98]

that in contrast to FB, DB has its minimum at 0° and remains quite small until about 30°.

To gain further insight into the bending process, Figure 5(a) presents an explicit comparison of the FB and DB results for bending toward the major and minor axes in B-DNA. First, we emphasize that the two collective variables, $\theta = 2\cos^{-1}(q_0)$ and $\phi = 2\tan^{-1}(\vec{q} \cdot \vec{e}/q_0)$, used in the FB and DB procedures, respectively, are equal only when $\vec{q}$, the vector part of the quaternion, aligns with the axis of rotation, $\vec{e}$, or when θ and ϕ are exactly zero, i.e., in the absence of rotation (see Supporting Note 3 for proof of this). The plot reveals that for small bending angles up to about 75°, FB resembles bending toward the major groove; for larger angles, the behavior reverses, with a crossover favoring the minor axis. This effect is due to a small asymmetry in the B-DNA helical structure, i.e., groove width. The difference in the minimum points of the free-energy profiles in the FB and DB procedures is due to the nature of the two collective variables—orientation angle and spin angle—used in these procedures.

Figure 5(b) shows the free-energy landscape as a function of bending toward the yaw and pitch principal axis (as obtained with the ABMD method). The plot shows that the preferred bending angle is somewhere between the two principal axes; while bending toward the pitch axis is symmetric, bending toward the yaw axis is asymmetric. This indicates a small but clear preference for bending toward the major groove at small angles.

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

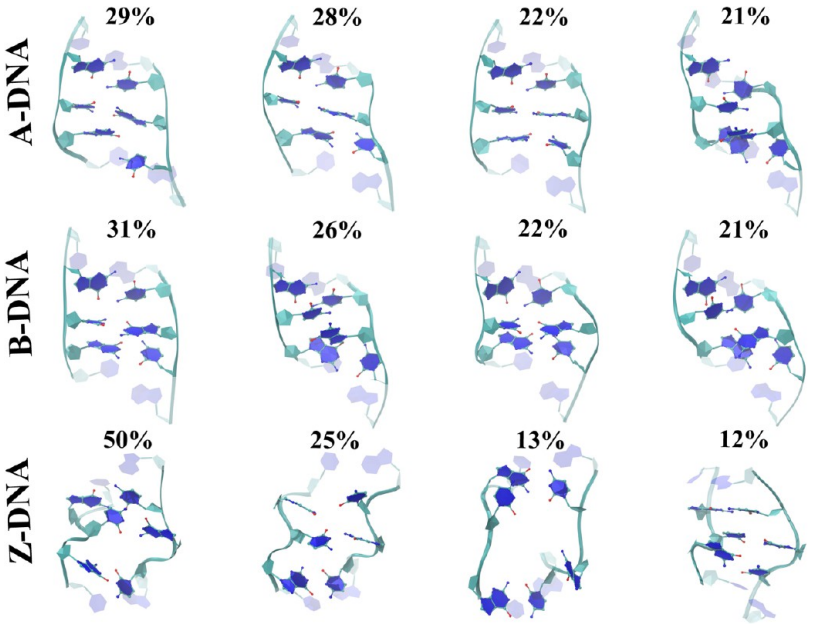


Figure 7. Representative structures for the middle bases of $(CG)_7C$ in A-, B-, and Z-form. The middle bases correspond to residues 7−9 on one strand and complementary residues 22−24 on the other strand. The percentages indicate the occurrence of the conformations represented by the shown structures at a high range of bending angle (110−130°) toward major groove.

Structural Characteristics of Bent DNA. We analyzed the structural characteristics of the bent DNA in terms of its local helical properties. Our analysis is restricted to the eight middle base pairs, which are free to move during the bending process. Base-pair/step descriptors are quantitatively reliable only while the helix is largely intact; thus, we treat them rigorously only for small bending angles ($\lesssim$50°). Beyond that range, where base pairs may be disrupted or sharp kinks appear, we quote these values qualitatively and rely on the global metrics discussed below. Their role here is simply to confirm the direction of the imposed bend and to illustrate broad structural trends. Bending can be divided into three regimes termed moderate (10−30°), high (80−100°), and severe (110−130°) given the bending angle. Results for restrained DB toward the major groove are shown in Figures 6 and S6 for the minor groove.

Bending toward the major groove is evidenced by the collapse of the major groove and expansion of the minor groove, particularly for the A- and B- forms (Figure 6). The reverse phenomenon occurs when bending toward the minor groove (SI Figure S6). This is in agreement with the previous RNA bending work,[86] which suggests that the collapse of the groove width can be the mechanical trigger for bending in equilibrium. Additionally, the transition from A-DNA to B-DNA is evidenced in terms of the slide and groove width even for small bending angles. On the other hand, Z-DNA shows a high degree of rigidity when subjected to bending in either the major or minor groove, signifying a distinct behavior compared to the A- or B-forms. Bending perturbs the base-pair parameters and the base-pair step parameters of the middle bases, but the effects on the other bases are small. Twist, tilt, and roll all describe the relative orientations of the steps. The average twist decreases when DNA bends, indicating helical unwinding. The tilt angle is also disrupted, particularly in A- and B-DNA, but the changes are less pronounced in Z-DNA. Positive roll accompanies bending toward the major groove, whereas negative roll tracks bending toward the minor groove. Roll thus records the direction of bending but is not its microscopic cause; it is simply a geometric consequence of the underlying deformation. It is interesting that the central steps of the A- and B-form show a substantial positive roll, suggesting discontinuous bending or kinks of type I. The significant roll deformation is associated with CpG steps, where the stronger GpC steps are more resistant to bending. The high positive roll is connected with the narrow major grooves.[87] We note that distinct backbone conformers can share nearly identical base-pair/step values; classification such as NtCs[88] would separate those states, but that additional analysis is complementary and lies outside the scope of the present work.

The conformational changes that occur during bending may also be assessed using the concept of handedness which was introduced in ref 83 (the definition of which can be found in the SI, Supporting Note 9, and Figure S7) and the radius of gyration, particularly when the base-pairs undergo severe disruptions. By definition, handedness is positive for right-handed A-DNA and B-DNA, zero for a nonhelical duplex, and negative for left-handed Z-DNA. The histograms of work versus these quantities are given in SI, Figures S8−S9. For the middle bases of all DNA forms, there is an increase in the radius of gyration with increasing bending angle, especially for configurations that require a lower amount of work. In other words, the middle section of DNA expands as it bends, although the two ends of DNA get closer, and the total radius of gyration decreases (not shown). While all forms unwind, Z-DNA remains left-handed and undergoes less evident untwisting and expansion because it requires more work for a given value of handedness and radius of gyration, which underscores its rigidity. The spread of the distributions in these maps indicates a significant amount of structural variability.

We now turn to the dynamic fluctuations of the double helices, which were analyzed in terms of Principal Component Analysis (PCA). We applied the formalism to all the heavy atoms during the bending from 10−130°, for bending toward the major groove. See SI for the details of calculations. The results indicate that the first four PCs account for more than 95% of the fluctuations. A visual examination of the trajectories

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

generated by the PCA reveals that the primary motion is related to the bending mode, which is unsurprising given that we are biasing the DNA structures to bend. The remaining modes correspond to other degrees of freedom that are generally not sampled during our biased sampling. Specifically, the second mode corresponds to tilting in a direction orthogonal to the bending direction, while the third mode highlights bending in the middle section of DNA and distortions corresponding to base opening, staggering, and propeller motion. Lastly, the fourth mode represents stretching. All of these dynamical fluctuations are in line with the structural characteristics already discussed.

To gain further insight into the microscopic aspects of the bent DNA, we have also probed the van der Waals energy and number of hydrogen bonds; the results are given in SI Table S4. As may be expected, as the bending angle increases, both quantities decrease for all the duplexes. Given that the average van der Waals energy is a measure of the stacking for the middle steps, it is clear that bending reduces the bases' ability to form preferred stacking stabilizing the double helical structures. The same is true for the hydrogen bond number. However, we found no significant correlation between backbone torsion angle change and bending in right and left-handed DNAs, indicating that backbone dihedral angles are stable under bending.

Generally speaking, for bending up to 100°, the helices remain unharmed. Above this threshold, localized kinks (one or two steps), wobbling/melting (a series of successive steps), flipped-out bases, or other disruptions were observed in different locations near the bends. We have organized the bent structures of the central bases in the highly bent regime toward the major groove into four groups based on the PCA data. The details of the clustering analysis are provided in SI (Supporting Note 10). Representative configurations and respective populations of these clustered groups are depicted in Figure 7. In the case of A- or B-DNA, the first cluster is characterized by configurations that have a high roll (~20−80°) and a broken base pair. The second and third clusters resemble a type I kink, with a large roll and the unstacking of a single base pair with minimal disruption of neighboring bases. The fourth cluster, on the other hand, represents conformations in which the central base pair is broken and the bases are on top of each other. For Z-DNA, the first cluster also exhibits a kink with a large roll angle. The second and third clusters display bending rigidity toward either the major or minor groove, thus differing from A- and B- forms. Both clusters display broken base pairs with more intense distortion in the third cluster (local melting). The fourth cluster is less distorted but unwound. These configurations show explicitly that DNA bending can result in various conformations, including type I kinks, broken base pairs, unstacked bases, and local melting (type II kinks).

Ion Distribution and Impact of Length on Bending. We also investigated the ion distribution around the bent $(CG)_7C$ DNA structures, and found that the cations (Na^+) from the surrounding solution gather on the concave side, screening the repulsion between phosphate groups and reducing resistance to further DNA bending, which is consistent with previous work.[89] In terms of the atomic structure, the sites associated with the O2, N3, N2 (in the minor grooves), and O6 (in the major grooves) and backbone atoms OP1,OP2 in A- and B-form (see SI Figure S10 for an atomic depiction) are able to localize the cations to varying degrees. Figure 8 gives a visual representation of the ion densities around the DNA structure under moderate (20°) and high (120°) bending angles toward the major groove. Given the

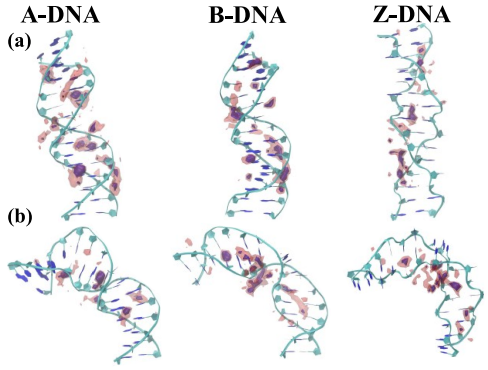


Figure 8. Ion cloud densities around the $(CG)_7C$ duplexes in two regions of bending: (a) weak (20° toward major groove, top panel) and (b) strong (120° toward major groove, bottom panel). The blue regions show volumes with a high ion density, while the red regions indicate areas of relatively low ion densities.

sensitivity of Z-DNA toward the presence of additional salts, we have also calculated the DB of Z-DNA toward the major groove with and without 0.2 M additional salt as shown in Figure 9b. The free energy is roughly 2−3 and 5−7 kcal/mol higher in the low and moderate bending regions, respectively, in the absence of additional salt (red curve), than in the 0.2 M NaCl case (black curve). Our data in the high bending range is too noisy to draw any firm conclusions.

Finally, we have also probed the length dependence of the FB of DNA. Figure 9a shows the corresponding free-energy curves. The results suggest that DNA bending is length-dependent, particularly for moderate-to-high bending angles. The results show that longer DNA segments tend to bend more easily compared to shorter segments, although this effect is modest. Most probably, this is because longer DNA segments have more degrees of freedom and can adopt a wider range of configurations, making them more flexible and slightly easier to bend.

Bending CCG and GGC Trinucleotide Repeat DNA. We now examine the effects of bending on DNA strands with CC and GG mismatches, which are characteristic of trinucleotide repeat duplexes. The bending of six trinucleotides listed in Table 1 were examined and results were compared to their canonical counterparts. It is important to remember that some of these sequences form complicated structures with extruded base pairs. Thus, $(GCC)_5$ (CpG steps between mismatches) forms a so-called e-motif structure in which the CC mismatches symmetrically flip out of the minor groove when the helix is in its B-form, while the $(CCG)_5$ (GpC steps between mismatches) remains stable with the mismatches remaining inside the helical core in an anti−anti conformation.[35] The G-rich sequences, i.e., $(CGG)_5$ and $(GGC)_5$ can exist in both the B- and Z-DNA forms. In the B form, the GG mismatches remain in the helical core with an anti-syn formation for both the CGG and GGC reading frames. Similarly, the GG mismatches also prefer to remain inside the helical core for CGG Z-DNA (GpC steps between mismatches), although the helix tends to stretch.[37] On the other hand, the GG mismatches are extruded for GGC Z-DNA (CpG steps between mismatches) and the structure favors the formation of pseudo-GpC steps.[37] We investigated the bendability of GGC Z-forms having longer lengths, as this particular structure demonstrates stability when featuring 8 or more mismatches. In addition, both GGC and CGG Z-forms require a high salt concentration (5.0 M) to maintain stability.[37]

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

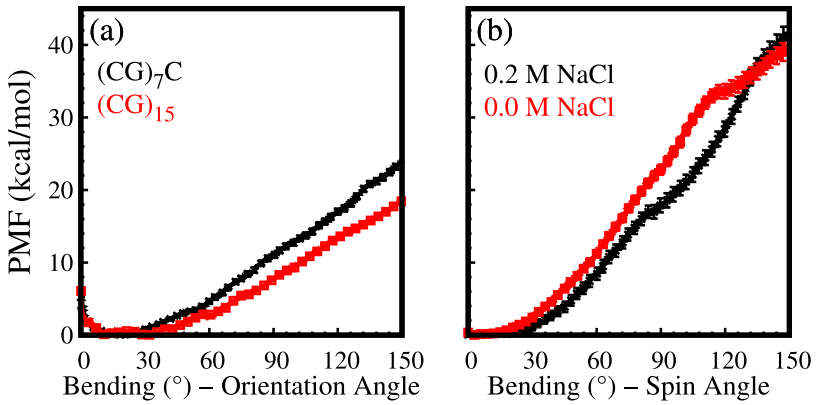


Figure 9. (a) A comparison of free bending for two different lengths: $(CG)_7C$ (black) and $(CG)_{15}$ (red) B-DNA. (b) A comparison of bending $(CG)_7C$ Z-DNA toward the major groove for two different additional salt concentrations: 0.2 M (black) and 0.0 (red). The bars indicate the standard error.

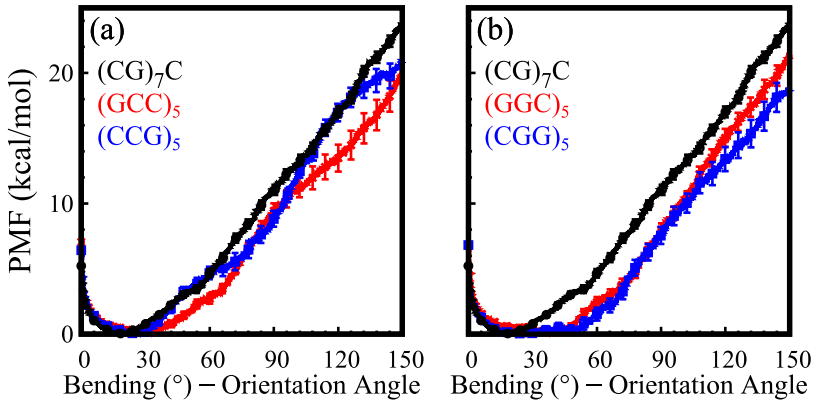


Figure 10. Bending free-energy profile of C-rich and G-rich B-form using the 'Orientation Angle' as the CV (free bending). (a) $(GCC)_5$ (red), $(CCG)_5$ (blue), and $(CG)_7C$ (black); (b) $(GGC)_5$ (red), $(CGG)_5$ (blue), and $(CG)_7C$ (black). The bars show the standard error.

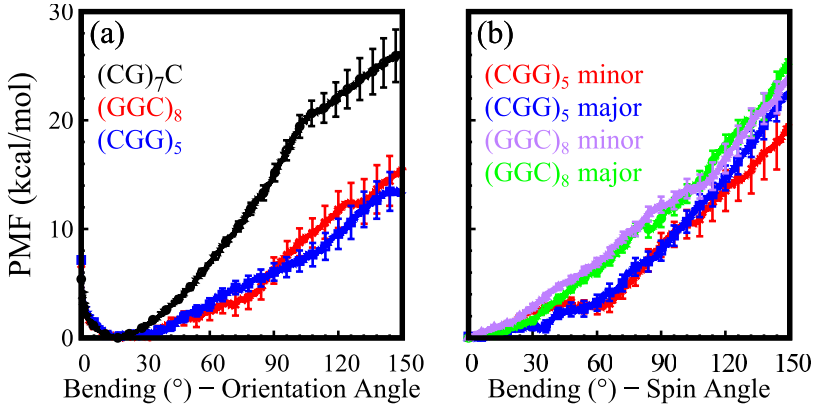


Figure 11. Bending free-energy profile of $(GGC)_8$, $(CGG)_5$, and $(CG)_7C$ Z-form using (a) the 'Orientation Angle' as the CV (free bending), (b) the 'Spin Angle' as the CV (directed bending). (a) Free bending: $(CGG)_5$ (blue), $(GGC)_8$ (red), and $(CG)_7C$ (black); (b) Directed bending: $(CGG)_5$ toward minor groove (red), $(CGG)_5$ toward the major groove (blue), $(GGC)_8$ toward the minor groove (purple), and $(GGC)_8$ toward the major groove (green). The bars show the standard error.

Free-Energy Profiles. FB free-energy curves for trinucleotide repeat structures that are C- or G-rich are shown in Figure 10 (for the B-form) and 11 (Z-DNA). These free-energy profiles resemble their (CG) counterparts: the free-energy minima are located at a nonzero bending angle, the curves are approximately quadratic up to around 50° and then continue to rise in a 'linear' fashion beyond that. Bending free energies for the B-DNA form with mismatches are less than those in ideal B-DNA, indicating that mismatched structures are more flexible and therefore easier to bend. While this effect is only modest for B-DNA, it is considerably larger for Z-DNA indicating that base extrusion considerably reduces the stiffness of the left-handed helices.

In C-rich B-DNA duplexes, the FB free-energy minima for $(GCC)_5$ and $(CCG)_5$ occur at 20 and 32°, respectively, indicating that the equilibrated configurations of these sequences (particularly $(CCG)_5$) bend more than the canonical

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

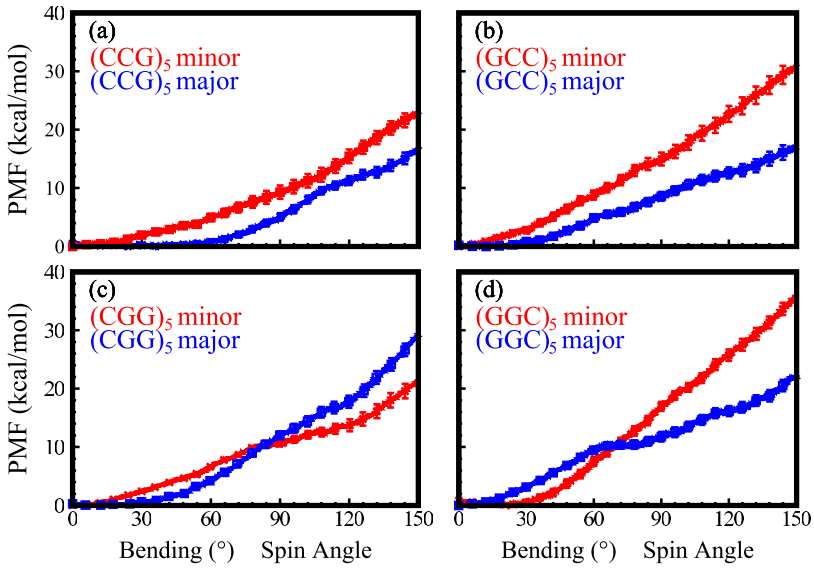


Figure 12. Bending free-energy profile of C-rich and G-rich B-form using the 'Spin Angle' as the CV, associated with restrained bending toward the major (blue) and minor (red) grooves. (a) $(CCG)_5$, (b) $(GCC)_5$, (c) $(CGG)_5$, and (d) $(GGC)_5$. The bars show the standard error.

CG B-DNA. The bending up to 150° requires roughly the same amount of free energy for both reading frames, and about 4 kcal/mol less than the corresponding ideal B-DNA. The bending free-energy curve of $(GCC)_5$ is typically lower than the profile of $(CCG)_5$, whereas the profile of $(GCC)_5$ is more similar to (CG) up to roughly 90°. This is most likely due to the extrusion of CC mismatches in $(GCC)_5$, which forms CpG steps and GpC pseudosteps like those in canonical B-DNA. At high bending angles, we observe a change in slope for both $(CCG)_5$ and $(GCC)_5$ reflecting the considerable deformation of the helices for these angles.

Figure 10b shows the impact of GG mismatches on the FB free-energy profile of $(CGG)_5$ and $(GGC)_5$ in their B-form. The GG mismatches are bulky, but more stable than their CC mismatch counterparts[36] and therefore tend to stay within their helical core. The presence of these mismatches causes the duplexes to naturally bend more, as the minimum free energy is now shifted to 42 and 28°, respectively. Both reading frames have a comparable free-energy profile, with a total free-energy difference of 3 and 5 kcal/mol less than that of the ideal CG B-DNA at 150°.

The FB free-energy profiles for CGG and GGC Z-DNA are shown in Figure 11a. The lowest free energy is found at 22° for both $(CGG)_5$ and $(GGC)_8$ Z-DNA. The free energy required to bend up to 150° for $(CGG)_5$ and $(GGC)_8$ is 13 and 16 kcal/mol, respectively, which is around 13 and 10 kcal/mol less than for regular Z-DNA. The G-rich Z-DNA TRs require the least amount of energy to bend among all TRs studied here. This is primarily due to the extrusion of the mismatched Gs from their helical core. At high bending angles, CGG Z-DNA appears to be more bendable than GGC; it tends to unwind and stretch more.

Using the DB protocol, we explore the preferred direction of bending and the relative stability of $(CCG)_5$, $(GCC)_5$, $(CGG)_5$, and $(GGC)_5$ in their B-form and $(CGG)_5$ and $(GGC)_8$ in their Z-form. Overall, several distinct bending processes were observed, with some structures losing their helical shape and melting. Figure 12 depicts free-energy profiles of C- and G-rich B-form sequences for DB toward the minor and major grooves; this process is sequence-dependent. The total free energy required to bend $(CCG)_5$ up to 150° toward the major groove is less than that to bend toward the minor groove. The same holds for $(GCC)_5$, so that bending toward the major groove is preferred for C-rich sequences. In contrast, the preferred bending direction for G-rich sequences is somewhat angle-dependent. We find that $(CGG)_5$ requires nearly zero free energy for bending toward its major groove up to about 60°, while this is true only for up to 30° for $(GCC)_5$. $(GCC)_5$ requires more work to bend toward the major groove than $(CCG)_5$ up to around 100°. The total free energy for bending these two duplexes up to 150° toward the major groove is similar. Turning to the G-rich sequences, we observe that for $(CGG)_5$ in B-form there is no substantial difference for bending toward either groove. In contrast, $(GGC)_5$ displays crossover behavior with bending toward the major groove preferred at large bending angles. Turning to the Z-form TRs, there is no preferred bending direction in terms of the major or minor groove, as expected (Figure 11b). Both $(CGG)_5$ and $(GGC)_8$ require similar amounts of free energy for DB. However, because $(GGC)_8$ is longer, it is in principle easier to bend than the corresponding $(CGG)_5$.

Structural Characteristics of Bent DNA with Mismatches. We analyzed the structural characteristics and dynamical fluctuations of the TR helices similar to those of the corresponding canonical helices. Turning to the latter, these were analyzed with PCA applied to all heavy atoms. The results indicate that the fluctuations are mostly sequence independent, although the intensity of the motion and base-pair distortions may vary. Our results show that the dynamical fluctuations in the presence of mismatches are somewhat more complex than those of regular DNA, as now the first five−seven PC components are needed to account for 95% of the variance. Similarly to regular DNAs, the first PC, which accounts for more than 65% of the movements, corresponds to the bending mode. The second and third PCs reveal tilting in two opposite directions; both being orthogonal to the bending direction with a numerical value of about 10%. The fourth PC highlights bending and distortions in the middle section of the DNA (about 5%); the remaining PCs, which account for less than 10% of the total motion, are related to stretching and/or twisting.

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

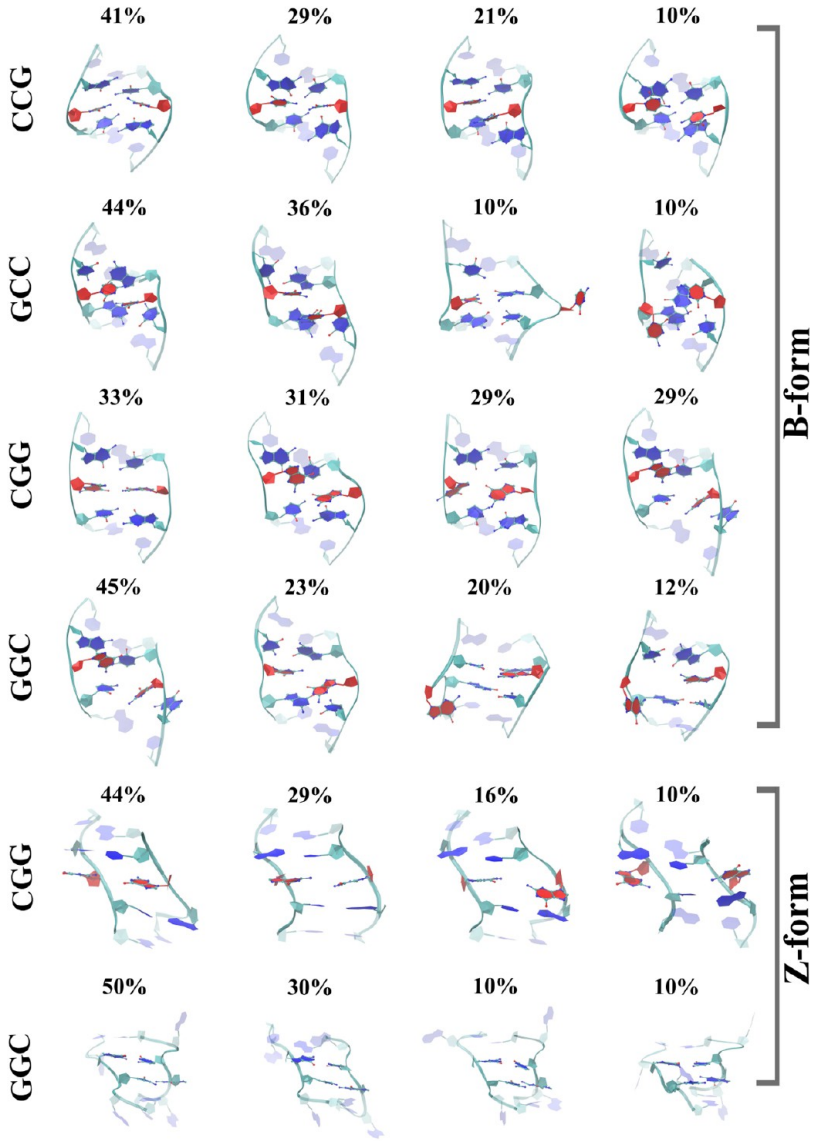


Figure 13. Representative bent conformations for clusters containing CC and GG mismatches. For the B-form conformations, shown are the middle bases of (CCG)5, (GCC)5, (CGG)5, and (GGC)5. For the Z-form, the representatives correspond to GG mismatches and CpG steps in the middle bases of (CGG)5 and (GGC)8. The middle bases correspond to residues 7−9 on one strand and complementary residues 22−24 on the other strand in all structures except $(GGC)_8$, where the middle bases correspond to residues 12,13 on one strand and complementary bases 36,37 on the other. The mismatched bases are shown in red. The percentages indicate the occurrence of the conformations represented by the representative structures at the high range of bending angle (110°-130°) toward the major groove. The numbers have not been reweighted.

We have also examined the atomic structures formed during bending and briefly discuss our observations at high bending angles (roughly 110−130°). Our analysis here focuses on the middle base pairs (two base pairs away from either side of the central base pairs). We have classified the different atomic configurations and the most prominent ones along with their frequency are shown in Figure 13. For the most part, we observe that the distortions at high bending angles consist of base pairs in a type II conformation. The exceptions are the first two clusters of CCG B-DNA; all other B-form clusters correspond to unstacked or flipped-out mismatched bases. Unstacking and base flipping are also common in bent CGG Z-DNA, while mismatches in GGC Z-DNA are extruded and dynamic. Here, structures with CpG steps are relatively stable, although they lose stacking with their neighboring pseudo-GpC steps.

While analyzing helical parameters in the presence of mismatches reveals patterns similar to those of the ideal forms (see SI Supporting Note 11 and SI Figure S11), we instead focus on the correlation of global parameters such as handedness, radius of gyration, C1′−C1′ distance, and the base flipping angle (CPD1 and CPD2,[90] see SI Figure S12) with bending. The correlation between these parameters and bending angle is illustrated in the scatter plots found in SI Figures S13−S15. As anticipated, local handedness and radius of gyration decrease and increase, respectively, as the bending angles are increased. We note that the data is quite noisy with considerable fluctuations. Bending tends to affect the C1′-C1′ distance, particularly for large bending angles, and is sequence-dependent (it may either increase or decrease). The plots of CPD1 and CPD2 indicate that base flipping occurs primarily at high bending angles (above 90°), which is consistent with previous findings.[91] Again, the results suggest that the relationship between bending and base flipping is sequence-dependent. Thus, the CGG Z-DNA helix experiences significant base-flip

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

fluctuations even at small bending angles. Finally, the results indicate that bending weakens hydrogen bonding and stacking energy (as seen in SI Table S4). However, there appears to be no significant correlation with the backbone angles (data not shown).

DISCUSSION

The ability of nucleic acids to bend is an important feature, with many biological implications. Bending plays an important role in gene replication, the repair of DNA, its interaction with many proteins, and its storage inside cells and in many other biological processes. Given the importance of bent DNA, we have undertaken a systematic investigation of the free-energy costs and structural characteristics of bent CG-rich DNA in its various helical forms, with and without mismatches.

Free Energy of DNA Bending Was Studied Using Novel Quaternion Collective Variables. The free energy of bent DNA was evaluated by three different methods, US, ABMD, and SMD − with all making use of novel quaternion collective variables. Just as complex numbers expand upon real numbers in two dimensions, quaternions elegantly extend this concept to four dimensions, offering a novel method to calculate three-dimensional rotations that is computationally more efficient. This approach not only proves convenient but also enables the bending of DNA in two distinct ways: the orientation angle or free bending, FB, in which the bending direction can move out of the plane, and the spin angle or directed bending, DB, in which the bending direction remains in the specified plane. Since the quaternion collective variables bias only the mass center of two ends of the DNA molecule, the method allows for the continuous bending of DNA which retains its full conformational flexibility, allowing the base pairs to adopt their preferred conformations. In terms of biasing methods, all three US, SMD, and ABMD approaches yielded similar free-energy curves. However, computationally we found relatively 'slow' SMD to be the most convenient for our investigations of DNA bending.

Free-Energy Curves for Bending A-, B-, and Z-DNA are Similar for Small Bending Angles, but Deviate from Each Other for Large Bending Angles. The free-energy curves for the different CG helical forms are quite similar, especially for small bending angles of up to around 45°. Bear in mind, A-DNA generally appears only under dehydrating or protein-bound conditions, and Z-DNA is favored by high-salt, negative supercoiling, or CpG methylation. In terms of FB, the free energy has a small nonzero value of about 5 kcal/mol at zero bending, which then drops to zero at about 15°, turning upward in a quadratic form to about 45−60°. The initial free energy dip for FB is due to an entropic increase associated with the large number of 'slightly' bent DNA configurations sampled during the simulations. Over this regime, the free-energy curves for the different CG helical forms track each other quite closely. The transition between A- and B-DNA helices resulted in converged free energy curves over the entire bending regime. On the other hand, for larger bending angles, the Z-DNA free-energy curve begins to show significant deviations from the A- and B-DNA curves. The left-handed Z-DNA helices, which tend to unwind at larger bends, are somewhat stiffer than the right-handed DNA helices.

DNA bending is anisotropic, and so we investigated bending toward the major and minor grooves. In contrast to FB, restrained bending toward a direction is characterized by a minimum at a zero bending followed by a harmonic increase in bending costs to about 40° for all the helical forms. For larger angles, we observe deviations between the right- and left-handed helical forms with the highest costs associated with Z-DNA, as in FB. Large angle bending costs for Z-DNA toward either groove are quite similar, as Z-DNA does not distinguish well between the two types of grooves.

Contrasting the FB and DB free-energy results for CG B-DNA bending, we find that FB costs somewhat less than DB. Moreover, there is crossover behavior associated with DB, with bending toward the major groove costing less at small bending angles and more at large bending angles than bending toward the minor groove.

We Have Characterized the Mechanisms Associated with the Bent Helical Forms. The overall pattern observed in bent structures suggests that bending coincides with helical untwisting and stretching. However, bending did not result in a transition from right-handed to left-handed DNA, or vice versa. There was some evidence that bending facilitates the transition from A- to B-DNA, although such a transition takes place, even in the absence of bending. A structural investigation of the duplexes reveals that the curvature is spread evenly over the helices for small to moderate bending angles. Generally speaking, the helices remained whole for bending angles up to about 100°; for larger bending angles, localized kinks, wobbling/melting, flipped-out bases, and other localized disruptions were observed depending on the form and sequence of the duplexes. Both Type I and Type II kinks were observed in our simulations. Type I kinks, with a high roll angle in a single base-pair, were mostly seen in A- or B-DNA, while base opening and flipping and type II kinks were more commonly associated with Z-DNA bending. Bending strongly affects the middle bases around a given kink, but its effects on bases farther away are relatively small. However, it is important to note that the inherent limitations of the force fields could exaggerate the effects of bending on base stacking/unstacking.[92−94] Therefore, the unstacking events observed in our simulations might not fully reflect the behavior of DNA under physiological conditions. While our observations provide valuable insights into the conformational changes during bending, they should be interpreted with caution and viewed primarily as a comparison among the different forms and sequences.

We also studied the ion distribution around the bent structures. For the most part, cations gather on the concave side of the bent helices, where they help screen the repulsion between backbone phosphate groups. This then works to reduce the resistance to further bending.

CCG and GGC Trinucleotide Repeat DNA Show Less Rigidity than the Ideal DNAs. We have studied the bending free energy of triplet repeat DNA based on CCG, GCC, GGC, and CGG duplexes characterized by either CC or GG mismatches. In broad terms, the FB and DB free-energy profiles resemble their CG counterparts for various helical forms. However, bending costs for duplexes with mismatches are less than those of helices without mismatches, showing that mismatched structures are more flexible and easier to bend, which in turn is due to the lower stability of the helices under the presence of mismatches. While this effect is relatively modest for B-DNA, it is larger for Z-DNA indicating that extruded bases (present in GGC Z-DNA) considerably reduce Z-DNA stiffness. This is interesting as DNA flexibility plays an essential role in the recognition and repair of mismatched nucleotides.[3,95−97] In terms of the DB, results are sequence-dependent. For C-rich sequences, bending toward the major groove is preferred, while for G-rich sequences, the bending direction varies for different

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

angle regimes. We note that G-rich TRs in their Z-form are the most flexible in terms of bending. Our induced bending reveals a pattern of hydrogen bond weakening, unstacking, base flipping, local melting, and distortion of helical parameters that is similar to, but not identical to, their matched counterparts.

SUMMARY AND CONCLUSIONS

While the persistence length of approximately 150 base pairs effectively describes DNA's flexibility at larger scales, such as chromatin packaging, its bending behavior at shorter-length scales displays intricate complexity. This fine-grained flexibility is critical for crucial protein−DNA interactions, as exemplified by the mismatch repair system, where precise DNA bending and manipulation are essential for error correction. In this study, we examined the bending free energy of short DNA helices, focusing on various helical forms and trinucleotide repeat sequences with and without mismatches.

By introducing a novel orientation quaternions formalism for the first time into the AMBER simulation package[51] and combining it with a variety of sampling methods, we quantified the energetic costs and structural characteristics of DNA bending. Our method allowed us to bend DNA in distinct ways and revealed quadratic variations in bending free energy for moderate bending, followed by linear trends for larger angles. Notably, left-handed Z-DNA exhibited the highest rigidity among canonical forms. We identified type I and type II kinks depending on sequence and helical form, with high flexibility observed in the presence of mismatches, particularly in Z-form CGG and GGC repeats. The use of quaternion collective variables not only enabled us to explore the orientation of bending but also allowed us to expand our calculations to a particular direction of bending, which sheds light on the mechanics of DNA bending, uncovering the energetic landscape and preferred bending directions. We observed distinct bending behaviors across helical forms and sequence variations, highlighting the role of mismatches in modulating DNA flexibility. Our findings offer new insights into the global conformational flexibility of short-length DNA molecules and have implications for various biological processes, including transcription factor binding and nucleosome formation. Furthermore, our study contributes to the understanding of DNA-repair mechanisms, particularly in the context of trinucleotide repeat expansion disorders, by elucidating the structural consequences of bending and its possible impact on mismatch recognition.

ASSOCIATED CONTENT

Data Availability Statement

The molecular dynamics simulations in this study were conducted using the AMBER software suite. The quaternion-based collective variables have been implemented in AMBER version 2020 and further. The input files and configuration scripts necessary to reproduce the results of this manuscript are available in the associated GitHub repository: https://github.com/afzq89/dna_bending

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.5c00541.

Notes 1−11 (implementation of quaternion-based CVs in AMBER; detailed description of pulling protocols along with sample input files; quaternion−Euler relations; comparison of quaternion-based bending angles with Curves+; measurement of groove widths and other helical parameters; equilibrium MD details; ABMD parameters; free energy and error analysis; definition of handedness; PCA and clustering; analysis of helical parameters in mismatched trinucleotide repeats); Tables S1−S4; Figures S1−S15 (PDF)

AUTHOR INFORMATION

Corresponding Author

Christopher Roland − *Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202, United States;* orcid.org/0000-0003-4562-8136; Email: cmroland@ncsu.edu

Authors

Ashkan Fakharzadeh − *Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202, United States;* orcid.org/0009-0006-9598-1973

Mahmoud Moradi − *Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, Arkansas 72701, United States;* orcid.org/0000-0002-0601-402X

Celeste Sagui − *Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202, United States;* orcid.org/0000-0002-2750-5186

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.5c00541

Author Contributions

A.F. ran simulations and analyzed results; A.F. and M.M. added the quaternion module into AMBER; all authors planned the project and contributed to writing of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work has been supported by NSF 1945465 and MCB-2409309 as well as NIH/NIGMS R35GM147423.

REFERENCES

(1) Vilar, J. M.; Saiz, L. DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. *Curr. Opin. Genet. Dev.* **2005**, *15*, 136−144.

(2) Modrich, P. Mechanisms in eukaryotic mismatch repair. *J. Biol. Chem.* **2006**, *281*, 30305−30309.

(3) Gorman, J.; Chowdhury, A.; Surtees, J. A.; Shimada, J.; Reichman, D. R.; Alani, E.; Greene, E. C. Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. *Mol. Cell* **2007**, *28*, 359−370.

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

(5) Prez, A.; Noy, A.; Lankas, F.; Luque, F. J.; Orozco, M. The relative flexibility of B-DNA and A-RNA duplexes: database analysis. *Nucleic Acids Res.* **2004**, *32*, 6144−6151.

(6) Ma, N.; van der Vaart, A. Anisotropy of B-DNA Groove Bending. *J. Am. Chem. Soc.* **2016**, *138*, 9951−9958.

(7) van der Vaart, A. Coupled binding-bending-folding: The complex conformational dynamics of protein-DNA binding studied by atomistic molecular dynamics simulations. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 1091−1098.

(8) Kratky, O.; Porod, G. Rontgenuntersuchung geloster Fadenmolekule. *Recl. Trav. Chim. Pays-Bas* **1949**, *68*, 1106−1122.

(9) Taylor, W. H.; Hagerman, P. J. Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure: II.

NaCl-dependence of DNA flexibility and helical repeat. *J. Mol. Biol.* **1990**, *212*, 363−376.
(10) Cloutier, T. E.; Widom, J. Spontaneous Sharp Bending of Double-Stranded DNA. *Mol. Cell* **2004**, *14*, 355−362.
(11) Wiggins, P. A.; van der Heijden, T.; Moreno-Herrero, F.; Spakowitz, A.; Phillips, R.; Widom, J.; Dekker, C.; Nelson, P. C. High flexibility of DNA on short length scales probed by atomic force microscopy. *Nat. Nanotechnol.* **2006**, *1*, 137−141.
(12) Mathew-Fenn, R. S.; Das, R.; Harbury, P. A. Remeasuring the double helix. *Science* **2008**, *322*, 446−449.
(13) Du, Q.; Smith, C.; Shiffeldrim, N.; Vologodskaia, M.; Vologodskii, A. Cyclization of short DNA fragments and bending fluctuations of the double helix. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5397−5402.
(14) Travers, A. DNA dynamics: bubble 'n' flip for DNA cyclisation? *Curr. Biol.* **2005**, *15*, R377−R379.
(15) Yan, J.; Marko, J. F. Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys. Rev. Lett.* **2004**, *93*, No. 108108.
(16) Lankaš, F.; Lavery, R.; Maddocks, J. H. Kinking occurs during molecular dynamics simulations of small DNA minicircles. *Structure* **2006**, *14*, 1527−1534.
(17) Harris, S. A.; Laughton, C. A.; Liverpool, T. B. Mapping the phase diagram of the writhe of DNA nanocircles using atomistic molecular dynamics simulations. *Nucleic Acids Res.* **2007**, *36*, 21−29.
(18) Crick, F. H. C.; Klug, A. Kinky helix. *Nature* **1975**, *255*, 530−533.
(19) Curuksu, J.; Zakrzewska, K.; Zacharias, M. Magnitude and direction of DNA bending induced by screw-axis orientation: influence of sequence, mismatches and abasic sites. *Nucleic Acids Res.* **2008**, *36*, 2268−2283.
(20) Sharma, M.; Predeus, A. V.; Mukherjee, S.; Feig, M. DNA bending propensity in the presence of base mismatches: implications for DNA repair. *J. Phys. Chem. B* **2013**, *117*, 6194−6205.
(21) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.
(22) Svozil, D.; Kalina, J.; Omelka, M.; Schneider, B. DNA conformations and their sequence preferences. *Nucleic Acids Res.* **2008**, *36*, 3690−3706.
(23) Schneider, B.; Černý, J.; Svozil, D.; Čech, P.; Gelly, J.-C.; de Brevern, A. G. Bioinformatic analysis of the protein/DNA interface. *Nucleic Acids Res.* **2014**, *42*, 3381−3394.
(24) Pan, F.; Roland, C.; Sagui, C. Ion distribution around left- and right-handed DNA and RNA duplexes: a comparative study. *Nucleic Acids Res.* **2014**, *42*, 13981−13996.
(25) Rich, A.; Nordheim, A.; Wang, A. H. The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.* **1984**, *53*, 791−846.
(26) Oh, D. B.; Kim, Y. G.; Rich, A. Z-DNA-binding proteins can act as potent effectors of gene expression in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16666−16671.
(27) Wang, G.; Christensen, L. A.; Vasquez, K. M. Z-DNA-forming sequences generate large-scale deletions in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2677−2682.
(28) Pearson, C. E.; Edamura, K.; Cleary, J. Repeat instability: Mechanisms of dynamic mutations. *Nat. Rev. Genet.* **2005**, *6*, 729−742.
(29) Mirkin, S. M. Expandable DNA repeats and human disease. *Nature* **2007**, *447*, 932−940.
(30) Kovtun, I. V.; McMurray, C. T. Features of trinucleotide repeat instability in vivo. *Cell Res.* **2008**, *18*, 198−213.
(31) McMurray, C. T. Hijacking of the mismatch repair system to cause CAG expansion and cell death in neurodegenerative disease. *DNA Repair* **2008**, *7*, 1121−1134.
(32) Usdin, K.; Woodford, K. CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA-synthesis in-vitro. *Nucleic Acids Res.* **1995**, *23*, 4202−4209.
(33) Paulson, H.Repeat Expansion Diseases. In *Handbook of Clinical Neurology*; Elsevier, 2018; Vol. *147*, p 105.
(34) Khristich, A. N.; Mirkin, S. M. On the Wrong DNA Track: Molecular Mechanisms of Repeat-mediated Genome Instability. *J. Biol. Chem.* **2020**, *295*, 4134−4170.
(35) Pan, F.; Zhang, Y.; Man, V. H.; Roland, C.; Sagui, C. E-motif Formed by Extrahelical Cytosine Bases in DNA Homoduplexes of Trinucleotide and Hexanucleotide Repeats. *Nucleic Acids Res.* **2018**, *46*, 942−955.
(36) Pan, F.; Man, V.; Roland, C.; Sagui, C. Structure and dynamics of DNA and RNA double helices obtained from the CCG and GGC trinucleotide repeats. *J. Phys. Chem. B* **2018**, *122*, 4491−4512.
(37) Fakharzadeh, A.; Zhang, J.; Roland, C.; Sagui, C. Novel eGZ-motif formed by regularly extruded guanine bases in a left-handed Z-DNA helix as a major motif behind CGG trinucleotide repeats. *Nucleic Acids Res.* **2022**, *50*, 4860−4876.
(38) Fu, Y. H.; Kuhl, D. P.; Pizzuti, A.; Pieretti, M.; Sutcliffe, J. S.; Richards, S.; Verkerk, A. J.; Holden, J. J.; Fenwick, R. G.; Warren, S. T.; et al. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* **1991**, *67*, 1047−1058.
(39) Sherman, S. L. Premature ovarian failure among fragile X premutation carriers: parent-of-origin effect? *Am. J. Hum. Genet.* **2000**, *67*, 11−13.
(40) Santos, C. B.; Lima, M. A. C.; Pimentel, M. M. G. A new PCR assay useful for screening of FRAXE/FMR2 mental impairment among males. *Hum. Mutat.* **2001**, *18*, 157−162.
(41) LaCroix, A. J.; Stabley, D.; Sahraoui, R.; et al. GGC repeat expansion and exon 1 methylation of XYLT1 is a common pathogenic variant in Baratela-Scott syndrome. *Am. J. Hum. Genet.* **2019**, *104*, 35−44.
(42) Gu, Y.; Shen, Y.; Gibbs, R. A.; Nelson, D. L. Identification of FMR2, a novel gene associated with the FRAXE CCG repeat and CpG island. *Nat. Genet.* **1996**, *13*, 109−113.
(43) Braida, C.; Stefanatos, R. K.; Adam, B.; Mahajan, N.; Smeets, H. J.; Niel, F.; Goizet, C.; Arveiler, B.; Koenig, M.; Lagier-Tourenne, C.; Mandel, J. L.; Faber, C. G.; de Die-Smulders, C. E.; Spaans, F.; Monckton, D. G. Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Hum. Mol. Genet.* **2010**, *19*, 1399−1412.
(44) Gao, X.; Huang, X.; Smith, G.; Zheng, M.; Liu, H. New antiparallel duplex motif of DNA CCG repeats that is stabilized by extrahelical basis symmetrically located in the minor-groove. *J. Am. Chem. Soc.* **1995**, *117*, 8883−8884.
(45) Zheng, M.; Huang, X.; Smith, G.; Yang, X.; Gao, X. Genetically unstable CXG repeats are structurally dynamic and have a high propensity for folding. An NMR and UV spectroscopic study. *J. Mol. Biol.* **1996**, *264*, 323−336.
(46) Renčiuk, D.; Kypr, J.; Vorlickova, M. CGG repeats associated with fragile X chromosome form left-handed Z-DNA structure. *Biopolymers* **2011**, *95*, 174−181.
(47) Fiorin, G.; Klein, M. L.; Henin, J. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* **2013**, *111*, 3345−3362.
(48) Moradi, M.; Tajkhorshid. Mechanistic picture for conformational transition of a membrane transporter at atomic resolution. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 18916−18921.
(49) Moradi, M.; Enkavi, G.; Tajkhorshid, E. Atomic-level characterization of transport cycle thermodynamics in the glycerol-3-phosphate:phosphate antiporter. *Nat. Commun.* **2015**, *6*, No. 8393.
(50) Moradi, M.; Tajkorshid, E. Computational recipe for efficient description of large-scale conformational changes in biomolecular systems. *J. Chem. Theory Comput.* **2014**, *10*, 2866−2880.
(51) Case, D. A.; Belfon, K.; Ben-Shalom, I.Y.et al. *AMBER 2020*; University of California: San Francisco, 2020.
(52) Dickerson, R. E. DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.* **1998**, *26*, 1906−1926.
(53) Strahs, D.; Schlick, T. A-Tract bending: insights into experimental structures by computational models. *J. Mol. Biol.* **2000**, *301*, 643−663.

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689

(54) Lankas, F.; Spackova, N.; Moakher, M.; Enkhbayar, P.; Sponer, J. A measure of bending in nucleic acids structures applied to A-tract DNA. *Nucleic Acids Res.* **2010**, *38*, 3414−3422.
(55) Coutsias, E. A.; Seok, C.; Dill, K. Using quaternions to calculate RMSD. *J. Comput. Chem.* **2004**, *25*, 1849−1857.
(56) Blanchet, C.; Pasi, M.; Zakrzewska, K.; Lavery, R. CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Res.* **2011**, *39*, W68−W73.
(57) Lu, X. J.; Olson, W. K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **2003**, *31*, 5108−5121.
(58) Zgarbová, M.; Sponer, J.; Otyepka, M.; Cheatham, T. E., III; Galindo-Murillo, R.; Jurecka, P. Refinement of the sugar−phosphate backbone torsion beta for amber force fields improves the description of Z- and B-DNA. *J. Chem. Theory Comput.* **2015**, *11*, 5723−5736.
(59) Galindo-Murillo, R.; Robertson, J. C.; Zgarbova, M.; Sponer, J.; Otyepka, M.; Jurecka, P.; Cheatham, T. E. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **2016**, *12*, 4114−4127.
(60) Zgarbová, M.; Sponer, J.; Jurecka, P. Z-DNA as a Touchstone for Additive Empirical Force Fields and a Refinement of the Alpha/Gamma DNA Torsions for AMBER. *J. Chem. Theory Comput.* **2021**, *17*, 6292−6301.
(61) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(62) Joung, I. S.; Cheatham, T. E., III Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **2008**, *112*, 9020−9041.
(63) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *Chem. Phys.* **1995**, *103*, 8577−8593.
(64) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(65) Zhang, J.; Fakharzadeh, A.; Pan, F.; Roland, C.; Sagui, C. Atypical structures of GAA/TTC trinucleotide repeats underlying Friedreich's ataxia: DNA triplexes and RNA/DNA hybrids. *Nucleic Acids Res.* **2020**, *48*, 9899−9917.
(66) Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K. *Steered Molecular Dynamics. Computational Molecular Dynamics: Challenges, Methods, Ideas*; Springer-Verlag: Berlin, Germany, 1998; pp 39−65.
(67) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. Free Energy Calculation from Steered Molecular Dynamics Simulations Using Jarzynskis Equality. *J. Chem. Phys.* **2003**, *119*, 3559−3568.
(68) Babin, V.; Roland, C.; Sagui, C. Adaptively Biased Molecular Dynamics for Free Energy Calculations. *J. Chem. Phys.* **2008**, *128*, No. 134101.
(69) Babin, V.; Karpusenka, V.; Moradi, M.; Roland, C.; Sagui, C. Adaptively biased molecular dynamics: an umbrella sampling method with a time dependent potential. *Int. J. Quantum Chem.* **2009**, *109*, 3666−3678.
(70) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(71) Grossfield, A. WHAM: The Weighted Histogram Analysis Method 2012 http://membrane.urmc.rochester.edu/wordpress/?page_id/126.
(72) Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120*, 5946−5961.
(73) Shirts, M. R.; Pande, V. S. Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *J. Chem. Phys.* **2005**, *122*, No. 144107.
(74) Jarzynski, C. Rare events and the convergence of exponentially averaged work values. *Phys. Rev. E* **2006**, *73*, No. 046105.
(75) Moradi, M.; Babin, V.; Roland, C.; Sagui, C. The Adaptively Biased Molecular Dynamics method revisited: New capabilities and an application. *J. Phys.:Conf. Ser.* **2015**, *640*, No. 012020, DOI: 10.1088/1742-6596/640/1/012020.
(76) Jarzynskia, C. Nonequilibrium work relations: foundations and applications. *Eur. Phys. J. B* **2008**, *64*, 331−340.
(77) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics. *J. Phys. Chem. B* **2006**, *110*, 3533−3539.
(78) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **2008**, *100*, No. 020603.
(79) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(80) Humphrey, W.; Dalke, A.; Schulten, K. VMD − Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(81) Curuksu, J.; Zacharias, M.; Lavery, R.; Zakrzewska, K. Local and global effects of strong DNA bending induced during molecular dynamics simulations. *Nucleic Acids Res.* **2009**, *37*, 3766−3773.
(82) Zgarbová, M.; Jurečka, P.; Šponer, J.; Otyepka, M. A- to B-DNA Transition in AMBER Force Fields and Its Coupling to Sugar Pucker. *J. Chem. Theory Comput.* **2018**, *14*, 319−328.
(83) Moradi, M.; Babin, V.; Roland, C.; Sagui, C. Reaction path ensemble of the B-Z-DNA transition: a comprehensive atomistic study. *Nucleic Acids Res.* **2013**, *41*, 33−43.
(84) Zgarbová, M.; Jurečka, P.; Šponer, J.; Otyepka, M. A- to B-DNA Transition in AMBER Force Fields and Its Coupling to Sugar Pucker. *J. Chem. Theory Comput.* **2018**, *14*, 319−328.
(85) Zgarbová, M.; Šponer, J.; Jurečka, P. Refinement of the sugar puckering torsion potential in the AMBER DNA force field. *J. Chem. Theory Comput.* **2025**, *21*, 833−846, DOI: 10.1021/acs.jctc.4c01100.
(86) Taghavi, A.; Yildirim, I. Computational Investigation of Bending Properties of RNA AUUCU, CCUG, CAG, and CUG Repeat Expansions Associated With Neuromuscular Disorders. *Front. Mol. Biosci.* **2022**, *9*, No. 830161.
(87) Marin-Gonzalez, A.; Aicart-Ramos, C.; Marin-Baquero, M.; Martin-Gonzalez, A.; Suomalainen, M.; Kannan, A.; Vilhena, J. G.; Greber, U. F.; Moreno-Herrero, F.; Perez, R. Double-stranded RNA bending by AU-tract sequences. *Nucleic Acids Res.* **2020**, *48*, 12917−12928.
(88) Černý, J.; Božíková, P.; Svoboda, J.; Schneider, B. A unified dinucleotide alphabet describing both RNA and DNA structures. *Nucleic Acids Res.* **2020**, *48*, 6367−6381.
(89) Spiriti, J.; Kamberaj, H.; de Graff, A. M.; Thorpe, M. F.; van der Vaart, A. DNA Bending through Large Angles Is Aided by Ionic Screening. *J. Chem. Theory Comput.* **2012**, *8*, 2145−2156.
(90) Song, K.; Campbell, A. J.; Bergonzo, C.; de Los Santos, C.; Grollman, A. P.; Simmerling, C. An Improved Reaction Coordinate for Nucleic Acid Base Flipping Studies. *J. Chem. Theory Comput.* **2009**, *5*, 3105−3113.
(91) Ma, N.; van der Vaart, A. Free Energy Coupling between DNA Bending and Base Flipping. *J. Chem. Inf. Model.* **2017**, *57*, 2020−2026.
(92) Banáš, P.; Mládek, A.; Otyepka, M.; Zgarbová, M.; Jurečka, P.; Svozil, D.; Lankaš, F.; Šponer, J. Can We Accurately Describe the Structure of Adenine Tracts in B-DNA? Reference Quantum-Chemical Computations Reveal Overstabilization of Stacking by Molecular Mechanics. *J. Chem. Theory Comput* **2012**, *8*, 2448−2460.
(93) Kilchherr, F.; Wachauf, C.; Pelz, B.; Rief, M.; Zacharias, M.; Dietz, H. Single-molecule dissection of stacking forces in DNA. *Science* **2016**, *353*, No. aaf5508, DOI: 10.1126/science.aaf5508.
(94) Häse, F.; Zacharias, M. Free energy analysis and mechanism of base pair stacking in nicked DNA. *Nucleic Acids Res.* **2016**, *44*, 7100−7108.
(95) Jeong, C.; Cho, W. K.; Song, K. M.; Cook, C.; Yoon, T. Y.; Ban, C.; Fishel, R.; Lee, J. B. MutS switches between two fundamentally

distinct clamps during mismatch repair. *Nat. Struct. Mol. Biol.* **2011**, *18*, 379−385.
(96) Cho, W. K.; Jeong, C.; Kim, D.; Chang, M.; Song, K. M.; Hanne, J.; Ban, C.; Fishel, R.; Lee, J. B. ATP alters the diffusion mechanics of MutS on mismatched DNA. *Structure* **2012**, *20*, 1264−1274.
(97) Wang, H.; Yang, Y.; Schofield, M. J.; Du, C.; Fridman, Y.; Lee, S. D.; Larson, E. D.; Drummond, J. T.; Alani, E.; Hsieh, P.; Erie, D. A. DNA bending and unbending by MutS govern mismatch recognition and specificity. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14822−14827.
(98) Lu, X. J.; Shakked, Z.; Olson, W. K. A-DNA Conformational Motifs in Ligand-bound Double Helices. *J. Mol. Biol.* **2000**, *300*, 819−840.

https://doi.org/10.1021/acs.jcim.5c00541
J. Chem. Inf. Model. 2025, 65, 5672−5689