2268–2283 *Nucleic Acids Research, 2008, Vol. 36, No. 7* *Published online 20 February 2008*
doi:10.1093/nar/gkm1135

Magnitude and direction of DNA bending induced by screw-axis orientation: influence of sequence, mismatches and abasic sites

Jeremy Curuksu[1], Krystyna Zakrzewska[2] and Martin Zacharias[1,*]

[1]School of Engineering and Science, Jacobs University, Campus Ring 1, D-28759 Bremen, Germany and [2]Bioinformatique et RMN structurales, Institut de Biologie et Chimie des Proteines, UMR 5086 CNRS, 7 passage du Vercors, 69367 Lyon, France

Received July 12, 2007; Revised December 5, 2007; Accepted December 6, 2007

ABSTRACT

DNA-bending flexibility is central for its many biological functions. A new bending restraining method for use in molecular mechanics calculations and molecular dynamics simulations was developed. It is based on an average screw rotation axis definition for DNA segments and allows inducing continuous and smooth bending deformations of a DNA oligonucleotide. In addition to controlling the magnitude of induced bending it is also possible to control the bending direction so that the calculation of a complete (2-dimensional) directional DNA-bending map is now possible. The method was applied to several DNA oligonucleotides including A(adenine)-tract containing sequences known to form stable bent structures and to DNA containing mismatches or an abasic site. In case of G:A and C:C mismatches a greater variety of conformations bent in various directions compared to regular B-DNA was found. For comparison, a molecular dynamics implementation of the approach was also applied to calculate the free energy change associated with bending of A-tract containing DNA, including deformations significantly beyond the optimal curvature. Good agreement with available experimental data was obtained offering an atomic level explanation for stable bending of A-tract containing DNA molecules. The DNA-bending persistence length estimated from the explicit solvent simulations is also in good agreement with experiment whereas the adiabatic mapping calculations with a GB solvent model predict a bending rigidity roughly two times larger.

INTRODUCTION

The structure and flexibility of DNA is important for its many biological functions including recognition by proteins, DNA repair, packaging and transient melting during transcription and replication (1–9). The investigation of the sequence-dependent bending deformability of nucleic acids is of particular interest since in many protein–DNA complexes the DNA adopts a curved or bent structure (10). Also, certain nucleotide sequences lead to intrinsic curvature, and this property can thus be an essential component in the recognition process (indirect readout) (11,12). For example, tracts of 3–6 consecutive Adenine (A) nucleotides repeated in phase with the helical repeat of DNA (A-tracts) lead to global curvature, which can cause an unusual slow migration during electrophoresis in an acrylamide gel (13–16). The reduction of electrophoretic mobility compared to a straight DNA of the same length has been used to quantify the sequence dependence of bending and for a single A-tract bend angles of 17°–21° have been measured (13). Both A-tract length and the nature of the spacer sequence between A-tracts can influence the bending magnitude (15). Some years ago, Hagerman (16) found that a sequence of the form $(dA_4T_4CG)_n$ has a significant electrophoretic abnormality indicating strong bending whereas a sequence with the same nucleotide content $(dT_4A_4CG)_n$ shows normal electrophoretic behavior. A caution is yet emphasized (17) regarding detailed measures by electrophoretic mobility since the contribution of out of plane bending to the extent of migration can bias the estimation.

Several models have been developed to explain the molecular origin of intrinsic A-tract-induced DNA curvature. Early models like the junction model (18) view the sequence-dependent curvature as a result of kinks created at junctions between two types of B-DNA structures [in this model, A-tracts adopt a heteronomous

*To whom correspondence should be addressed. Tel: +494212003541; Fax: +394212003249; Email: m.zacharias@jacobs-university.de

B′-structure that differs from the normal B-DNA, see Ref. (19)]. Supports for this model come in part from premelting transitions of DNA that can be observed in A-tract containing DNAs but not in other sequences (20) and from early modeling studies (21). The wedge model explains the curvature as the effect of dinucleotide-specific roll and tilt angles that if appropriately spaced along the helix can give rise to overall DNA curvature (22,23). Structural analysis of A-tract containing DNA using X-ray crystallography indicated a straight A-tract DNA and located the origin of curvature within the intervening sequences (24). Other high-resolution techniques such as NMR spectroscopy have been applied but the relatively low proton density of nucleotides limit the number of Nuclear Overhauser effects for nucleic acid fine structure determination. These limits of NMR have been partially overcome by the application of residual dipolar couplings and lead recently to the proposal of a more unified model for A-tract induced curvature where bending results from phased combinations of roll and tilt contributions with relatively low amplitude, delocalized over the whole sequence (Delocalized Bend Model) (25). Similar to A-tracts some other sequences can also adopt intrinsically curved structures or may also exhibit a higher bending flexibility of DNA (12,26–29). Furthermore, the bending deformability of mismatches and abasic sites may play a decisive role for distinguishing damaged and undamaged DNA during DNA repair processes. In addition to biophysical and structural studies, molecular modeling and simulation methods have been used to explain sequence-dependent curvature of DNA. These include unrestrained molecular dynamics (MD) simulations of A-tract containing DNA (30–35). Studies based on unrestrained MD simulations suffer, however, from two drawbacks: First, currently accessible timescales may not be sufficient to sample all relevant states compatible with bent DNA structures even around the equilibrium state. Secondly, many interesting DNA-bending deformations seen, for example, in protein–DNA complexes go beyond equilibrium fluctuations of isolated DNA. Hence, it is unlikely to observe and characterize such deformations using conventional MD simulations. One possibility to study DNA-bending deformations systematically is to employ energy minimization including constraints to bend DNA (36,37). For example Sanghani *et al.* (37) used the JUMNA [Junction Minimization of Nucleic Acids, (38)] program to compare induced DNA bending on the Hagerman sequences $(dA_4T_4CG)_n$ and $(dT_4A_4CG)_n$. The authors defined a curved superhelical axis that constrained the DNA to follow a superhelical pathway of defined radius of curvature. Constrained energy minimization using different radii of curvature for the superhelical path was used to calculate the bending properties of the Hagerman sequences (37). A disadvantage of such approach is that the constraint applied to the complete DNA segment may restrict the possibilities of the structure to relax the bending induced strain.

In order to systematically study the microscopic effects giving rise to DNA curvature and the associated energetic costs, we have developed a new restraint to induce continuous bending deformations of DNA. The method is based on a screw axis of double helical segments and a restraint for the angle between them. We implemented this restraint in the JUMNA program for bending a DNA fragment during energy minimization. The bending angle variable has been defined in such a way that all DNA conformational parameters can distribute optimally according to the specific dependence on the sequence without disturbing the helix locally. An additional advantage of the method is the possibility to restrain also the bending direction. This allows a systematic mapping of the global bendability of a given DNA sequence as a function of every bending direction and representing it on a polar plot.

The approach has been applied to several DNA molecules including examples of A-tract containing sequences and sequences that are supposed to influence protein–DNA recognition indirectly due to sequence-dependent bending flexibility. The results indicate good agreement with experimental data and offer atomic level explanations when investigating the bending deformability of nucleic acids. In addition to regular duplex DNA, the effect of central mismatches or an abasic site on bending deformability was also investigated. The calculations indicate a significantly altered bending deformability at abasic sites and some mismatches compared to regular DNA that could play a role during DNA damage recognition. Finally, a MD version of the approach has been implemented and applied to an A-tract containing DNA oligonucleotide for comparison with the adiabatic mapping calculations. Free energy simulations of induced bending on this system indicate a softer penalty for DNA bending but good qualitative agreement with adiabatic mapping on the mechanism of A-tract DNA bending. The coupling of DNA bending to other global variables such as twisting, stretching and groove width was also investigated.

MATERIALS AND METHODS

Geometric variable associated with DNA bending

The bending angle of a double-stranded oligomer is defined as an angle between two screw axes of the adjacent few base pairs (referred to as handles below). A restraining potential of the form $V(\theta) = k(\theta - \theta_{ref})^2$ is then added to the force field where θ is an effective angle between two vectors corresponding to the two handles.

Screw-axis bending variable. The orientation of one nucleotide is defined by a reference axis system located on the base with unit vectors (e_1, e_2, e_3) where e_1 is the bond N9–C8 for purines or N1–C6 for pyrimidines and thus approximately parallel to the direction of the central dyad axis of the double helix. The e_2 vector is defined by the cross-product between e_1 and the glycosidic bond $C_{1'}$–N (opposite sign for the complementary strand) and thus perpendicular to the base plane, and e_3 is the normal with respect to e_1 and e_2 and points approximately in the direction of the long base-pair axis.

Space-invariant rotation vectors between adjacent base reference axis sytems [Figure 1A, (39–42)] can be obtained

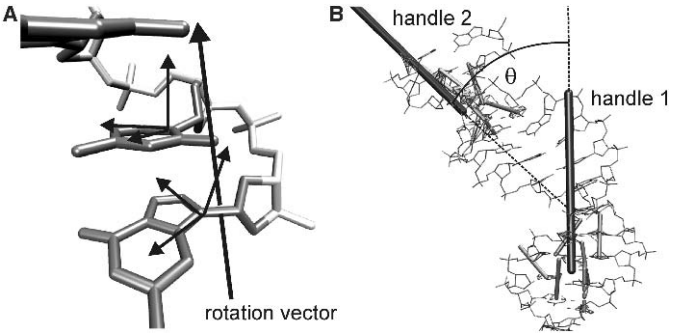

Figure 1. Definition of bending angle restraining coordinates for adiabatic mapping and umbrella sampling calculations. (**A**) Illustration of the local base associated coordinate systems (three orthogonal vectors at each base) of two consecutive nucleotides and the associated rotation vector (each dinucleotide step defines one rotation vector). (**B**) Definition of the bending angle: Each handle vector (long and bold sticks) is the sum of six rotation vectors (short sticks) for 4-bp terminal fragments. The global bend angle is given by the angle between the two handles.

from the so-called 'rotation formula' (43,44). Equation (1) expresses the rotation matrix as a function of the rotation angle θ and the rotation vector $\vec{u}(u_x, u_y, u_z)$:

$$q = I + \sin\theta \cdot S(\vec{u}) + (1 - \cos\theta) \cdot S^2(\vec{u}) \quad \mathbf{1}$$

I is the identity matrix and $S(\vec{u})$ the antisymmetric matrix associated with vector $\vec{u}$. The matrix expression in Equation (1) can be used to define the vector $\vec{u}$ as a function of elements of the rotation matrix q in the following way (42):

$$u_x = \frac{(q_{32} - q_{23})}{\varepsilon}, \quad u_y = \frac{(q_{13} - q_{31})}{\varepsilon} \quad \text{and} \quad u_z = \frac{(q_{21} - q_{12})}{\varepsilon} \quad \mathbf{2}$$

ε is a constant ($= 2{\cdot}\sin\theta$) and can thus be set as $||\vec{u}||$. Since the coordinates of the base-reference axis system are known, the rotation matrix q can be obtained by:

$$q = B' \times B^{-1} = B' \times B^{\mathrm{T}} \quad \mathbf{3}$$

Hence, Equation (2) can now be solved. As illustrated in Figure 1B, space-fixed internucleotide rotation vectors are calculated for several neighboring dinucleotide steps inside a fragment of n base pairs and their sum defines a vector 'handle' [analog to a local axoid in Ref. (41)] oriented as a function of $(2n-2)$ specific inter-nucleotide single rotations. Typically, the four terminal base pairs at each end of a DNA molecule were included to define two (terminal) handles. The global bend angle is given by the angle between the two handle vectors.

Directional bending variable. A specific direction of bending can be imposed by the following construction. The two handles are kept perpendicular to a vector $\vec{r}$, obtained by rotating the long axis of a chosen base pair (we take the vector between the C1′-atoms of the two complementary bases) by the angle α around the average vector of the two handles $\vec{m}$. A direction $\alpha = 0$ corresponds to bending in the direction perpendicular to the reference base pair C1′–C1′ direction (approximately in the directions of the base-pair dyad axis), pointing toward the minor groove. An angle $\alpha = 90^\circ$ corresponds to bending toward the phosphates of the base pair, in the direction of the leading 5′–3′ strand.

During directional bending two quadratic restraining potentials of the form $V(\Omega_j) = k(\Omega_j - 90)^2$ keep the two handles perpendicular to $\vec{r}$. The angle Ω_j between handle Hj and reference vector $\vec{r}$ obtained from the scalar product has no sign. To control the unidirectional nature of bending inside the plane normal to vector $\vec{r}$, one can restrain the 'rotation' of a handle with respect to the other around a vector $\vec{r}$, which means using Equation (1) to restrain the sine of the bending angle θ. The bending angle is thus $\theta_i = \tan^{-1}(\sin\theta_i/\cos\theta_i)$ and for computational simplicity when calculating analytic derivative of θ, independent harmonic biasing potential are used to restrain $\cos\theta_i$ and $\sin\theta_i$.

$$\cos\theta = \vec{H}_1 \cdot \vec{H}_2 \qquad \textit{scalar product}$$

$$\sin\theta = \frac{1}{\vec{r} \cdot \vec{H}_1}\left[\vec{H}_2 - \vec{H}_1 - (1 - \cos\theta) \cdot \left[\vec{r} \cdot \left(\vec{r} \cdot \vec{H}_1\right)\right]\right]$$

Rotation formula

Energy minimization and adiabatic mapping along the bending coordinate

The bending angle restraint was implemented in the program JUMNA (38). JUMNA employs a combination of helicoidal and internal variables to describe and energy-minimize the structure of DNA oligonucleotides. In the current calculations, the Amber Parm98 force field (45) and a Generalized Born (GB) model based on a pairwise descreening approximation was used to implicitly account for solvent effects (46,47). Bending energy curves were generated by a series of simulation 'windows' for stepwise increasing values of the bending restraining angle θ_{ref} (5°-steps; force constant: $k = 400\,\mathrm{kcal\ mol^{-1}\ rad^{-2}}$) in the harmonic restraining energy term. Backward simulations starting from the final structure of the last windows were also performed in each case.

Two-dimensional (2D)-bendability maps that included the bending direction were obtained by a series of bending minimizations for each possible bending direction. The bending direction, α, was changed in 20° steps from 0° to 360°. Structures along the DNA bending pathways were analyzed in terms of local helical parameters using the Curves program (48,49).

MD umbrella sampling along the bending coordinate

The current bend angle-restraining coordinate was also implemented in the MD program Amber version 8.0 (50) and used to search the conformational space associated with the DNA bending pathway. Simulation were performed on an oligonucleotide with a central A-tract (5′-dCGCGCA$_5$CGCGC)$_2$. Standard B-DNA starting structures were generated using the nucgen program of the Amber8 package (50). For consistency with previous

MD simulations on DNA bending and a large-scale effort to characterize the sequence dependence of DNA flexibility ['ABC' initiative, (51, 52)], simulations were carried-out using the Parm94 force field (53). The molecule was neutralized by 16 K^+ counterions and solvated with 3519 TIP3P water molecules (54), corresponding to a solvent layer $\geq$10 Å, within a truncated octahedral box which had a face-to-face dimension of 63 Å. Simulations were performed at constant temperature and pressure applying periodic boundary conditions and the particle-mesh Ewald approach (55) with a 9-Å direct space sum cutoff. Bond lengths involving hydrogen atoms were constrained using Shake (56); the equations of motion were integrated using the Verlet algorithm and a 2 fs time step.

After an initial stage of energy minimization, the solvent and counterions were allowed to equilibrate during 0.6 ns at constant volume while progressively increasing the temperature up to 300 K and restraining the solute to the start structure (standard B-DNA). The solute was then partially relaxed by progressively decreasing the harmonic restraints applied to each atom from 15 to 0 kcal mol^{-1} $Å^{-2}$ over a total period of 1.2 ns at constant pressure and temperature, including a final stage of 0.2 ns without any restraint.

After equilibration of the relaxed oligomer, a series of simulation 'windows' for steadily increasing or decreasing values of θ_{ref} were carried-out for Umbrella Sampling (57–59). For each window i, θ_{ref} was modified by 5° and restrained to the new value using the harmonic biasing potential $V_i(\theta)$ with a force constant of $k = 0.01$ kcal mol^{-1} $degrees^{-2}$ and equilibrated within 0.2 ns. The final structure obtained at the end of this sampling was used as a starting point for the next window generating the starting conformations for the entire bending pathway. Individual window's production runs were then performed simultaneously with a force constant of $k = 0.2$ kcal mol^{-1} $degree^{-2}$ for at least 1.0 ns. Backward simulations starting from the final structure (100° bent) were carried-out using the same protocol.

The Weighted Histogram Analysis Method (WHAM) method (60) was used to calculate a potential of mean force (PMF) or free energy associated with global DNA bending. The WHAM equations express the optimal estimate for the unbiased probability distribution $P(\theta)$ using all data point available over each individual biased probability histogram $P^*_i(\theta)$,

$$P(\theta) = \frac{\left[\sum n_i \cdot P_i^*(\theta)\right]}{\left[\sum n_i \cdot \exp\left\{\left[F_j - V_j(\theta)\right]/k_B T\right\}\right]} \quad \mathbf{4}$$

where n_i is the number of data points used to construct the biased distribution function in window i, $k_B T$ is the temperature expressed in energy units and all sums run over the total number N of windows sampled. The free energy constant F_i is itself determined using the expression of the optimal estimate for the probability distribution function $P(\theta)$,

$$F_i = -k_B T \cdot \int \ln P(\theta) \cdot \exp\left[\frac{-V_i(\theta)}{k_B T}\right] d\theta \quad \mathbf{5}$$

The above two equations have to be solved self-consistently: starting from an initial guess for the N free energy constants F_i to estimate the unbiased distribution function $P(\theta)$ via Equation (4), the latter is subsequently used in Equation (5) to generate new estimates for the N constants F_i, these are used in turn in Equation (4) and so on. The iteration cycle is repeated until both equations are satisfied, i.e. the largest change in the set of values F_i on two consecutive iterations is below a tolerance index (here 10^{-3}). The relative free energy (PMF) at a given bend angle is defined by $W(\theta) = -k_B T \cdot \ln P(\theta)$.

Structures obtained along the DNA-bending pathways were analyzed in term of helical parameters using the Curves program (48,49).

DNA oligonucleotides

Energy minimization studies were performed on several AT-rich doubled-stranded (ds)DNA dodecamers. This included a dodecamer with a central A-tract (5′-dGG $CA_6CGG)_2$ (61), an alternating pyrimidine–purine sequence (5′-$d(AT)_6)_2$ and a pair of sequences, d(5′-CG $A_4T_4CG)_2$ and (5′-$dCGT_4A_4CG)_2$ studied experimentally by Hagerman (16) and referred to as Hagerman H1/H2 sequences (17), respectively. In addition, the recognition sequences of the human and bovine papillomaviruses E2 proteins, (5′-$dACCGAATTCGGT)_2$: HPV sequence) and (5′-$dACCGACGTCGGT)_2$: BPV sequence), respectively, were also studied [referred to as HPV/BPV sequences, (62–65)]. In addition to regular dsDNA, the approach was applied to modified or damaged DNA oligonucleotides with sequences 5′-dCGTAC-CATGC/5′-dGCATGAGT ACG [central abasic site, (66)], 5′-dGCTTCAGTCGT/ 5′-dACGACGGAAGC [central G:A mismatch, (67)], 5′-dGCCACCAGCTC/5′-dGAGCTCGTGGC [central C:C mismatch, (68)] and 5'-dCCATGCGTGG/5′-dCC ATGCGTGG [tandem G:T mismatches, (69)]. For all these sequences experimental structures are available. However, to allow for an unbiased comparison of experiment and calculation, all calculations (on all DNA molecules) were started from canonical B-DNA structures.

RESULTS AND DISCUSSION

Induced DNA bending during restrained energy minimization

The current definition of a DNA bending coordinate based on two screw axis (handles) associated with the two ends of a DNA oligonucleotide (Figure 1B) allows continuous bending of DNA and at the same time full conformational flexibility of the molecule to relax toward a stable conformational state. The induced bending angle based on the angular orientation of the two handles correlates with the global axis curvature as defined by the program 'Curves' (values with the latter are yet systematically lower; Figure 2). Note, that in Curves the global bending is given by the angle between the first and last segments of the curved helical axis. Even better correlation can be seen for the angle between two linear axis obtained from a Curves analysis of the two terminal 4-bp fragments

(the same nucleotides that define the two handles; Figure 2, circles).

Application to DNA oligonucleotides

The application of the restrained bending minimization method (energy minimization in 5° steps) for the Hagerman sequences (H1: $(5'\text{-}dA_4T_4CG)_n$ and H2: $(5'\text{-}dT_4A_4CG)_n$ with $n = 1, 2$) resulted in an energy minimum near 0° for the H2 sequences whereas energy minima at bending angles of ~15° ($n = 1$) and 25° ($n = 2$), respectively, were found in case of the H1 sequence (Figure 3A, onset). Only a small stabilization energy ~–0.3 kcal/mol per A_4T_4-tract compared to a straight structure was obtained. Such a small stabilizing energy is not surprising since experimentally already small changes in the temperature can significantly affect the curvature of DNA ['melting' of bent DNA structure (20)]. Based on the energetic stabilization it is possible to calculate a probability distribution that takes into account the intrinsic geometric probabilities for bend angles (Figure 3). The probability distribution includes a Jacobian factor [sin (bend angle)] so that the probability distribution reflects the underlying free energy of bending which includes the bending energy but also the bending entropy (higher intrinsic probability of bending angles near 90° versus bending angles around 0° or 180°). This results in distinct probability differences between H1 ($n = 1$), H1 ($n = 2$) and H2 sequences and also a slight shift of the free energy minimum (maximum of the probability density) compared to the potential energy minimum. Probability maxima of ~17° for H2 ($n = 1$), 21° for H1 ($n = 1$) and 31° for the H1 sequence ($n = 2$) cases, respectively, with longer tails toward larger bend angles in the probability curves in case of the H1 sequences were obtained (Figure 3). The calculated bend angle maximum for the H1 sequence is close to experimental A-tract bending angle estimates of 17–21° per A-tract [based on gel electrophoresis, (13)]. The results differ quantitatively from the study of Sanghani *et al.* (37) who employed a superhelical bending variable, a different force field (Flex force field) and a sigmoidal distance-dependent dielectric function representing solvent effects. Using the same sigmoidal distance-dependent dielectric function with the parm98 force field we found an energy minimum for H1 at a bending angle of 25° compared to ~10° for H2 (Figure 4). A distance-dependent dielectric model appears

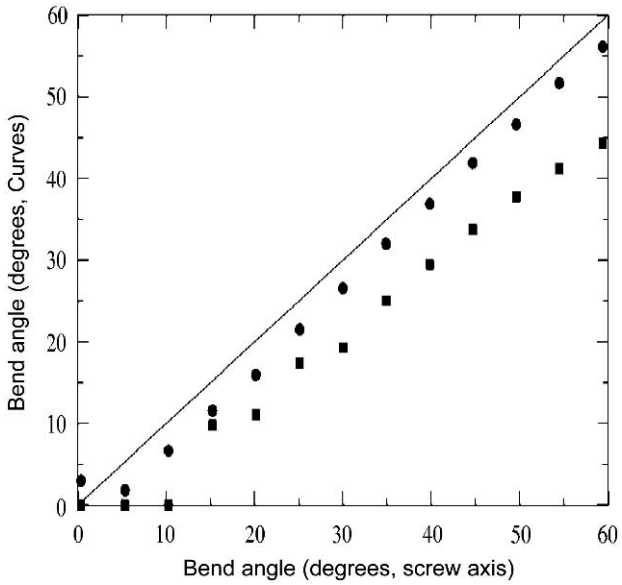

Figure 2. Correlation of DNA bend angles (15-bp polyCG) based on the screw-axis description versus global bending calculated using the program Curves5.0 [squares, Ref. (49)]. In this case, the global bending angle is defined by the angle between the helical axis direction at the first base pair versus axis direction at the last base pair. The circles correspond to a Curves-based bend angle between the linear helical axes of the terminal 4-bp fragments of the DNA oligonucleotide. The line indicates perfect correlation.

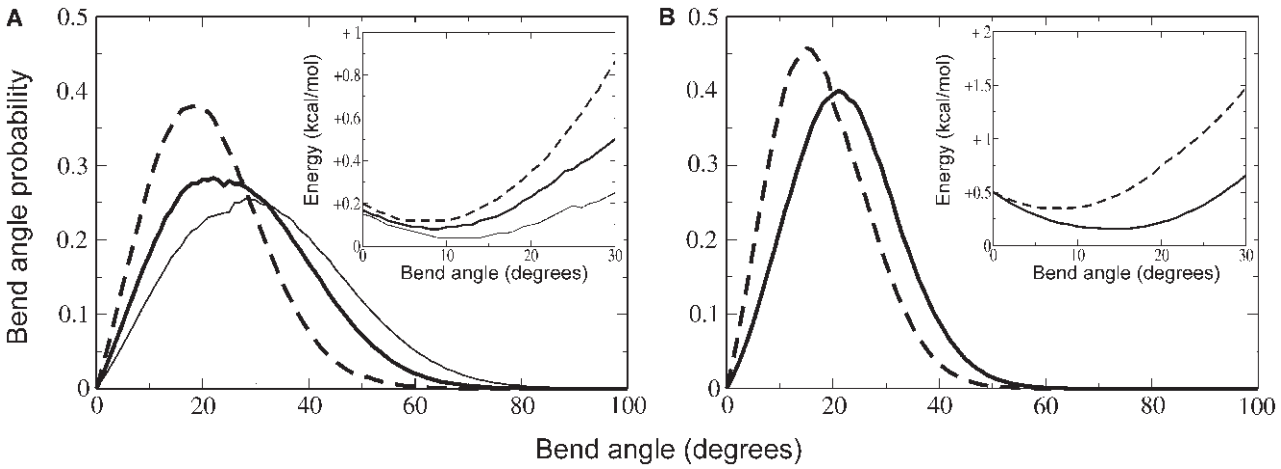

Figure 3. Bending angle probability distributions obtained from the potential energy curve as defined by, $p(\theta) = \sin\theta \exp[-E(\theta)/RT]$, versus bending angles. **(A)** Bend angle probability for the Hagerman 1 sequence [bold line, $(5'\text{-}dCGA_4T_4CG)_2$], two H1 sequences in phase with helical repeat [thin line, $(5'\text{-}dCGA_4T_4CGA_4T_4CG)_2$] and Hagerman 2 (H2) sequence [dashed line, $(5'\text{-}dCGT_4A_4CG)_2$]. (**B**) Same for the HPV-E2-recognition sequence (bold line, $(5'\text{-}dACCGAATTCGGT)_2$) and BPV-E2 recognition sequence [dashed line, $(5'\text{-}dACCGACGTCGGT)_2$]. The corresponding potential energy versus bend angle plots calculated using restraint energy minimization within Jumna (38) are shown as panel insets (same line types as for probability distributions).

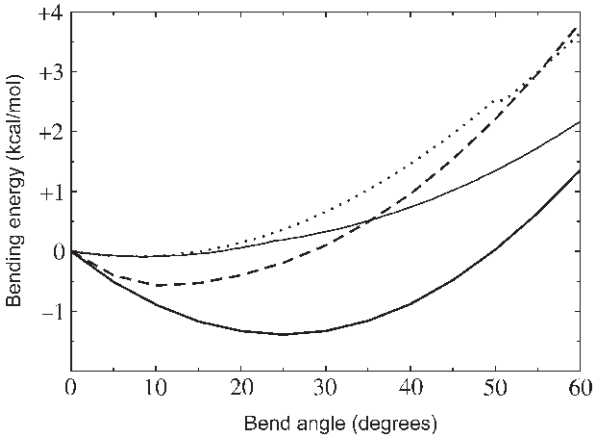

Figure 4. Potential energy versus global bending: comparison of the Generalized Born (GB) model and a sigmoidal distance-dependent dielectric (sddd) function [see text and Ref. (37)]. Calculations were performed on the H1-(5′-dCCGA$_4$T$_4$CG)$_2$-sequence (thin line: GB-model, bold line: sddd-model) and the H2-(5′-dCCGT$_4$A$_4$CG)$_2$-sequence (dotted line: GB-model, dashed line: sddd-model).

to stabilize the curved H1 structure more strongly (>1 kcal mol^{-1}compared to a straight structure) than using the present GB model.

The papilloma virus E2 protein binds to two sequence elements on DNA that are separated by a 4-bp linker not in contact with the protein (70). Efficient interaction requires the DNA to bend toward the protein in the direction of the central minor groove (62,70). It has been shown that the E2 recognition element in human DNA (HPV) adopts already a (pre)bent conformation (toward the minor groove) in the absence of the E2 protein whereas the corresponding bovine DNA recognition sequence is straight in solution influencing the binding affinities of HPV- and BPV-DNA for the E2 protein (62,71,72).

Indeed, restrained energy minimizations along the bending coordinate resulted in an energy minimum with a bending angle of ~18° for the HPV sequence (probability maximum at ~21°) and a steeper rise of the bending energy with a minimum near 0° for the BPV sequence (probability maximum at ~15°, Figure 3B). The curvature estimated with Curves was ~15° for HPV which is in good agreement with previous modeling studies on human and bovine E2 protein papillomavirus-binding sites (70–72). The current bending angle calculated for the X-ray structures (62,63) is, however, smaller for both HPV and BPV (7° and 4°, respectively). A possible reason for the disagreement might be crystal-packing effects that enforce a more straight conformation of the DNA in the crystal compared to free solution [as has been observed for A-tract DNA, (24)].

The restrained bending minimization studies described above were all performed starting from ideal B-DNA with zero initial bend angle and without any restriction on the initial bending direction. It is possible that the initial bending direction has a significant influence on the subsequent bending steps and the resulting structures and energies. In order to control the influence of the initial bending direction on the restrained bending minimization results, the start structures (H1-sequence) were pre-bent in various possible directions (see Materials and Methods section and paragraph on directional DNA bending). The structures were then subjected to restrained bending minimization without any restraints on the bending direction. Interestingly, for small initial bending angles (5° bending magnitude and all possible bending directions in 10° steps) the same bending energy curve and identical minimized structures were obtained as for the bending minimization starting from ideal B-DNA (no pre-bending, Figure 5). At a pre-bending of 50°, slightly different bending energy curves depending on the direction of the pre-bending were obtained (Figure 5). However, even at a large pre-bending of 50° in various directions, the structures undergo large axial rotations such that the final bending direction (upon restraining only the bending magnitude not the direction) was very similar (directed toward the central minor groove, Figure 5A and B). This result indicates that there is a significant directional preference for DNA bending (analyzed in more detail in the paragraph on directional DNA bending) and the energy surface for DNA bending is relatively smooth such that relatively large axial rotations are possible upon removal of a restraint on the bending direction. The differences in the final structures (and energies) upon removal of the directional restraint are mainly due to differences in the conformational DNA backbone substates in the structures obtained after the initial directional bending. It is of interest to note, that if one restrains the dihedral backbone of the DNA during initial directional bending (50° case) close to standard B-DNA and then removes the directional restraint (only keep the restraint on the bending magnitude) all start structures relax to the same structure (bent toward the central minor groove).

Helical parameters of bend DNA structures

We analyzed the molecular origin of bending in terms of local helical parameters of the optimally bent DNA structures calculated with the program Curves (Figure 6). Several studies have already been reported to highlight the structural origin of curvature for DNA fragments containing runs of adenines (A-tract, see Introduction section). A major finding of X-ray crystallography is that A-tracts and the H1 sequence adopt an exceptionally narrow minor groove and the A:T base pairs are highly propeller twisted (73). In agreement with experiment (17,73,74) the bending energy minimum observed for the H1 sequence showed significant (negative) propeller twisting of the complete central segment (Figure 6). The calculated bending direction was, indeed, toward the central minor groove. An important point concerning the H1 sequence is that the central AT step does not disrupt the structural uniformity of A-tract stacking [negative propeller twist pattern, (73,74)]. In addition, for the H1 sequence the positive roll angles at the CG steps flanking the central A-tract add up in phase with a small negative roll at the central AT (half a turn away) resulting in the observed global

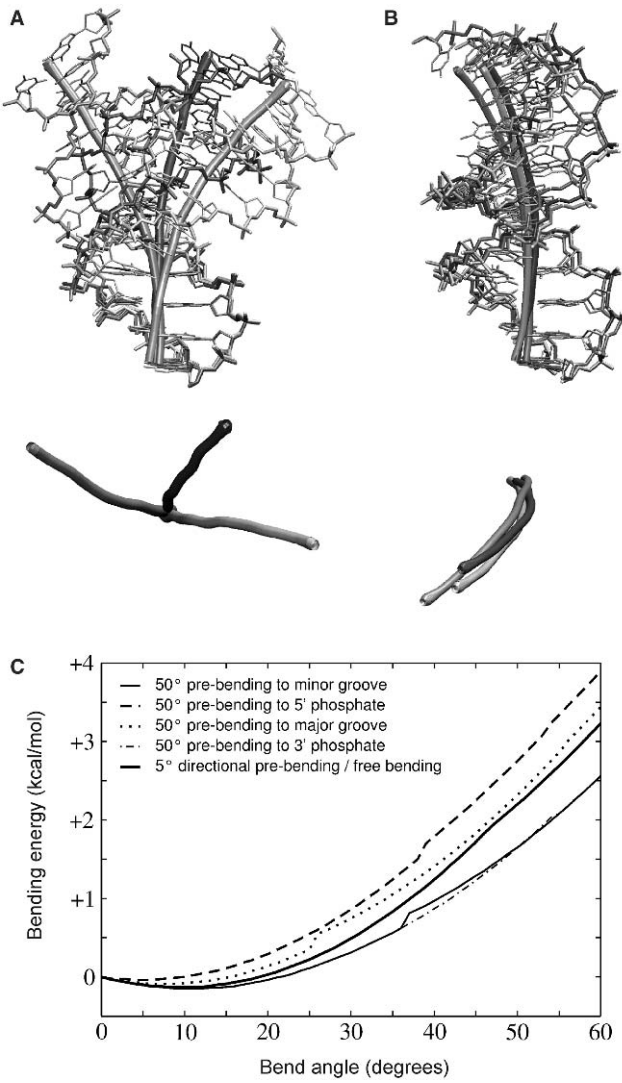

Figure 5. Influence of initial DNA-bending direction (H1-sequence) on results of restrained bending minimization including only a restraint on bending magnitude. (**A**) Superposition of three DNA start structures (stick models at different gray levels) bent in different directions by 50°. Bending directions were toward the central major groove (darkest gray) and toward the phosphate groups closest to the 5′- or 3′-ends of one DNA end, respectively (bending direction approximately perpendicular to bending toward central minor or major grooves). The helical axis of each DNA calculated using the program 'Curves' is indicated (bold lines). The view in the upper panel is orthogonal to the helical axis whereas in the lower panel a perpendicular view approximately along the helical axis is given (only the helical axis for each DNA is shown). (**B**) Same for the energy minimized structures [starting from the structures shown in A] after removal of the restraint on bending direction but keeping a restraint on bending magnitude. (**C**) Calculated potential energy versus bending magnitude starting from various DNA structures pre-bend in different directions and by different magnitudes (according to line type indicated in the inset of the Figure). Free bending corresponds to the bending energy curve obtained by starting from ideal B-DNA (same curves as obtained by starting from any of the 5° pre-bent structures; the 5° pre-bent structures were generated in 10° steps of the bending direction).

bending (Figure 6). This bending mechanism is in good qualitative agreement with experiment [NMR structure of the H1-sequence, (17)] that also shows continuous negative propeller twisting of the central segment. However, the magnitude of the negative roll at the central AT step is larger than for the calculated structure.

For H2, both the calculations and the analysis of the experimental NMR structure (17) indicate on average more positive propeller twist at the central segment (smaller fluctuation in the calculated structure). The calculated roll-angle pattern agrees almost quantitatively with experiment. The positive roll angles at flanking CG steps are outbalanced by a large central positive roll. Consequently, significantly larger local roll changes are necessary in order to add up to achieve global bending compared to H1 (energetically unfavorable). In addition, the central negative propeller twist pattern (necessary to narrow the minor groove) is interrupted at the central TA step of H2 by a reduced magnitude of negative propeller twisting (almost zero for the experimental structure). In Figure 7, we illustrate the dependence between negative propeller twist of H1/H2 sequences and the minor groove narrowing. In the case of A tract and H1-sequences the negative propeller represents a sterically favoring mechanism, whereas for H2 the reduced magnitude of propeller twisting can be explained by the steric hindrance of the two central adenine amino groups upon negative propeller twisting, in agreement with classical Calladine rules (82).

When looking at the HPV/BPV pair, where the unique difference between the two sequences is a substitution of a central AT by a central CG, one can draw up basically the same interpretation of sequence effects as detailed above for H1/H2. Indeed, positive rolls at CG steps add up with a small negative roll at the central AT in HPV, but are opposed by a large positive roll at the central CG in BPV. Together with the smaller variation in tilt, this results in a net curvature in the former sequence and confirms the structural uniformity of AT step in different sequence contexts. The results on the calculated structure are in qualitative agreement with the experimental free DNA structure (62,63) and in even better agreement with the DNA bound to the E2-protein (64,65). This concerns the pattern of roll and tilt angles as well as the on-average negative propeller twist angles of the central segment (Figure 6). For both the HPV and BPV sequences the propeller twisting in the experimental structure shows larger fluctuations compared to the calculated structures.

Directional bending of B-DNA

In addition to free relaxation upon induced DNA bending it is also possible to add an additional restraint for restricting the bending in a preselected direction. A 2D directional bendability map provides an instantaneous picture of the propensity of a DNA to bend in any possible direction. Bending directions were systematically checked using both Curves (version 5) (49) and Madbend (30) (data not shown). Instead of the potential energy, we use the probability density for bending with respect to the bending amplitude θ for a given direction α, expressed as $p_\theta = \exp[-E_\theta/k_B T]/\sum \exp[-E_\theta/k_B T]$, where the sum runs over all value of θ corresponding to one particular direction of bending. This variable enhances contrasts and highlights more selectively global bending properties compared to a plot of the bending energy. In addition, the

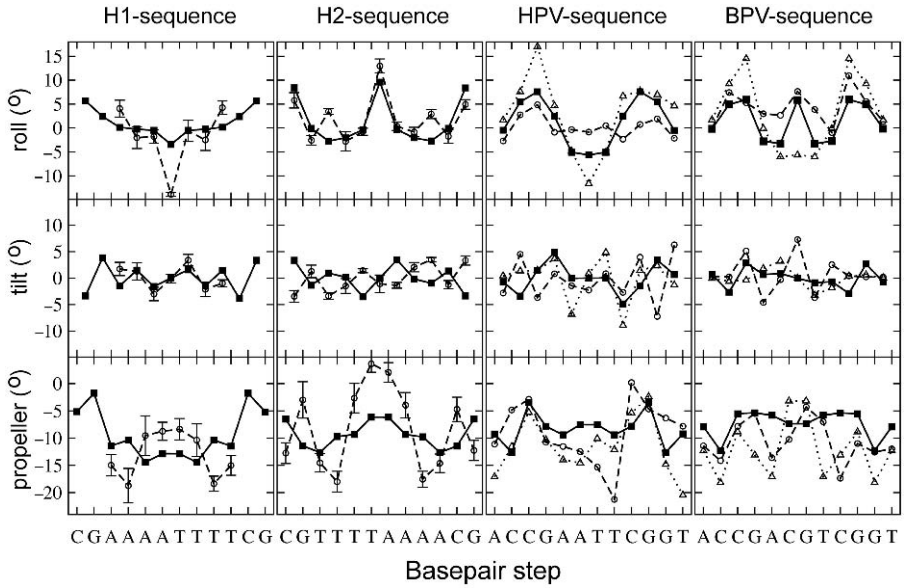

Figure 6. Helical parameters roll, tilt and propeller twist (filled squares in each plot) versus base sequence for the most stable bent conformations (probability maxima in Figure 3). All calculations were performed on palindromic (single repeat) sequences. Experimental data for NMR-derived structures of the H1 [5′-dCCGA$_4$T$_4$CG)$_2$] and H2-[5′-dCCGT$_4$A$_4$CG)$_2$] sequences and X-ray structures of the unbound HPV-E2 and BPV-E2 recognition sequences are indicated as circles (dashed lines). Peripheral base pairs are not reported for the NMR structure of H1, which contains GC instead of CG flanking steps. Error bars have been derived from the analysis of all published NMR structures in each pdb-entry. For the HPV-E2 and BPV-E2 recognition sequences experimental data on the DNA molecules in complex with E2-proteins are plotted as triangles (dotted line).

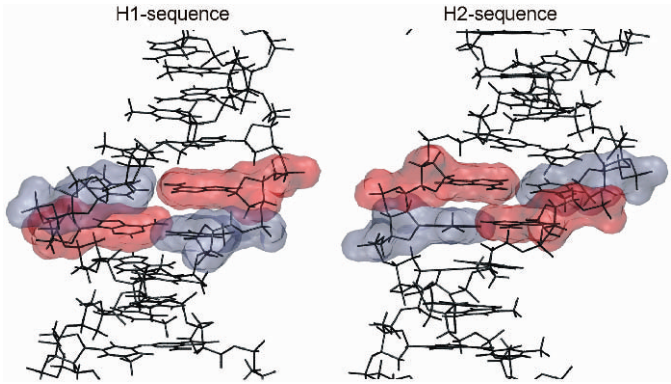

Figure 7. View into the minor grooves of the central AT and TA base-pair steps of the Hagerman H1 and H2 sequences, respectively (bond stick and opaque molecular surface representation). The central adenine/thymine nucleotides are indicated in red/blue, respectively. In case of the H2 sequence, a central bending toward the minor groove interferes with a potential sterical clash of the central adenine bases (partial cross stacking; not the case for the H1-sequence).

probability maximum (most likely conformation) with respect to the total 2D-probability distribution is also reported (squares in Figures 8 and 9).

For the oligonucleotides with a central A$_6$-tract and for the H1 sequence the directional 2D-bendability map indicate bending probability contours that extend to large values for bending toward the minor groove (Figure 8). The result agrees with unrestrained MD simulation studies and NMR spectroscopy of the A-tract containing DNA fragments that also indicated curvature toward the DNA minor groove (31–34). Other bending directions are significantly less favorable. Note, that only a directional bending map as calculated in the present study can give an ultimate indication of a preferred DNA-bending preference since both bending magnitude and direction are treated as variables that can be changed separately. Almost quantitative agreement between experiment and most probable DNA conformation in terms of bending magnitude and direction have been obtained (square and cross in the plot for the A-tract and other DNA structures, Figure 8). The H1 sequence exhibits a slightly different bending direction and slightly reduced bendability compared to the A-tract with six consecutive adenines. Experimental studies have indeed reported maximum electrophoretic anomalies for A-tracts of length 6 (15). In case of the H2 sequence and the d[5′-(AT)$_6$]$_2$ oligonucleotide, a smaller overall bending tendency toward the minor groove was obtained. The $P = 0.1$ contour was in both cases $<25°$ whereas in case of the A6-tract and H1 sequence the $P = 0.1$ contour reached bend angles of $>35°$ (Figure 8).

Noticeably, for bending toward the central major groove (direction angle $\sim 180°$) the $P = 0.1$ contour for the alternating AT sequence was slightly larger than for the H1 or A-tract case. The energetic cost was similar ($\sim$5 kcal/mol) for a 60° bending in the direction of the central minor or major groove for the alternating AT sequence but roughly doubled ($\sim$10 kcal/mol) for the latter direction in case of A-tract, H1 and H2 sequences. Hence, the directional bendability map predicts a much easier bending of alternating TA sequences toward the major groove. Indeed, alternating AT(or TA) sequences are often found as a part of recognition elements for DNA minor-groove-binding proteins. Binding to the DNA minor groove requires minor groove opening and bending toward the major groove. Hence, the easier bending

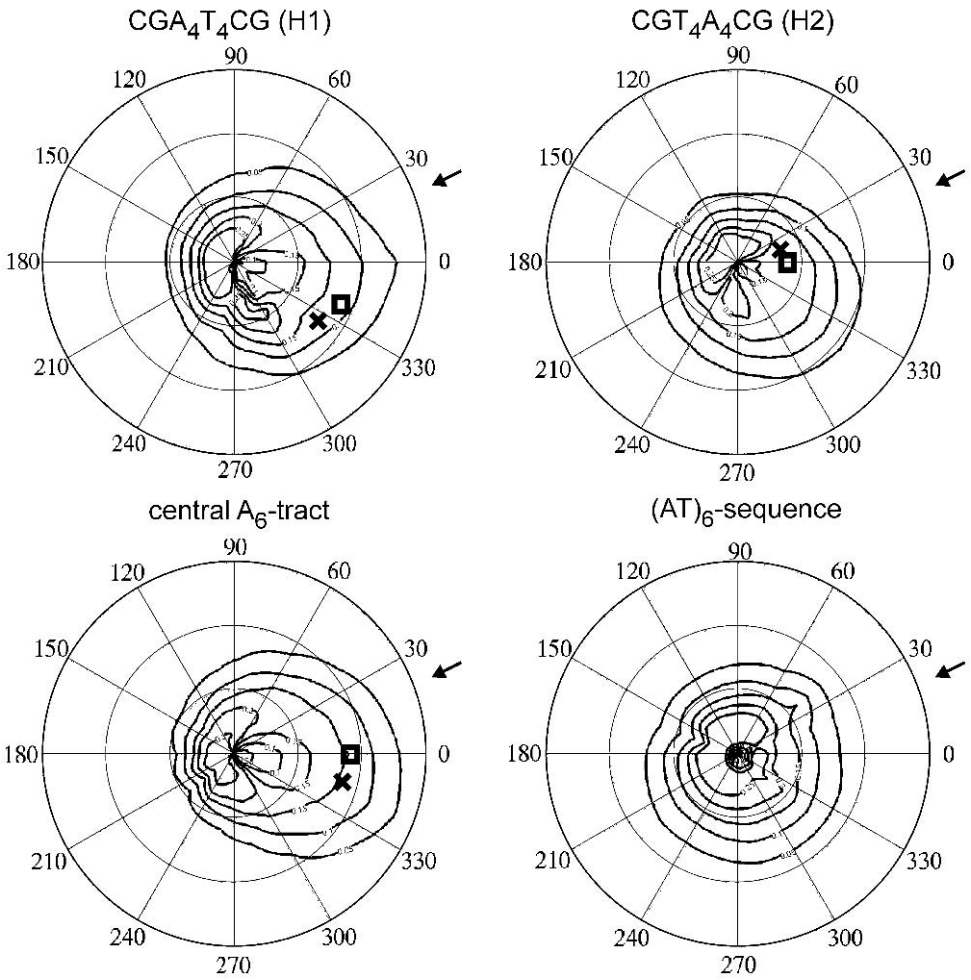

Figure 8. Iso-probability surfaces for directional DNA bending. The H1 and H2 sequences are given in the legend of Figure 3. Bend angles up to 45° have been considered (largest circle; inner circles correspond to 15° and 30° bending angle magnitude, respectively). Meridians correspond to bending directions (0–360°; tics on the outer circle). The direction of bending toward the central minor groove has been marked by an arrow. The contours (in steps of 0.05 with a maximum of 0.25) are constant Boltzmann probabilities normalized independently in each bending direction as discussed in Materials and Methods, and Results sections. The squares in each plot correspond to the probability maximum of the complete 2D-probability distribution (most stable bent conformation). The cross represents the directional bending state of the experimental conformation for each case [except for the $(AT)_6$-sequence for which no experimental data are available].

deformability of short alternating AT sequences toward the major groove may contribute indirectly to the recognition by the minor groove-binding proteins as also suggested by other studies (30).

Directional bending of damaged and mismatched DNA

Common examples of lesions in DNA include mismatches and the formation of single abasic sites. Changes in the global structure and bending deformability may play a role for the recognition of damaged DNA by repair enzymes (75,76). The presence of abasic sites in DNA can also affect structural and dynamic properties of the adjacent duplex DNA. MD (77,78) and EM (79) studies indicate significant bending associated with a decrease in interstrand interaction energy at the site of the lesion. However, a systematic analysis of the bendability of abasic sites and mismatches in all possible bending directions has so far not been performed. In order to do such an analysis, we chose DNA oligomers for which experimental structures are available with either a central abasic site, a single C:C mismatch, a single G:A mismatch or a tandem G:T mismatch. In the case of the abasic site, the experimental structure indicates a bend conformation with similar bending magnitude as the most probable conformation of our 2D bending map but with opposite bending direction (square and cross in Figure 9). However, the experimental conformation lies within a region of high-calculated bending probability. This result indicates that the abasic site might be compatible with multiple possible conformations that are overall bent in different directions (79).

The directional bendability map for the tandem G:T mismatch indicates little difference to regular B-DNA and a slight preference for bending toward the central major groove. The predicted most probable bending magnitude and direction agree very well with the experimental results (Figure 9). On the contrary, the 2D bendability maps for both a single central G:A and a single C:C mismatch differ significantly from regular B-DNA, showing much more

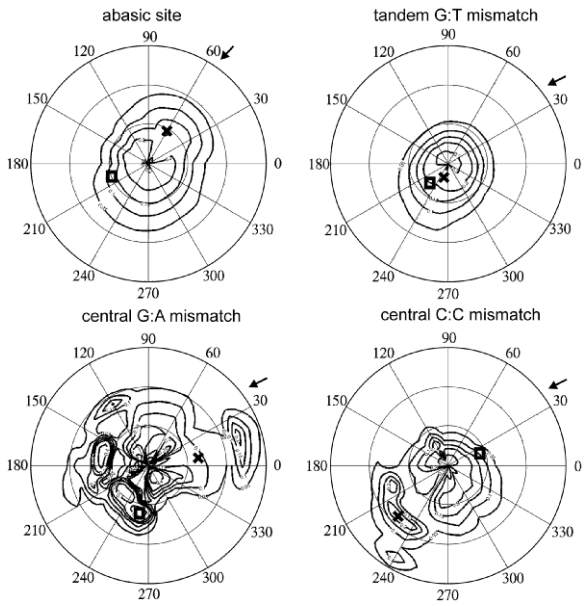

Figure 9. Iso-probability surfaces for directional bending of abasic DNA [5′-dCGTAC-CATGC/dGCATGAGTACG, (66)], a tandem G:T mismatch [(5′-dCCATGCGTGG)$_2$, (69)], a central G:A mismatch [5′-dGCTTCAGTCGT/5′-dACGACGGAAGC, (67)] and a C:C mismatch [5′-dGCCACCAGCTC/5′-dGAGCTCGTGGC, (68)]. Same lines and symbols as in Figure 8.

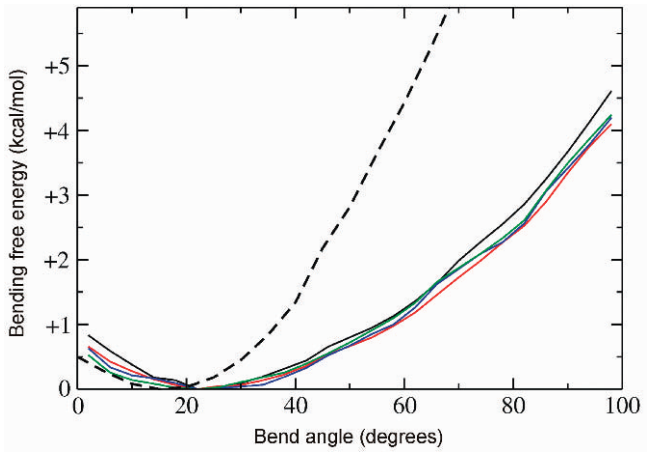

Figure 10. Potential of mean force for bending of an A-tract DNA oligomer with the sequence (5′-dCGCGCA$_5$CGCGC)$_2$ using molecular dynamics umbrella sampling simulations and a bend angle restraining step size of 5° (see Materials and Methods section for details). Calculated free energy curves are plotted for forward simulations after 0.5 ns (black) and 1 ns (red) of data collecting time for each umbrella sampling window. Backward simulations were started from the fully bent structures and free energy curves were calculated independently from the forward simulations (blue and green curves for 0.5 ns and 1 ns sampling time per window, respectively). The variation of each independently calculated free energy at each bend angle was $<\pm 0.5$ kcal/mol. The dashed line corresponds to the potential bending energy as obtained from restrained energy minimization (adiabatic mapping) of the same DNA molecule.

complex behavior. In both cases several probability submaxima are observed indicating several possible stable substates with different global bend angles and bending directions. The calculated most probable bending magnitude and direction differ from experiment, which is not surprising, since the balance between different minima may easily be changed. Hence, our results show that a greater variety of alternative stable bending directions can be accommodated by these two sequences if external factors (here a virtual bending coordinate) act on these structures. This property may play a decisive role for the recognition of these elements by repair enzymes.

Induced A-tract bending during MD umbrella sampling

The adiabatic mapping of DNA bending employing an implicit continuum solvent model allows the systematic investigation of DNA deformability as a function of bending direction and magnitude. In order to investigate the effect of a more realistic representation of surrounding aqueous solvent and ions we performed dynamics simulations of DNA bending. In order to calculate free energies of bending we used the umbrella sampling technique. This approach allows in principle also the sampling of several conformational DNA substates compatible with a given global bend angle (not only energy minima as during the adiabatic mapping calculations). The MD/umbrella sampling approach is, however, computationally much more demanding than the adiabatic mapping EM method. Therefore, the MD simulations were limited to one A-tract-containing DNA and results were directly compared to adiabatic mapping using the same molecule (5′-dCGCGCA$_5$CGCGC)$_2$. To control the convergence of the simulations, forward simulations starting from B-DNA and backward umbrella sampling simulations starting from the final structure of the forward simulation (but otherwise independent) were performed. In addition, the free energy curves were calculated for two simulation times of 0.5 and 1 ns per umbrella sampling window (spacing of windows: 5°), respectively. The force constant to restrain the bending angle to a window reference angle was small enough ($k = 0.2$ kcal mol^{-1} degrees^{-2}) to allow for sufficient overlap between neighboring angle windows. The resulting free energy curves were very similar with a variation of <0.5 kcal/mol at each data point and a free energy minimum close to $\theta = 20°$ (Figure 10). The average bend angle calculated with Curves was also ~20° (average of conformations from the corresponding bend angle restraining window). The optimal bend angle from the MD umbrella sampling calculations shows a good agreement with the global minimum of bending estimated for this sequence when using the EM approach (dashed line in Figure 10) and is also in good agreement with experimental observations (13).

Bending up to $\theta = 100°$ required a total free energy change of 4 kcal/mol which is significantly smaller than the calculated bending energy change from the adiabatic mapping with a GB solvent model. This result indicates that either the explicit solvent representation significantly softens the bending deformability of DNA or the MD simulations (in Cartesian coordinates instead of internal coordinates in JUMNA) provide many more degrees of

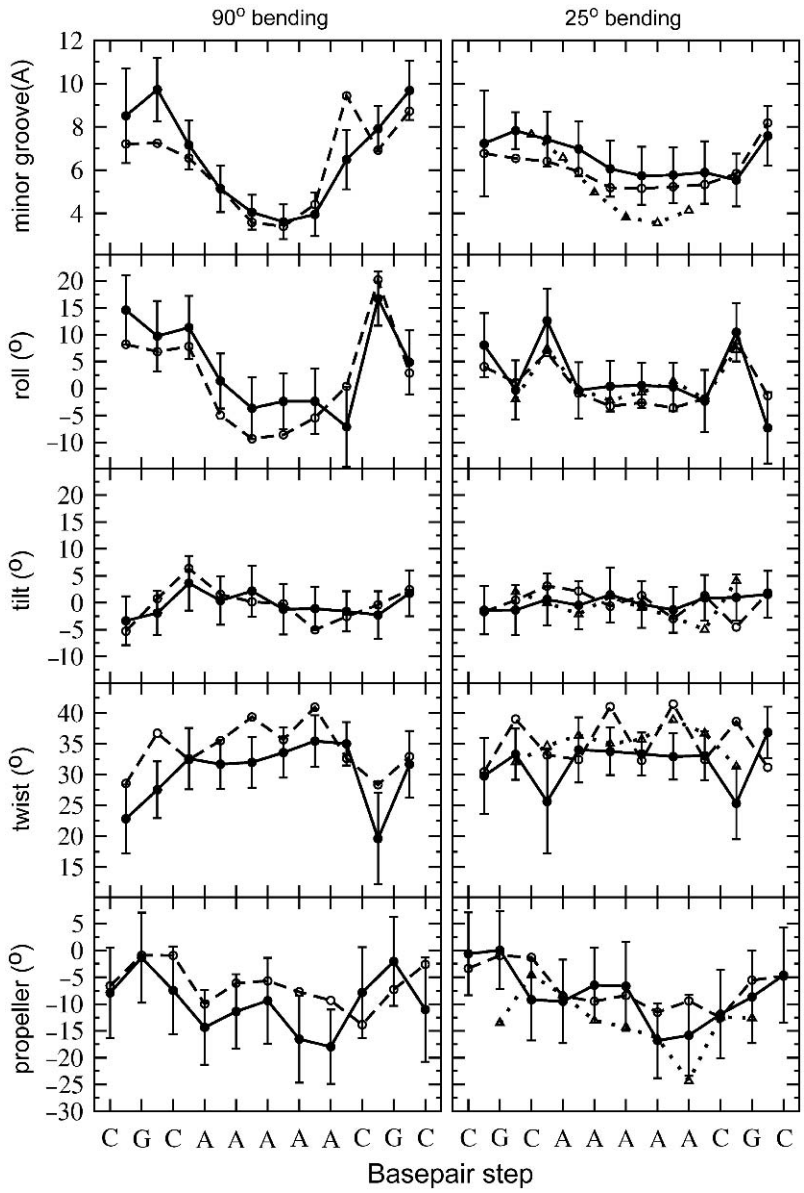

Figure 11. Average helical parameters in the regime of high and small curvature for the central sequence (5′-dCGCA$_5$CGC)$_2$ during MD free energy simulations (filled circles and full line) and during adiabatic mapping using energy minimization (open circles and dashed line). Averages and standard deviations (indicated as error bars) were taken from DNA conformations sampled in the umbrella sampling windows in the range of 20° to 40° for small (near optimal) curvature and in the range of 80° to 100° for tight bending. Parameters obtained from the adiabatic mapping approach were calculated for energy-minimized structures corresponding to 25° and 80° bending, respectively. For comparison, helical parameters extracted from the experimental structure of a very similar A-tract DNA [(5′-dCGCA$_6$GCG)$_2$, (61)] that adopts a bent conformation in solution (~20°) was also included (triangles and dotted curve). To allow direct comparison the central helical step for the experimental structure is not shown.

freedom to relax the bending-induced strain on the DNA conformation. It is possible to fit a quadratic function to the bending free energy curve near the free energy minimum. From the associated curvature it is possible to calculate an estimate of the persistence length P of the DNA molecule [using $P = 2 \times$ contour length (in base-pair steps)/<bend angle variance at room temperature>]. In case of the MD simulations, a persistence length between 170 and 180 bp was obtained in good agreement with experiment [~150 bp, (80)]. The adiabatic mapping predicts a stiffer DNA with a persistence length of 350 bp, more than twice the experimental value for DNA (80). During both, the free energy simulation and the adiabatic mapping, bending was strongly anisotropic, directed toward the minor groove of the central A-tract.

Helical conformational changes induced upon A-tract bending during MD simulations

Helical DNA parameters of conformations near the optimal bending angles (20–30°) and at higher-induced bend angles of ~80–100° were analyzed (Figure 11). For comparison the helical parameters of the structures corresponding to 80° and 20° obtained by adiabatic mapping for the same oligonucleotide were also

calculated. An experimental structure of exactly the same sequence is not available, however, the structure of a close sequence, with a central A6 instead of an A5 tract has been determined by NMR spectroscopy (61). The helical parameters for this structure (except for the central AA step) have also been calculated for comparison with experiment. Both types of calculations (EM adiabatic mapping and MD simulations) suggest qualitatively the same bending mechanism in line with the analysis of the Hagerman sequences (see above). Both in the regime of optimal bending as well as in the regime of strongly induced bending, the results from the EM calculations are close or within the error bars obtained from the analysis of the MD umbrella sampling simulations (Figure 11).

At the optimal bending angle, both the adiabatic EM mapping and the umbrella sampling/MD simulations indicate significant negative propeller twisting of the central base pairs (~–12°) and a significant narrowing of the minor groove relative to the flanking sequences. This is also seen for the experimental structure (Figure 11). The average propeller twist in the central segment indicates that interstrand bifurcated hydrogen bonds linking the adenine residue N6-H amino group to adjacent thymine O4 group across the major groove are at most transient, since they require a propeller close to –20°. This supports the notion that bifurcated hydrogen bonds are not a driving force in the A-tract-induced bending phenomenon (33,81). Slightly negative roll angles within the A-tract and on average slightly decreasing tilt angles along the A-tract were observed. However, large positive rolls were seen at both peripheral CG and CA steps, respectively, associated with a reduction of twist. Almost quantitative agreement between the experimentally observed pattern of roll and tilt angles and the calculations at the optimal bending angle (for both EM and MD) was obtained (Figure 11).

In the regime of high curvature (80–100°), the above picture of local changes in the conformation does not change dramatically (no completely new or different bending mechanism is seen). It evolves continuously so that trends observed at optimal bending can be seen even more clearly. The pattern of roll angles along the sequence shows significantly increased positive values on both sides of the A-tract except for the 3′ purine–pyrimidine junction, and negative values (~–5°) at each AA step (for both adiabatic mapping and free energy MD simulation). The positive roll on both sides of the A-tract is coupled to an unwinding of the helix especially at the CG steps flanking the A-tract (Figure 11). Also, the tilt profile indicates slightly positive values at the beginning of the A-tract (CA step) and slightly more negative values at its 3′ side. The minor groove narrowing and the negative propeller-twist angles along the A-tract are more pronounced than for the small bending regime.

Overall, the trends observed during both adiabatic mapping and during the MD free energy simulations agree qualitatively very well with experimental results on A-tract bending (25,61) and free MD simulation studies (32–34). Solution studies of oligonucleotides containing A_4 tract (25) and A_6 tract (61) by NMR spectroscopy reported a very similar roll pattern along the DNA molecules with significant positive rolls at base-pair steps flanking the A-tract, coupled to an unwinding at those steps and small negative roll angles within the A_6 tract (61). In addition, the trend observed for tilt (positive near the 5′-end of the A-tract and negative at the 3′-end) was also observed in both NMR studies (25,61).

The above analysis of the helical parameters is largely compatible with the junction model and a recent extension, termed Delocalized Bend (DB) Model (25) which combines elements of the wedge (12,13) and junction models (22,23). The junction model assigns the main cause of bending to elements at the junction of A-tracts. In the DB model, significant contributions to the overall bend arise also from roll and tilt of the dinucleotides inside the A-tract. Such contributions can be seen in the present simulations especially in the regime of strong DNA curvature.

Coupling of bending deformations to other global parameters of DNA

In contrast to free MD simulations or other restraining approaches, the present methodology allows the continuous smooth global DNA bending much beyond equilibrium levels as it can for example occur upon protein–DNA binding. It is also possible to investigate systematically the coupled changes of other global parameters upon inducing DNA curvature. Unexpectedly, the average minor groove width of the A_5-tract containing DNA fragment that was studied by EM and umbrella sampling simulations showed only an insignificant reduction with increasing bend angle. However, this was largely due to an opposing effect of the central A-tract (narrowing of the minor groove) and the flanking sequences. From the analysis of complexes of proteins that bind in the minor groove of DNA, it is known that minor groove opening results in unwinding (untwisting) of DNA. This coupling was also seen for the central A-tract segment that showed an increase of the average twist with increasing bend angle (and decreasing minor groove width, Figure 12). In contrast, the average twist and average rise (path length/number of base-pair steps) per base pair of the flanking sequences started to decrease upon bending beyond 40°. Opposing trends seen for the central and the flanking segments appear to reduce the overall coupling of bending to other global conformational parameters of the DNA molecule. The specific trends for A-tract and flanking sequences concerning the stretching, winding and average minor groove widths were found in both EM and MD calculations (Figure 12). Interestingly, the fluctuation of twist, rise and minor groove width of the central A-tract (error bars in Figure 12) do not change with increasing bend angle and are always significantly smaller for the central A-tract compared to the flanking sequences. This result indicates that the central A-tract forms a relatively rigid segment that largely keeps its rigidity even under significant bending stress.

CONCLUSIONS

A new restraint to induce DNA curvature, based on the definition of an average screw axis for the terminal

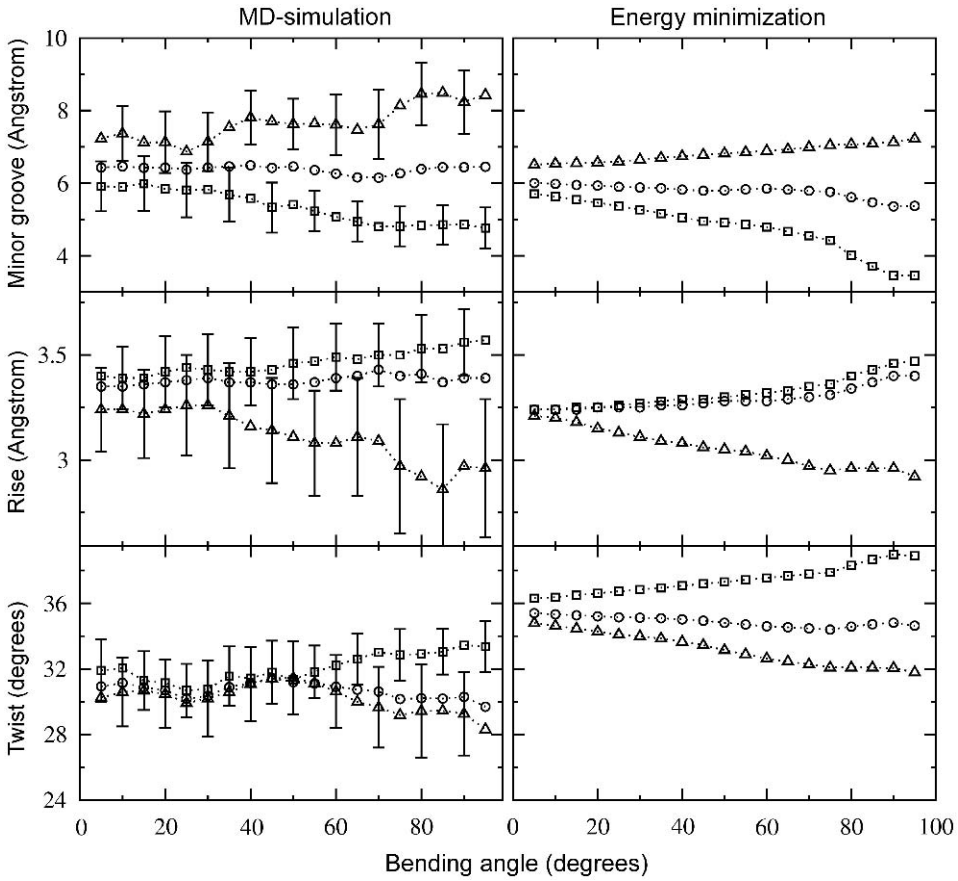

Figure 12. Correlated changes of global conformational variables upon induced DNA bending of the (5′-dCGCA$_5$CGC)$_2$ oligonucleotide. The average minor groove width, the rise per base-pair step and the average DNA twist are plotted versus induced global bend angle (circles) for both the MD umbrella sampling simulations and the adiabatic mapping energy minimization calculations (using 'Curves' global parameters). In addition to the average along the complete DNA, corresponding averages for the central (5-bp, squares) and terminal parts (triangles) of the DNA are also shown.

segments of DNA oligonucleotides has been developed. In principle, DNA bending could also be induced by simple distance or angle restraints involving selected atoms at both ends and in the middle of a DNA fragment. However, application of such restraints can result in significant distortion of the local structure near the atoms involved in the restraining coordinate and, depending on the choice of atoms, create a directional bias to bending. The present average screw-axis definition is based on the local coordinate frames of each nucleotide of the terminal fragments. This definition allows an optimal distribution of the bending restraining force on the whole terminal segments and all DNA conformational parameters can distribute optimally according to the specific base sequence, without disturbing the helix locally. The method allowed us to systematically investigate the bending deformability of various DNA molecules without any restrictions on the conformational relaxation of the DNA upon addition of the bending restraint.

In addition to free relaxation upon induced bending, the algorithm was extended to add a restraint on the bending direction and to rapidly create a 2D DNA-bendability map. Note, that only such an approach allows a systematic exploration of the sequence-dependent bendability of DNA in contrast to 1D approaches or free simulations that may suffer from sampling only one or a limited number of bending directions. Interestingly, pre-bending in various possible directions followed by removal of the restraint on bending direction resulted in axial rotation toward a preferred bending direction indicating a significant bending anisotropy of DNA.

The method is applicable not only to DNA but also to other chain molecules to investigate global bending properties. This includes possible applications to study RNA and the influence of nonhelical elements that frequently interrupt base-paired regions in RNA on global and directional bending properties. For example, the ribosome consists of many RNA helices connected by nonhelical elements. Understanding the global deformability of these elements is critical for an understanding of the ribosome function. Other application examples are the bending and kinking flexibility of long helical segments in proteins that are often of functional importance for the protein molecule.

Results on A-tract containing sequences and the Hagerman H1 sequence predict preferential bending toward the minor groove and resulted in helical parameter changes in good agreement with experiment. In addition,

the direction and magnitude of bending was in very close agreement with experimental solution structures. For the $(AT)_6$ sequence, the calculations predict no preferred bending toward the minor groove but a greater bendability toward the major groove relative to the other DNA molecules under study. This is compatible with the preference of TA containing segments at the center of the recognition sequences for minor groove-binding proteins that need to bend toward the major groove in order to open the minor groove. The energetic preference of bend versus straight conformations was of the same order or smaller than the average energy per degree of freedom at room temperature for H1 and A-tract sequences. This small energetic preference indicates that DNA is in a dynamic equilibrium with many different states and a sequence-specific distribution of possible bend angles.

Significant changes in bending magnitude and preferred bending direction (compared to regular DNA) were found for duplex DNA with central mismatches. In particular for a central G:A or central C:C mismatches, the calculations predict a significantly enlarged variety of possible structures that showed different directions and magnitudes of global bending compared to B-DNA. The altered bending flexibility may contribute to the recognition properties of modified DNA by repair enzymes.

The approach was extended to study DNA bending during MD simulations including surrounding ions and water explicitly. It should be emphasized that this approach goes significantly beyond free MD simulations of DNA where at most bending flexibility close to the equilibrium state can be investigated. Our approach allows for the first time to induce smooth global bending in a DNA fragment during an MD simulation and to directly extract the associated change in free energy. The calculated bending free energy profile for an A_5-tract containing oligonucleotide shows an optimal bending angle close to experiment and the corresponding minimum obtained from the adiabatic mapping calculations. However, the calculated free energy penalty for strong DNA bending was calculated significantly smaller for the explicit solvent free energy simulations than for the adiabatic mapping. The helical parameter changes that occurred upon induced global bending showed good agreement between adiabatic mapping and MD as well as with experimental data form an NMR structure of a similar A-tract DNA. The global bending flexibility estimated in terms of a persistence length was in close agreement with experiment for the free energy simulations whereas the bending rigidity obtained from adiabatic mapping was much larger. The simulations predict that the mechanism of DNA bending is similar at higher bending angles that go beyond the equilibrium-bending angle for the A-tract DNA. It is important to note that even at these higher bend angles no unstacking (kinking) of DNA was observed for the A-tract DNA studied. It is likely that the onset of unstacking events due to strong bending fundamentally alters the pattern of helical parameters associated with bending. A systematic analysis of induced kinking of DNA under 'bending stress' is possible using the present approach and this will be a subject of future studies. Another interesting finding of the present simulations is that although there is some coupling of the global bending with other global variables characterizing the DNA, the fluctuations of these global variables do not change significantly under bending stress. The rigidity of the central A-tract (in terms of groove width, twist and rise fluctuations) was found to be very similar for relaxed and significantly bent DNA.

It should be noted that the current method is mainly limited by the possible incomplete sampling of conformational substates that are compatible with a given bend DNA conformation. In case of the EM adiabatic mapping approach only one conformational state (nearest local energy minimum) is considered for each bending angle (and direction). In case of the MD umbrella sampling approach, several substates compatible with a bend conformation can in principle be sampled but due to barriers between the substates it may not be possible to achieve an exhaustive exploration. The good qualitative agreement between EM and MD results for the A-tract DNA indicates, however, that the EM adiabatic mapping method may allow for a rapid and systematic exploration of many other DNA sequences (and possibly other types of biomolecules such as RNA) to get at least qualitative insight into the bending deformability. Work in progress concerns the directionally restrained bending during free energy MD simulations. The approach allows us to explore systematically the sequence-dependent bending flexibility of DNA without influencing any other structural features of the double helix. The problem is of particular importance in view of the recent studies that DNA can form mini-circles with a much stronger curvature than expected from the persistence length of DNA (27–29).

ACKNOWLEDGEMENTS

We thank Dr R. Lavery for helpful discussions and encouragement. This study was performed using computational resources of the CLAMV (Computational Laboratories for Analysis, Modeling and Visualization) at Jacobs University, Germany, and supported by a grant from the VolkswagenStiftung to M.Z. Funding to pay the Open Access publication charges for this article was provided by VolkswagenStiftung and Jacobs University Bremen.

Conflict of interest statement. None declared.

REFERENCES

1. Hagerman,P.J. (1990) Flexibility of DNA. *Annu. Rev. Biochem.*, **59**, 755–781.
2. Travers,A.A. (2004) The structural basis of DNA flexibility. *Phil. Trans. R. Soc. Lond. A.*, **362**, 1423–1438.
3. Rhodes,D., Schwabe,J.W., Chapman,L. and Fairall,L. (1996) Towards an understanding of protein-DNA recognition. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **351**, 501–509.
4. Olson,W.K., Gorin,A.A., Lu,X.J., Hock,L.M. and Zhurkin,V.B. (1998) DNA sequence dependent deformability deduced from protein–DNA crystal complexes. *Proc. Natl Acad. Sci. USA*, **95**, 11163–11168.

5. Deremble,C. and Lavery,R. (2005) Macromolecular recognition. *Curr. Opin. Struct. Biol.*, **15**, 171–175.
6. Zakrzewska,K. (2003) DNA deformation energetics and protein binding. *Biopolymers*, **70**, 414–423.
7. von Hippel,P.H., Rees,W.A., Rippe,K. and Wilson,K.S. (1996) Specificity mechanisms in the control of transcription. *Biophys. Chem.*, **59**, 231–246.
8. Hunter,C.A. (1993) Sequence-dependent DNA structure. The role of base stacking interactions. *J. Mol. Biol.*, **230**, 1025–1054.
9. Calladine,C.R., Drew,H., Luisi,B. and Travers,A. (2004) *Understand DNA: the Molecule and How it Works*. Academic Press, London, pp. 31–54.
10. Dickerson,R.E. and Chiu,T.K. (1997) Helix Bending as a Factor in Protein/DNA recognition. *Biopolymers*, **44**, 361–403.
11. Gromiha,M.M., Siebers,J.G., Selvaraj,S., Kono,H. and Sarai,A. (2005) Role of inter and intramolecular interactions in protein-DNA recognition. *Gene*, **364**, 108–113.
12. Crothers,D.M. (1998) DNA curvature and deformation in protein-DNA complexes: a step in the right direction. *Proc. Natl Acad. Sci. USA*, **95**, 15163–15165.
13. Koo,H.S., Drak,J., Rice,J.A. and Crothers,D.M. (1990) Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry*, **29**, 4227–4234.
14. Hagerman,P.J. (1985) Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. *Biochemistry*, **24**, 7033–7037.
15. Koo,H.S., Wu,H.M. and Crothers,D.M. (1986) DNA bending at adenine.thymine tracts. *Nature*, **320**, 501–506.
16. Hagerman,P.J. (1986) Sequence-directed curvature of DNA. *Nature*, **321**, 449–450.
17. Stefl,R., Wu,H., Ravindranathan,S., Sklenar,V. and Feigon,J. (2004) DNA A-tract bending in three dimensions: solving the dA4T4 vs. dT4A4 conundrum. *Proc. Natl Acad. Sci. USA*, **101**, 1177–1182.
18. Wu,H.M. and Crothers,D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature*, **308**, 509–513.
19. Arnott,S., Chandrasekaran,R., Hall,I.H. and Puigjaner,L.C. (1983) Heteronomous DNA. *Nucleic Acids Res.*, **11**, 4141–4155.
20. Park,Y.W. and Breslauer,K.J. (1991) A spectroscopic and calorimetric study of the melting behaviors of a ''bent'' and a ''normal'' DNA duplex: [d(GA4T4C)]2 versus [d(GT4A4C)]2. *Proc. Natl Acad. Sci. USA*, **88**, 1551–1555.
21. Selsing,E., Wells,R.D., Alden,C.J. and Arnott,S. (1979) Bent DNA: visualization of a base-paired and stacked A-B conformational junction. *J. Biol. Chem.*, **254**, 5417–5422.
22. Trifonov,E.N. and Sussman,J.L. (1980) The pitch of chromatin DNA is reflected in its nucleotide sequence. *Proc. Natl Acad. Sci. USA*, **77**, 3816–3820.
23. Ulanovsky,L.E. and Trifonov,E.N. (1987) Estimation of wedge components in curved DNA. *Nature*, **326**, 720–722.
24. Dickerson,R.E., Goodsell,D. and Kopka,M.L. (1996) MPD and DNA bending in crystals and in solution. *J. Mol. Biol.*, **256**, 108–125.
25. Barbic,A., Zimmer,D.P. and Crothers,D.M. (2003) Structural origins of adenine-tract bending. *Proc. Natl Acad. Sci. USA*, **100**, 2369–2373.
26. Bolshoy,A., McNamara,P., Harrington,R.E. and Trifonov,E.N. (1991) Curved DNA without A-A: experimental estimation of all 16 DNA wedge angles. *Proc. Natl Acad. Sci. USA*, **88**, 2312–2316.
27. Cloutier,T.E. and Widom,J. (2004) Spontaneous sharp bending of double-stranded DNA. *Mol. Cell*, **14**, 355–362.
28. Lankas,F., Lavery,R. and Maddocks,J.H. (2006) Kinking occurs during molecular dynamics simulations of small DNA minicircles. *Structure*, **14**, 1527–1534.
29. Wiggins,P.A., Van Der Heijden,T., Moreno-Herrero,F., Spakowitz,A., Phillips,R., Widom,J., Ceekers,C. and Nelson,P.C. (2006) High flexibility of DNA on short length scales probed by atomic force microscopy. *Nat. Nanotechnol.*, **1**, 137–141.
30. Strahs,D. and Schlick,T. (2000) A-tract bending: insights into experimental structures by computational models. *J. Mol. Biol.*, **301**, 643–663.
31. Sprous,D., Young,M.A. and Beveridge,D.L. (1999) Molecular dynamics studies of axis bending in d(G5-(GA4T4C)2-C5) and d(G5-(GT4A4C)2-C5): effects of sequence polarity on DNA curvature. *J. Mol. Biol.*, **285**, 1623–1632.
32. McConnell,K.J. and Beveridge,D.L. (2001) Molecular dynamics simulations of B'-DNA: sequence effects on A-tract-induced bending and flexibility. *J. Mol. Biol.*, **314**, 23–40.
33. Beveridge,D.L., Dixit,S.B., Barreiro,G. and Thayer,K.M. (2004) Molecular dynamics simulations of DNA curvature and flexibility: helix phasing and premelting. *Biopolymers*, **73**, 380–403.
34. Dixit,S.B., Pitici,F. and Beveridge,D.L. (2004) Structure and axis curvature in two dA6 xdT6 DNA oligonucleotides: comparison of molecular dynamics simulations with results from crystallography and NMR spectroscopy. *Biopolymers*, **75**, 468–479.
35. Mazur,A.K. (2006) Evaluation of elastic properties of atomistic DNA models. *Biophys. J.*, **91**, 4507–4518.
36. Ulyanov,N.B. and Zhurkin,V.B. (1984) Sequence-dependent anisotropic flexibility of B-DNA. A conformational study. *J. Biomol. Struct. Dyn.*, **2**, 361–385.
37. Sanghani,S.R., Zakrzewska,K., Harvey,S.C. and Lavery,R. (1996) Molecular modelling of (A4T4NN)*n* and (T4A4NN)*n*: sequence elements responsible for curvature. *Nucleic Acids Res.*, **24**, 1632–1637.
38. Lavery,R., Zakrzewska,K. and Sklenar,H. (1995) JUMNA: Junction Minimisation of Nucleic Acids. *Comp. Phys. Comm.*, **91**, 135–158.
39. Babcock,M.S., Pednault,E.P. and Olson,W.K. (1994) Nucleic acid structure analysis. Mathematics for local Cartesian and helical structure parameters that are truly comparable between structures. *J. Mol. Biol.*, **237**, 125–156.
40. von Kitzing,E. and Diekmann,S. (1987) Molecular mechanics calculations of dA12.dT12 and of the curved moleculed(GCTCG AAAAA)4.d(TTTTTCGAGC)4. *Eur. Biophys. J.*, **15**, 13–26.
41. Srinivasan,R., Geetha,V., Seetharaman,J. and Mohan,S. (1993) A unique or essentially unique single parametric characterisation of biopolymeric structures. *J. Biomol. Struct. Dyn.*, **11**, 583–596.
42. Mazur,J. and Jernigan,R.L. (1995) Comparison of rotation models for describing DNA conformations: Application to static and polymorphic forms. *Biophys. J.*, **68**, 1472–1489.
43. van Dam,H., Louck,J.D. and Biederharn,L.C. (1990) Rotation and angular momentum. In Lerner,R.G. and Trigg,G.L. (eds), *Encyclopedia of Physics, 2nd edn.* VCH Publishers, New York, pp. 1076–1081.
44. Goldstein,H. (1981) *Classical Mechanics, 2nd edn*. Addison Wesley publishing Co., New York, pp. 255–310.
45. Cheatham,T.E., Cieplak,P. and Kollman,P.A. (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.*, **16**, 845–862.
46. Hawkins,G.D., Cramer,C.J. and Truhlar,D.J. (1995) Pairwise solute descreening of solute charges from a dielectric continuum. *Chem. Phys. Lett.*, **246**, 122–129.
47. Zacharias,M. (2001) Conformational analysis of DNA-trinucleotide-hairpin-loop structures using a continuum solvent model. *Biophys. J.*, **80**, 2350–2363.
48. Lavery,R. and Sklenar,H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.*, **6**, 63–91.
49. Lavery,R. and Sklenar,H. (1989) Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.*, **6**, 655–667.
50. Case,D.A., Darden,T.A., Cheatham,T.E. III, Simmerling,C.L., Wang,J., Duke,R.E., Luo,R., Maerz,K.M., Wang,B. *et al.* (2004) AMBER 8, University of California, San Francisco.
51. Beveridge,D.L., Barreiro,G., Byun,K.S., Case,D.A., Cheatham 3rd,T.E., Dixit,S.B., Giudice,E., Lankas,F., Lavery,R. *et al.* (2004) Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and results on d(CpG) steps. *Biophys. J.*, **87**, 3799–3813.
52. Dixit,S.B., Beveridge,D.L., Case,D.A., Cheatham 3rd,T.E., Giudice,E., Lankas,F., Lavery,R., Maddocks,J.H., Osman,R. *et al.* (2005) Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. II: sequence context effects on the dynamical structures of the 10 unique dinucleotide steps. *Biophys. J.*, **89**, 3721–3740.
53. Cornell,W.D., Cieplak,P., Bayley,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. *et al.*

(1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
54. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
55. Darden,T., York,D. and Pedersen,L. (1993) Particle mesh Ewald: an N*log(N) method for computing Ewald sums. *J. Chem. Phys.*, **98**, 10089–10092.
56. Ryckaert,J.P., Ciccotti,G. and Berendsen,H.J.C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, **23**, 327–341.
57. Torrie,G.M. and Valleau,J.P. (1977) Non physical sampling distribution in Monte Carlo free-energy estimation: umbrella sampling. *J. Comput. Phys.*, **23**, 187–199.
58. Roux,B. (1995) The calculation of the potential of mean force using computer simulations. *Comput. Phys. Comm.*, **91**, 275–282.
59. Giudice,E., Varnai,P. and Lavery,R. (2003) Base pair opening within B-DNA: free energy pathways for Gc and AT pairs from umbrella sampling simulations. *Nucleic Acids Res.*, **31**, 1437–1443.
60. Kumar,S., Rosenberg,J.M., Bouzida,D., Swendsen,R.H. and Kollman,P.A. (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.*, **13**, 1011–1021.
61. MacDonald,D., Herbert,K., Zhang,X., Pologruto,T. and Lu,P. (2001) Solution structure of an A-tract DNA bend. *J. Mol. Biol.*, **306**, 1081–1098.
62. Hizver,J., Rozenberg,H., Frolow,F., Rabinovich,D. and Shakked,Z. (2001) DNA bending by an adenine–thymine tract and its role in gene regulation. *Proc. Natl Acad. Sci. USA*, **98**, 8490–8495.
63. Rozenberg,H., Rabinovich,D., Frolow,F., Hegde,R.S. and Shakked,Z. (1998) Structural code for DNA recognition revealed in crystal structures of papillomavirus E2-DNA targets. *Proc. Natl Acad. Sci. USA*, **95**, 15194–15199.
64. Kim,S.S., Tam,J.K., Wang,A.F. and Hegde,R.S. (2000) The structural basis of DNA target discrimination by papillomavirus E2 proteins. *J. Biol. Chem.*, **275**, 31245–31254.
65. Hegde,R.S., Grossman,S.R., Laimins,L.A. and Sigler,P.B. (1992) Crystal structure at 1.7 A of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature*, **359**, 505–512.
66. de Los Santos,C., El-Khateeb,M., Rege,P., Tian,K. and Johnson,F. (2004) Impact of the C1 configuration of abasic sites on DNA duplex structure. *Biochemistry*, **43**, 15349–15357.
67. Sanchez,A.M., Volk,D.E., Gorenstein,D.G. and Lloyd,R.S. (2003) Initiation of repair of A/G mismatches is modulated by sequence context. *DNA Repair*, **2**, 863–878.
68. Boulard,Y., Cognet,J.A. and Fazakerley,G.V. (1997) Solution structure as a function of pH of two central mismatches, C. T and C. C, in the 29 to 39 K-ras gene sequence, by nuclear magnetic resonance and molecular dynamics. *J. Mol. Biol.*, **268**, 331–347.
69. Isaacs,R.J., Rayens,W.S. and Spielmann,H.P. (2002) Structural differences in the NOE-derived structure of G-T mismatched DNA relative to normal DNA are correlated with differences in (13)C relaxation-based internal dynamics. *J. Mol. Biol.*, **319**, 191–207.
70. Djuranovic,D. and Hartmann,B. (2005) Molecular dynamics studies on free and bound targets of the bovine papillomavirus type I e2 protein: the protein binding effect on DNA and the recognition mechanism. *Biophys. J.*, **89**, 2542–2551.
71. Byun,K.S. and Beveridge,D.L. (2004) Molecular dynamics simulations of papilloma virus E2 DNA sequences: dynamical models for oligonucleotide structures in solution. *Biopolymers*, **73**, 369–379.
72. Rohs,R., Sklenar,H. and Shakked,Z. (2005) Structural and energetic origins of sequence-specific DNA bending: Monte-Carlo simulations of papillomavirus E2-DNA binding sites. *Structure*, **13**, 1499–1509.
73. Shatzky-Schwartz,M., Arbuckle,N.D., Eisenstein,M., Rabinovich,D., Bareket-Samish,A., Haran,T.E., Luisi,B.F. and Shakked,Z. (1997) X-ray and solution studies of DNA oligomers and implications for the structural basis of A-tract-dependent curvature. *J. Mol. Biol.*, **267**, 595–623.
74. Mack,D.R., Chiu,T.K. and Dickerson,R.E. (2001) Intrinsic bending and deformability at the T-A step of CCTTTAAAGG: a comparative analysis of TA and AT steps within A-tracts. *J. Mol. Biol.*, **312**, 1037–1049.
75. Demple,B. and Harrison,L. (1994) Repair of oxidative damage to DNA: enzymology and biology. *Annu. Rev. Biochem.*, **63**, 915–948.
76. Elledge,S.J. (1996) Cell cycle checkpoints: preventing an identity crisis. *Science*, **274**, 1664–1672.
77. Barsky,D., Foloppe,N., Ahmadia,S., Wilson,D.M. III and MacKerell.A.D. Jr, (2000) New insights into the structure of abasic DNA from molecular dynamics simulations. *Nucleic Acids Res.*, **28**, 2613–2626.
78. Fujimoto,H., Pinak,M., Nemoto,T., O'Neill,P., Kume,E., Saito,K. and Maekawa,H. (2005) Molecular dynamics simulation of clustered DNA damage sites containing 8-oxoguanine and abasic site. *J. Comput. Chem.*, **26**, 788–798.
79. Ayadi,L., Coulombeau,C. and Lavery,R. (1999) Abasic sites in duplex DNA: molecular modeling of sequence-dependent effects on conformation. *Biophys. J.*, **77**, 3218–3226.
80. Hagerman,P.J. (1988) Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.*, **17**, 265–286.
81. Diekmann,S., Mazzarelli,J.M., McLaughlin,L.W., von Kitzing,E. and Travers,A.A. (1992) DNA curvature does not require bifurcated hydrogen bonds or pyrimidine methyl groups. *J. Mol. Biol.*, **225**, 729–738.
82. Calladine,C.R. (1982) Mechanics of sequence-dependent stacking of bases in B-DNA. *J. Mol. Biol.*, **161**, 343–352.