Theoretical Studies of the Possible Origin of Intrinsic Static Bends in Double Helical DNA

Alexey K. Mazur*

Contribution from the Laboratoire de Biochimie Théorique, CNRS UPR9080, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, Paris, 75005 France

Received May 10, 2000. Revised Manuscript Received October 20, 2000

Abstract: The macroscopic curvature induced in the double helical B-DNA by regularly repeated adenine tracts (A-tracts) is a long-known, but still unexplained, phenomenon. This effect plays a key role in DNA studies because it is unique in the amount and the variety of the available experimental information and, therefore, is likely to serve as a gate to unknown general mechanisms of recognition and regulation of genome sequences. In this paper, conformations of a 25-mer A-tract repeat have been studied by molecular dynamics simulations. It is shown that properly directed static curvature emerges spontaneously in independent MD trajectories starting from straight canonical A- and B-DNA forms. Dynamics converge to the same bent state in conditions excluding any initial bias except the base pair sequence. The ensemble of curved conformations, however, appears microscopically heterogeneous, in contradiction to all existing theoretical models of bending. Analysis of these unexpected observations leads to a new, significantly different hypothesis of the possible mechanism of sequence-directed bends in double helical DNA.

Introduction

Strong static bends in DNA were originally discovered in sequences containing regular repeats of A_nT_m, with $n + m > 3$, called A-tracts.[1] The A-tract-induced bending plays an important role in structural DNA studies because it is the strongest such effect and the most thoroughly investigated. It was first noticed in restriction fragments from the kinetoplast body of *Leishmania tarentolae*[2,3] and confirmed by electric birefringence decay[4] and electron microscopy.[5] Every A-tract deviates the helical axis by approximately 18° toward its minor groove,[6,7] and if they are repeated in the sequence in phase with the helical screw, the macroscopic curvature emerges. The curvature is reduced with temperature and ionic force, while divalent metal ions increase it in some sequences, but reduce in others.[8,9] It strongly varies with the length and composition of A-tracts and depends on the sequences between them.[6] Detailed analysis of these results can be found in comprehensive reviews published in various years.[1,10−14]

According to many independent experimental observations, the structure of A-tract sequences should differ significantly from the "random" B-DNA. In solution, the poly-dA double helix is overwound to a twist of around 36° from around 34° of a random sequence.[15−17] The fiber diffraction data suggest consistently that the poly-dA double helix is characterized by a very narrow minor groove and a high propeller twist.[18−20] Yet another distinction is an apparently large negative inclination of base pairs.[19] Several A-tracts that are avaliable in single-crystal structures of B-DNA oligomers have irregular conformations but exhibit similar trends toward their centers.[21−25] Interestingly, in X-ray structures, A-tracts look less prone to bending than do other sequences. Some indirect observations also support this view; notably, poly-dA fragments move faster than random DNA in gel migration assays[6] and avoid wrapping around nucleosome particles.[26,27]

* Fax: +33[0]1 58 41 50 26. Email: alexey@ibpc.fr.

(1) Crothers, D. M.; Shakked, Z. In *Oxford Handbook of Nucleic Acid Structure*; Neidle, S., Ed.; Oxford University Press: New York, 1999; pp 455−470.
(2) Marini, J. C.; Levene, S. D.; Crothers, D. M.; Englund, P. T. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 7664−7668.
(3) Wu, H.-M.; Crothers, D. M. *Nature* **1984**, *308*, 509−513.
(4) Hagerman, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4632−4636.
(5) Griffith, J.; Bleyman, M.; Rauch, C. A.; Kitchin, P. A.; Englund, P. T. *Cell* **1986**, *46*, 717−724.
(6) Koo, H.-S.; Wu, H.-M.; Crothers, D. M. *Nature* **1986**, *320*, 501−506.
(7) Zinkel, S. S.; Crothers, D. M. *Nature* **1987**, *328*, 178−181.
(8) Diekmann, S.; Wang, J. C. *J. Mol. Biol.* **1985**, *186*, 1−11.
(9) Diekmann, S. *Nucleic Acids Res.* **1987**, *15*, 247−265.
(10) Diekmann, S. In *Nucleic Acids and Molecular Biology*; Eckstein, F., Lilley, D. M. J., Eds.; Springer-Veralg: Berlin Heidelberg, 1987; *Vol. 1*; pp 138−156.
(11) Hagerman, P. J. *Annu. Rev. Biochem.* **1990**, *59*, 755−781.
(12) Crothers, D. M.; Haran, T. E.; Nadeau, J. G. *J. Biol. Chem.* **1990**, *265*, 7093−7096.
(13) Crothers, D. M.; Drak, J. *Methods Ensymol.* **1992**, *212*, 46−71.
(14) Olson, W. K.; Zhurkin, V. B. In *Structure and Dynamics. Vol. 2: Proceedings of the Ninth Conversation, State University of New York, Albany, NY, 1995*; Sarma, R. H., Sarma, M. H., Eds.; Adenine Press: New York, 1996; pp 341−370.
(15) Peck, L. J.; Wang, J. C. *Nature* **1981**, *292*, 375−378.
(16) Rhodes, D.; Klug, A. *Nature* **1981**, *292*, 378−380.
(17) Strauss, F.; Gaillard, C.; Prunell, A. *Eur. J. Biochem.* **1981**, *118*, 215−222.
(18) Aymami, J.; Coll, M.; Frederick, C. A.; Wang, A. H.-J.; Rich, A. *Nucleic Acids Res.* **1989**, *8*, 3229−3245.
(19) Lipanov, A. A.; Chuprina, V. P.; Alexeev, D. G.; Skuratovskii, I. Y. *J. Biomol. Struct. Dyn.* **1990**, *7*, 811−826.
(20) Chandrasekaran, R.; Radha, A. *J. Biomol. Struct. Dyn.* **1992**, *10*, 153−167.
(21) Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761−786.
(22) Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. *Nature* **1987**, *330*, 221−226.
(23) DiGabriele, A. D.; Sanderson, M. R.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 1816−1820.
(24) Edwards, K. J.; Brown, D. G.; Spink, N.; Skelly, J. V.; Neidle, S. *J. Mol. Biol.* **1992**, *226*, 1161−1173.
(25) DiGabriele, A. D.; Steitz, T. A. *J. Mol. Biol.* **1993**, *231*, 1024−1039.
(26) Simpson, R. T.; Künzler, P. *Nucleic Acids Res.* **1979**, *6*, 1387−1415.

10.1021/ja001605y CCC: $19.00
Published on Web 12/06/2000

Despite a large body of experimental information that has accumulated during the last twenty years, the possible physical mechanism of this effect remains unclear. Because every base pair in the B-DNA stack interacts only with the two neighbors, any sequence specificity should result from the short-range interactions confined in a single-base-pair step. Nonlocal effects are also possible, however, due to the base-backbone interactions and the propagation of correlations along the backbone. The initial experimental data on A-tract bending were interpreted in terms of two alternative mechanisms, namely, the wedge model[28] and the junction model.[29] Both had to be modified significantly as and when new information appeared, and other theories were also proposed, but none of them explains all experimental data.[1] The overall pattern became yet more confusing when it was found that certain non-A-tract sequences also exhibit distinguishable curvature.[30]

As a theoretical problem, the sequence-dependent DNA bending presents a challenge that is analogous to the protein folding concerning both the unclear physical origin and the potential practical importance. The atom level molecular modeling is virtually the only theoretical approach that is potentially able to reveal the underlying mechanisms and predict the sequence-dependent macroscopic DNA forms. If a model system based upon general atom−atom potentials could reproducibly yield properly curved DNA conformations, one must be able to disclose the mechanism of bending in the model and hypothesize, with certain confidence, that a similar mechanism may take place in nature. The foregoing scheme, however, is too difficult, and until now, most of the modeling studies used other strategies. Much has been learned about DNA bending mechanics by using energy minimization[31−34] and Monte Carlo.[35] The possibilities of these methods are limited by the multiple minima problem, especially when it is necessary to consider solvent molecules as well. In recent years, multi-nanosecond free MD simulations of long DNA fragments became feasible, owing to the progress in computer power and improved full atom force fields.[36,37] A few reported simulations of phased A-tract sequences demonstrated that, without any a priori bias, the DNA double helix bends anisotropically and spontaneously acquires certain structural features known from experiments.[38,39] The free MD simulations represent a promising line of research in this field, and we continue it here by using the recently proposed minimal model of B-DNA.[40]

The minimal B-DNA consists of a double helix with the minor groove filled with water. It does not use explicit counterions and treats long-range electrostatics semiempirically, that is, with a distance dependent dielectric permittivity and scaled phosphate charges. Despite these drastic reductions, the minimal model gives B-DNA conformations which compare well with experimental data, and moreover, it is free from a systematic negative bias of the average twist observed in conventional simulations with AMBER94 potentials.[41] It was interesting to check if this model can reproduce the A-tract bending and how well its results agree with different theoretical mechanisms. Figure 1 shows how the DNA sequences were constructed. In a series of preliminary tests with short oligomers, we looked for A-tracts that would readily adopt in MD their characteristic conformations with narrow minor grooves. Originally, we took a 15-mer from the bending center in Figure 1, which is the first curved DNA fragment studied in vitro,[3] but found that its left-hand A-tract behaves much better than the right one. This might be due to an intrinsic difference between 3′- and 5′-end A-tracts, therefore, the "good" A-tract was inverted and the sequence continued by using the same AAAATAG motive. According to experiments, inversion of A-tracts should not affect bending.[6] We considered it important to place A-tracts at both ends in order to specify the boundary conditions. Because the A-tracts have a characteristic local structure, the boundary conditions could be at least qualitatively checked, which would not be the case for GC-rich sequences.

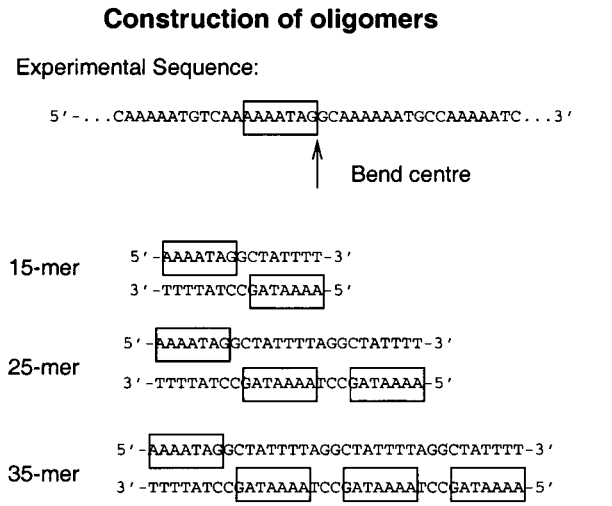

Figure 1. Construction of oligomers with A-tract repeats. The heptamer motive AAAATAG is taken from the experimental sequence of the trypanosome kinetoplast DNA.[3] The centers of the A-tracts are shifted by approximately 10 base pairs in phase with the helical screw.

All three sequences shown in Figure 1 exhibited static bending in simulations, but here we consider only the 25-mer AAAAT-AGGCTATTTTAGGCTATTTT which was studied most extensively. We demonstrate spontaneous development of stable static curvature in good agreement with experimental data concerning both the bending direction and magnitude and compare our results with predictions of theoretical models of bending. This analysis leads to a new, significantly different hypothesis of the physical origin of intrinsic bends in the DNA molecule.

Methods and Simulation Protocols

Molecular dynamics simulations were performed with the internal coordinate method (ICMD),[42,43] including a special technique for

(27) Rhodes, D. *Nucleic Acids Res.* **1979**, *6*, 1805−1816.
(28) Trifonov, E. N.; Sussman, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 3816−3820.
(29) Levene, S. D.; Crothers, D. M. *J. Biomol. Struct. Dyn.* **1983**, *1*, 429−435.
(30) Bolshoy, A.; McNamara, P.; Harrington, R. E.; Trifonov, E. N. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2312−2316.
(31) Zhurkin, V. B.; Lysov, Y. P.; Ivanov, V. I. *Nucleic Acids Res.* **1979**, *6*, 1081−1096.
(32) von Kitzing, E.; Diekmann, S. *Eur. Biophys. J.* **1987**, *14*, 13−26.
(33) Chuprina, V. P.; Abagyan, R. A. *J. Biomol. Struct. Dyn.* **1988**, *1*, 121−138.
(34) Sanghani, S. R.; Zakrzewska, K.; Harvey, S. C.; Lavery, R. *Nucleic Acids Res.* **1996**, *24*, 1632−1637.
(35) Zhurkin, V. B.; Ulyanov, N. B.; Gorin, A. A.; Jernigan, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7046−7050.
(36) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(37) MacKerell, A. D., Jr; Wiórkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.
(38) Young, M. A.; Beveridge, D. L. *J. Mol. Biol.* **1998**, *281*, 675−687.
(39) Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Mol. Biol.* **1999**, *285*, 1623−1632.
(40) Mazur, A. K. *J. Am. Chem. Soc.* **1998**, *120*, 10928−10937.
(41) Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.
(42) Mazur, A. K.; Abagyan, R. A. *J. Biomol. Struct. Dyn.* **1989**, *6*, 815−832.

flexible sugar rings.[44] The time step was 10 fs, and the effective inertia of planar sugar angles was increased by 140 amu Å^2. AMBER94[36,41] force field and atom parameters were used with TIP3P water[45] and no cut off schemes. The starting fiber A- and B-DNA models were constructed from the published atom coordinates,[46] with only the B-DNA structure energy-minimized before hydration. The hydration protocol[40] requires around 16 water molecules per base pair for B-DNA. For the A-DNA start, a somewhat smaller initial number of waters was increased after equilibration to make it the same for all trajectories. All other conditions and protocols were as described earlier.[40] Programs Curves,[47] XmMol,[48] and Mathematica by Wolfram Research, Inc., were employed for graphics and data analysis.

During the runs, after every 200 ps, water positions were checked in order to identify those penetrating into the major groove and those completely separated. These molecules, if found, were removed and next reintroduced in simulations by putting them with zero velocities at random positions around the hydrated duplex, so that they could readily re-join the core system. This procedure ensures stable conditions, notably, a constant number of molecules in the minor groove hydration cloud and the absence of water in the major groove, which is necessary for fast sampling (unpublished results of the author). The interval of 200 ps between the checks is small enough to ensure that on average less than one molecule is repositioned and, therefore, the perturbation introduced is considered negligible.

Results and Discussion

Convergence of Trajectories. The concept of static bending implies that in dynamics, the structure should fluctuate around a state with a distinguishable bend and a definite bending direction. Because any MD simulation is limited in time, there is no way to prove rigorously that some specific conformation is representative. A certain degree of confidence can be achieved, however, if independent trajectories converge to the same state. To check this, three long trajectories were computed from different initial conformations. The first trajectory, referred to below as TJBa, started from the fiber canonical B-DNA structure and continued to 10 ns. When it appeared that TJBa converged to a bent state, two more trajectories (TJBb and TJA) were computed for verification. TJBb started with random velocities from a reminimized straight conformation that was taken after the equilibration phase of TJBa and was also continued to 10 ns. The last 20 ns trajectory (TJA) started from the fiber canonical A-DNA conformation to get rid of any initial bias implicitly involved in the choice of the initial state.

The following three figures display the conformational dynamics in TJBa. The two surface plots in Figure 2 exhibit the time evolution of the shape of the helical axis. It is clear that the molecule maintained a planar bent shape during a substantial part of the trajectory, with an increase in bending during the last 5 ns, as shown by the *x*-surface. This occurred after a long period of fluctuations near a straight conformation, which excludes an occasional bending caused by initial conditions or equilibration procedures. In the perpendicular *y*-projection, the helical axis is locally wound, but straight on average. The *y*-projection sometimes reveals two local maxima between A-tracts, which correspond to two independent bends with slightly divergent directions. One may note also that there were at least two relatively long periods when the axis was almost perfectly straight, namely, around 3 ns and during the fifth nanosecond.

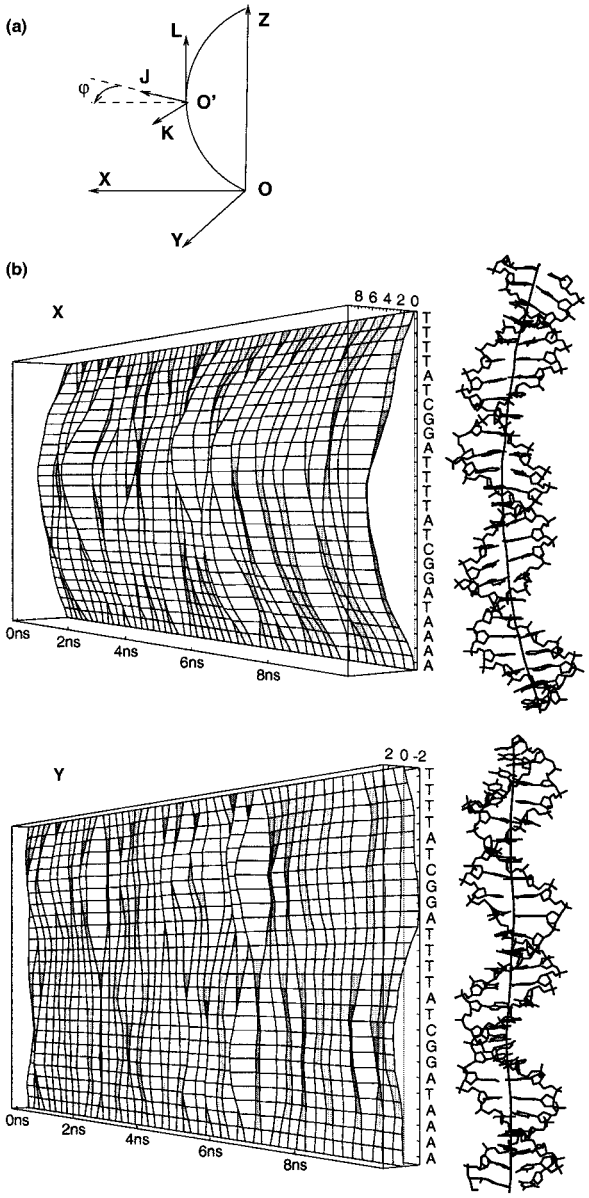

Figure 2. (a) Geometric constructions used in evaluating DNA bending dynamics. The axis of the curved double helix is computed as the best fit common axis of coaxial cylindrical surfaces passing through sugar atoms, which gives solutions close to those produced by the Curves algorithm.[47] This space curve is rotated with the two ends fixed at OZ to put the middle point in plane *XOZ*. Note that this procedure does not keep the structures superimposed, and the same curve can correspond to different bending directions. The bending direction is measured by the angle *n* between plane *XOZ* and the **J** vector of the local DNA coordinate frame.[66] This vector points to the major DNA groove along the short axis of the central base pair. Therefore, the zero φ value corresponds to the overall bend toward the minor groove at this point. (b) The time evolution of the overall shape of the helical axis in TJBa. The X and Y surfaces are constructed by using projections upon planes *XOZ* and *YOZ*, respectively, in plate (a). Any time section of these surfaces gives the corresponding projection averaged over a time window of 75 ps. The horizontal deviation is given in angströms and, for clarity, its relative scale is increased 2 times with respect to the true DNA length. Shown on the right are the two perpendicular views of the last 1-ns average conformation in the orientation corresponding to that in the surface plots at the end of the trajectory.

The time variation of the rmsd from the canonical B-DNA is shown in Figure 3. Its value fell to 2 Å during straightenings and went beyond 5 Å only in strongly bent structures. The same figure displays the evolution of the minor groove profile. It is

(43) Mazur, A. K. *J. Comput. Chem.* **1997**, *18*, 1354−1364.
(44) Mazur, A. K. *J. Chem. Phys.* **1999**, *111*, 1407−1414.
(45) Jorgensen, W. L. *J. Am. Chem. Soc.* **1981**, *103*, 335−340.
(46) Arnott, S.; Hukins, D. W. L. *Biochem. Biophys. Res. Communs.* **1972**, *47*, 1504−1509.
(47) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91.
(48) Tuffery, P. *J. Mol. Graph.* **1995**, *13*, 67−72.

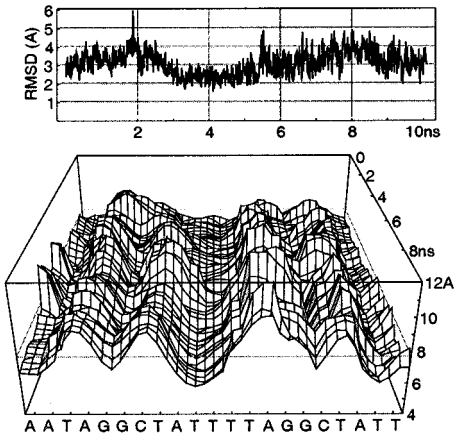

Figure 3. Time evolution of the non-hydrogen atom rmsd from the canonical B-DNA form and of the profile of the minor groove in TJBa. The surface is formed by 75-ps time-averaged successive minor groove profiles, with that on the front face corresponding to the final DNA conformation. The groove width is evaluated by using space traces of C5′ atoms as described elsewhere.[67] Its value is given in angströms and the corresponding canonical B-DNA level of 7.7 Å is marked by the straight dotted lines on the faces of the box.

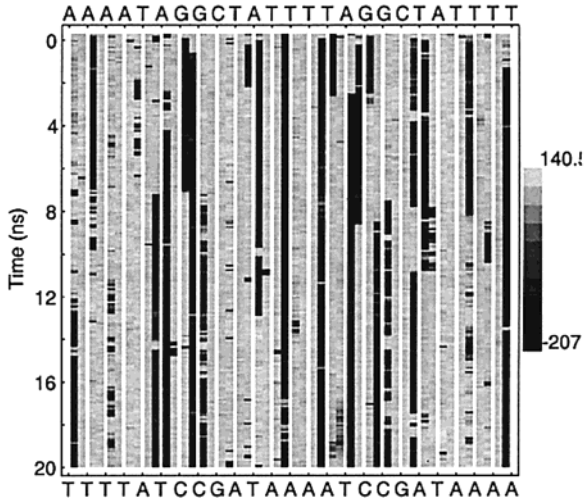

Figure 4. Dynamics of B_I and B_{II} backbone conformers in TJA. The B_I and B_{II} conformations are distinguished by the values of two consecutive backbone torsions, ϵ and ζ. In a transition, they change concertedly from (t,g^-) to (g^-,t). The difference $\zeta-\epsilon$ is, therefore, positive in the B_I state and negative in B_{II}, and it is used as a monitoring indicator, with the corresponding gray scale levels shown on the right. Each base pair step is characterized by a column consisting of two sub-columns, with the left sub-columns referring to the sequence written at the top in 5′−3′ direction from left to right. The right sub-columns refer to the complementary sequence shown at the bottom.

seen that the overall groove shape established during the first 2 ns and remained stable later, with insignificant local fluctuations. In all A-tracts, the groove strongly narrows toward 3′ ends. During the last nanoseconds of the three trajectories, its average width in the narrowest places was between 6.33 and 6.45 Å.[49] These values agree well with experimental data; notably, in the experimental single crystal dodecamer structures, the A-tract minor grooves, when measured by the same method, narrow down to 6.12−6.51 Å.[21−25]

The $B_I \leftrightarrow B_{II}$ backbone dynamics is exhibited in Figure 4. The overall pattern involves fluctuations in all time scales accessible in simulations. Many $B_I \leftrightarrow B_{II}$ transitions are reversed within hundreds of picoseconds, but there are also very long-

(49) See the supplementary matherial to this paper.

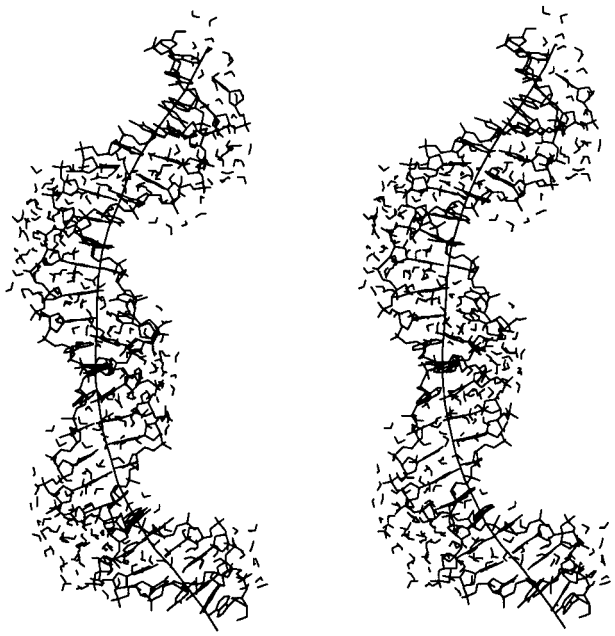

Figure 5. Stereo snapshot of the system at approximately 8.5 ns of TJBb.

living conformers. Examples of concerted transitions may be seen when a given step switches from B_{II} to B_I simultaneously with an opposite transition in one of the neighboring steps. There are sites where either B_I or B_{II} states are referred. For instance, all TpT steps are always in B_I. However, in a few sites where the B_{II} conformation is referred, this is apparently determined by a broader sequence context. In A-tracts, the B_{II} conformers found in A-strands tend to alternate with B_I in consecutive ApA steps. The $B_I \leftrightarrow B_{II}$ activity remains high throughout the trajectory, which indicates that the seemingly static shape of the surfaces in Figures 2 and 3 actually corresponds to a slowly changing quasi-equilibrium state.

The foregoing results were qualitatively similar in all three trajectories.[49] Dynamics converged to excellent B-DNA structures concerning the helicoidal parameters, with the average twist and rise of 34.8° and 3.5 Å, respectively. In TJA, the transition to the B-form was completed during the first 100 ps. In conformations averaged over 1-ns intervals, the usual bending angle was 35−45°, with the rmsd from the canonical B-DNA around 3 Å. The corresponding maximal values were 71° and 6.5 Å, respectively. Figure 5 shows a snapshot of the absolute rmsd maximum that occurred around 8.5 ns of TJBb. The strong smooth bend toward the minor grooves of the three A-tracts is evident, with the bending angle around 61°.

The only important qualitative difference between the three trajectories was observed in the dynamics of the bend direction which are shown in Figure 6. In both TJBa and TJBb, the final bending direction occurred early and remained more or less stable, excluding the temporary increase of directional fluctuations in TJBa during straightening. The local amplitude of scattering in these plots presents a sensitive indicator of the magnitude of bending because in strongly bent states, the bending plane is particularly well-determined and the measured direction shows fewer fluctuations. Figure 6 indicates, therefore, that in both TJBa and TJBb, there was a broad but single range of angles around 60 ± 60° where strong bends occurred. In contrast, during the first 10 ns of TJA, the bending plane was also well-defined, but it made almost a half-turn with respect to the O′ coordinate frame in Figure 2. Thus, a transition between oppositely bent conformations occurred via a pathway

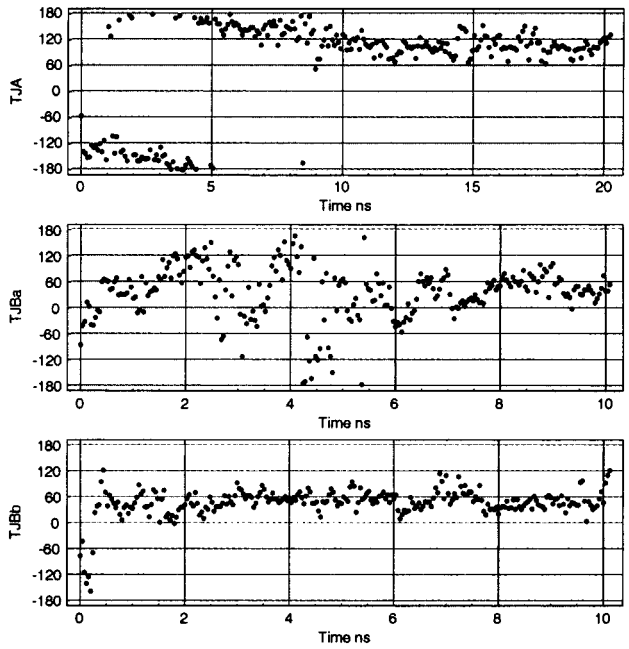

Figure 6. Time evolution of the bending direction in the tree trajectories. The bending direction is measured in degrees as shown in Figure 2.

that avoided the straight state. This rotation was very steady, almost regular, and it gradually slowed, becoming indistinguishable in the last 5 ns.

The above results show that the static bending emerges spontaneously in this model DNA fragment. Its direction is well-defined, and it can only be determined by the base pair sequence because the initial conformations were straight and symmetrical. The bending occurs toward the narrowed minor grooves of A-tracts and its magnitude corresponds to 13° per A-tract or 20° per a junction zone between them, which agrees well with experimental estimates.[1] The unexpectedly demonstrative pattern in Figure 6 strongly suggests that the quasi-equilibrium bent state represents a strong attractor of trajectories, with a basin of attraction comprising both canonical A- and B-DNA forms.

Comparison to Earlier Theories of Bending: Unexpected Micro-Heterogeneity of the Bent State. All theoretical explanations of the sequence-dependent DNA bending proposed during the last 20 years agree with some experimental data and disagree with the other. Below we check how the bending dynamics observed in MD agree with the most popular models, regardless of their experimental validation. Comparisons to experiments have been the subject of many reviews.[1,10−14]

The wedge model of DNA bending[28] followed the ideas developed in the 70s to explain the ability of the double helix to wrap around nucleosome particles. The first idea was that this can occur due to axis kinks phased with the helical screw.[31,50] The kinks are obtained by destacking in fable base pair steps. The second idea pointed out that it should be easy to bend the double helix without destacking by means of small local deformations.[51] The wedge model merges the two by postulating that in every specific dinucleotide, stacking is slightly nonparalle, and this causes bending in the same way as kinks do. This model can be further developed by increasing the number of wedge degrees of freedom and by considering triplets, tetraplets, and so forth instead of dinucleotides. Depending upon the specific wedge parameters, the curvature may appear inside A-tracts or between them[52,53] and even in sequences without A-tracts.[30] All such mechanisms can be united in one group that is characterized by the tacit emphasis on the stacking interactions as the single cause of bending.

The junction model[29] assumes that there is a distinct A-tract form of the double helix that is characterized by strong inclination of base pairs with respect to the helical axis. At the junction with the normal B-DNA, planar base pair stacking would induce a kink in the helical axis. This effect can be accounted for with the wedge model as well, but the junction theory points to the specific A-tract form of DNA as the principal physical cause. Within the framework of this theory, particular roles were sometimes attributed to bifurcated hydrogen bonds,[18] the water spine in the minor groove,[33] or the NH_2 groups of adenines.[32]

Both of these models predict that bending in sequence repeats must be accompanied by repetition of certain base pair orientations in identical sequence fragments, which can be checked in distributions of computed helicoidal parameters. Among them, inclination, roll, and tilt are the most important. The last two are directly related to bending. In an ideal stack of parallel base pairs, roll and tilt angles are zero. The positive and negative roll angles, in a given step, bend the stack toward the major and minor groove, respectively, while the tilt bends in the perpendicular plane. However, the bending direction fluctuates significantly even in the converged phase of dynamics (see Figure 6), and one may reasonably expect that large dynamic fluctuations could hide sequence references in helicoidal parameters. For a more sensible analysis, therefore, we checked substates in the zone $90 \pm 30°$, which was visited by all three trajectories, and took one strongly bent conformation from each trajectory. Figure 7 shows them in the schematic representation produced by the Curves program.[47] These states were energy minimized, with water included, to suppress thermal fluctuations. The minimization was nearly exhaustive, with the energy decrease beyond 1100 kcal/mol, which exceeds the average kinetic energy around 1060 kcal/mol available in the system. The corresponding trajectories stayed within 2 Å rmsd from these local minima during at least 1 ns. The structures are clearly bent in an identical direction, and at the same time, their shapes are rather different. Notably, in the TJBb structure, the helical axis has two distinct kinks that are centered approximately at the first and the third TpA steps. In the TJBa conformation the same two kinks look smoothed, but still distinguishable. Finally, the TJA structure is smoothly curved without kinks.

Figure 8 displays sequence variations of the main helicoidal parameters in the three structures. They obviously disagree with the views of all edge models, even with dinucleotide blocks replaced by multiplets. There is no significant similarity between identical multiplets in the three structures, which means that the sequence context determines the bend but does not fix the helicoidal parameters. Roll and tilt strongly fluctuate around zero, with little similarity between identical sequence fragments. Moreover, their signs generally do not correspond to the local bending direction. One clear example of such is given by the roll in the four TpA steps. In the TJBa structure, all four TpA steps have a strong positive roll. Paradoxically, two of these TpA steps have their major grooves almost exactly at the outside edge of the curved axis; that is, a high positive roll accompanies the bending in an exactly opposite direction. This paradox is readily resolved when one looks at the neighboring roll values.

(50) Crick, F. H. C.; Klug, A. *Nature* **1975**, *255*, 530−533.
(51) Namoradze, N. Z.; Goryunov, A. N.; Birshtein, T. M. *Biophys. Chem.* **1977**, *7*, 59−70.
(52) Maroun, R. C.; Olson, W. K. *Biopolymers* **1988**, *27*, 585−603.
(53) Goodsell, D. S.; Dickerson, R. E. *Nucleic Acids Res.* **1994**, *22*, 5947−5503.

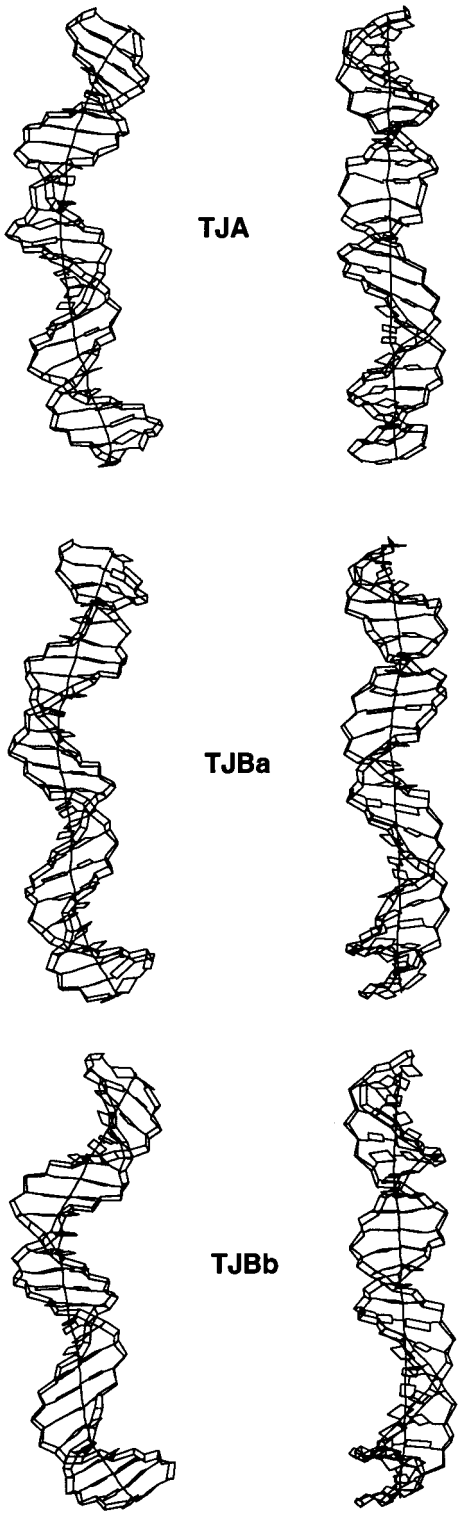

Figure 7. Representative bent energy minima from the three trajectories. To avoid artifact convergence of the bending planes, the structures have been separately superimposed with the straight canonical B-DNA. Two orthogonal views are shown.

A TpA step with a high positive roll is normally preceded or followed by a step with a low negative roll. The higher is the maximum, the lower is the neighboring minimum, so that the two nearly cancel each other. In the other two structures, the corresponding values are lower, and accordingly, they are higher in the neighboring steps. Other parameters also exhibit jumping alterations of high and low values in consecutive steps, which suggests that these fluctuations are correlated and tend to cancel.

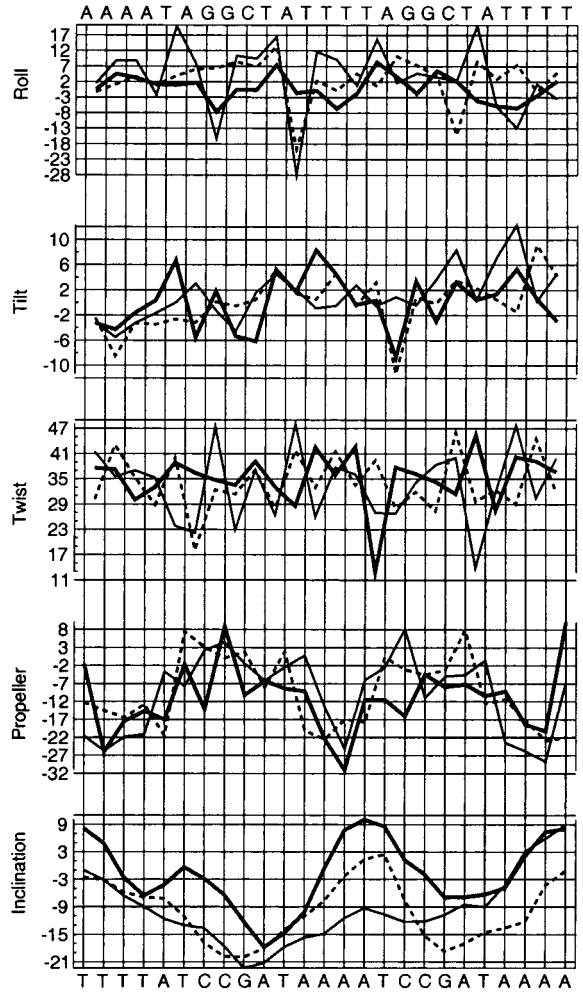

Figure 8. Sequence variations of helicoidal parameters in the bent structures shown in Figure 7. The sequence of the first strand is shown on the top in the 5′−3′ direction. The complementary sequence of the second strand is written on the bottom in the opposite direction. All parameters were evaluated with the Curves program[47] and are given in degrees. TJA, thick line; TJBa, thin solid line; and TJBb, dashed line.

Sometimes their phases are linked to the sequence, like for the roll in the GGCTATT heptaplet, but even in this rare example, its local values differ by as much as 25°. The inclination traces are exceptional in that they exhibit smooth quasi-sinusoidal variations, but their phases differ as well.

The junction model also poorly agrees with these results. Figure 7 shows that curved conformations are sometimes smoothly bent and sometimes kinked at the two insertions between the A-tracts. The kinks, however, are not centered at the A-tract boundaries, and they are not sharp. In the most kinked TJBb structure, A-tracts are less bent than are regions between them; that is, the bend is localized but still smooth. In Figure 8 the inclination traces oscillate smoothly, without kinks or jumps, always decreasing from 5′ to 3′ ends of A-tracts. All parameters vary along A-tracts and do not repeat at consecutive steps. Thus, although the structures are bent, the specific regular A-tract and non-A-tract conformations suggested by the junction model are not observed.

Nonlocal mechanisms of bending agree with our results only to a certain extent. The modified junction theory[54] assumes that the junction deformations can propagate for several base pair steps. The A-tracts in the sequence studied here may be too

(54) Nadeau, J. G.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2622−2626.

short for their ingenious structure to establish. The second nonlocal theory[55] suggested that the bending is caused by the modulations of the minor groove width. The double helix is usually bent toward the major groove at the minor groove widenings. Figures 2 and 3 show, however, that the profile of the minor groove was established earlier than bending and that it remained unchanged during the long period of straightening. It is understood, however, that these two nonlocal models are incomplete. Actually, they cannot be verified or disproved because the issue of the physical origin of bending is tacitly dropped. Simple geometrical considerations dictate that the grooves must be narrower at the inside edge of the bend.[56] Therefore, postulating that groove modulations cause bending or, vice versa, that it is bending that causes groove modulations does not solve the problem. Similarly, in the modified junction theory, the boundary deformations can exist without the structures and boundaries themselves, which makes it applicable to any DNA structure and suggests that these deformations are caused by some other force.

The last model, which was first mentioned in the context of the junction theory[57] but became popular only in recent years,[58] attributes the cause of bending to solvent counterions. If they bind specifically to the minor grooves of A-tracts, phosphate groups in a phased sequence appear neutralized at one side of the double helix. Repulsion of phosphates at the opposite side should bend DNA toward minor grooves of A-tracts. We mention this theory here for completeness only because the origin of bending in the minimal model is certainly different. Here counterions are treated implicitly, assuming that their effect is not sequence specific. The fact that this model produces static bends suggests that bending can occur without specific interactions with counterions. At the same time, such interactions certainly take place and they can modulate the bending.

The complicated patterns discussed above cannot easily be fit in with the usual views of the DNA structure. On one hand, all helicoidal parameters exhibit a certain degree of sequence dependence. On the other hand, they differ significantly even though the bends are similar, and moreover, every base pair step is put in an identical sequence context. This microheterogeneity is also observed in the average MD conformations, in the distributions of B_{II} and B_I conformers, and in the profiles of the minor grooves.[49] The fact that similar macroscopic bends may have dissimilar local structures argues against the common belief that they are built from small blocks with asymmetric conformations determined by internal interactions. Instead, the evident compensation of fluctuations in neighboring steps noticed in Figure 8 suggests that they are due to thermal motions that occur under some external constraints that require bending. In the next section, we analyze this possibility in detail.

The Possible Physical Origin of Spontaneous Static Bends in Double Helices. The hypothesis outlined below is based upon our computational results as well as upon analysis of well-known experimental data. Although it does not answer all unclear questions concerning DNA bending, we consider it most likely and describe it here for discussion and further investigation.

Let us first ask a simple geometric question: What is the shortest line that joins two points on a cylindrical surface? To answer it, one should first cut the cylinder parallel to its axis, unfold its surface onto a plane, join the two points by a straight line and then fold the surface back upon the cylinder. The resultant curve represents an interval of a spiral trace with a constant inclination to the cylinder axis. Now consider an ideal canonical B-DNA model in which the base pair stack forms the core of a cylinder and the sugar−phosphate backbone follows an ideal spiral trace on its surface, that is, the shortest line that joins the "surface" nitrogens N_1 and N_9. Assuming that the backbone is a stiff elastic characterized by a certain specific length, we are obliged to conclude that this model implies that the backbone is stretched and tends to reduce its length. Our next question is, What would happen if the preferred backbone length were longer than allowed by the canonical model? A simple answer is that it would try to extend by pushing bases. The extension can be accommodated without breaking the helical symmetry, notably, by increasing the helical twist. This option, however, is opposed by the loss in the stacking energy, and when it becomes difficult to extend in this way, the backbone will try to deviate from the ideal spiral trace. In this case, the two grooves on the surface of the Watson−Crick double helix can no longer have constant widths.

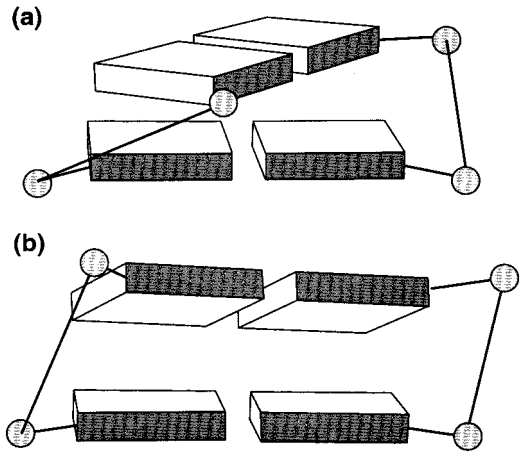

Figure 9. Simplified mechanical representation of a base pair step in B-DNA viewed toward the minor groove.[68] The fixed backbone length enforces destacking of the base pairs from the minor groove side when the twist is reduced.

It is possible that, in physiological conditions, the equilibrium specific length of the DNA backbone is actually larger than that in a regular B-DNA structure with the average helical twist of 34.5°. The backbone appears "frustrated" and is forced to break the symmetry and wander along the cylindrical surface. The parallel stacking has to be perturbed, and it is probably this effect that eventually bends the double helix. Let us consider one simple mechanism. We take an ideal B-DNA and vary only the helical twist. The model in Figure 9a makes clear that by smoothly increasing and decreasing the twist one narrows and widens the minor groove. The desired backbone waving emerges, and a larger length can be accommodated on the same cylindrical surface. This figure also indicates, however, that if the parallel stacking is maintained, the backbone must be compressed when the twist is reduced and stretched in the opposite phase. Because the backbone is stiff and it cannot be compressed significantly, the parallel stacking should suffer. This is illustrated schematically in Figure 9b. In the widenings of the minor groove, the backbone pushes off the neighboring base pairs at C1′ atoms, enforcing destacking from the minor groove side and bending DNA toward the major groove.

The above hypothesis is supported by several well-known experimental facts. Modulations of the DNA grooves is an immediate indicator of the compressed state of the backbone.

(55) Burkhoff, A. M.; Tullius, T. D. *Cell* **1987**, *48*, 935−943.
(56) Drew, H. R.; Travers, A. A. *J. Mol. Biol.* **1985**, *186*, 773−790.
(57) Levene, S. D.; Wu, H.-M.; Crothers, D. M. *Biochemistry* **1986**, *25*, 3988−3995.
(58) Hud, N. V.; Sklenář, V.; Feigon, J. *J. Mol. Biol.* **1999**, *286*, 651−660.

This feature is ubiquitous in X-ray B-DNA structures regardless of the sequence. If we had to decide whether the backbone in stretched, relaxed, or compressed by looking only at X-ray B-DNA oligomers, we would have to conclude that the first option is unlikely, the second is possible, and the third is the most probable. The backbone is a charged polymer; therefore, its equilibrium specific length is maximal with low temperature, and it should decrease as the temperature grows. According to our model, this should cause reduction of intrinsic bending regardless of the sequence, which agrees with experiments.[8] The state of the backbone is affected by water activity. It is known that the transition from the B- to A-DNA form can be induced by different solvent additives that reduce the specific backbone length.[59] Any such factor should also reduce the bending; therefore, one may expect that solvent conditions that shift the equilibrium toward the A-form should abolish the bending first. Actually, regardless of the sequence, the bending is suppressed by small concentrations of monovalent counterions and alcohols.[8,9,60,61] It can be noted, finally, that the effect shown in Figure 9b is the well-known mechanism responsible for twist−roll correlations which are common in experimental B-DNA structures.[62]

The twist−roll model in Figure 9b illustrates only one of the ways the stacking can be perturbed by the compressed backbone. It is clear that ensembles of such perturbations may be very heterogeneous. Actually, the overall backbone length does not completely determine its conformation. Moreover, any given backbone profile may be compatible with a number of base pair orientations. The whole canonical ensemble can also include straight conformations in which all asymmetrical perturbations compensate one another. Therefore, groove modulations may occur without bending, which agrees with the recent experimental data[63] and was exemplified here in TJBa.

This view of the DNA molecule assumes that groove modulations and, accordingly, stacking perturbations are its inherent features that always exist regardless of the sequence. However, the macroscopic curvature emerges only if phases of these modulations are coherent with the helical screw. If, in addition, groove widenings and narrowings cannot migrate along the molecule, the bending direction is fixed, as well. Sequences with A-tract repeats fulfill these conditions simply because A-tracts prefer higher twist[15−17] and naturally accommodate narrowings of the minor groove, while the common intervening sequences prefer lower twist and fix the widenings. Therefore, the static bending should be possible in repeats of any other sequences with alternating properties.[30] A special A-tract form of DNA is not necessary, and moreover, the regular poly-dA structure may not exist in solution. Even with the average twist of 36°, the backbone may still be compressed and continue to wander, although with a longer characteristic wavelength.

If this theory is true, it would entail important biological as well as physical conclusions. A homopolymer DNA with a waving backbone exemplifies a system with symmetry that is broken due to an intrinsic frustration.[64] It may have very interesting physical properties, notably, never-ending migration of the backbone waves along DNA and glass transitions. The theory also predicts that the DNA structure is not completely determined by the stacking interactions in dinucleotides, trinucleotide, and so forth. The backbone waves may put every base pair step in a medium range context because their phases can correlate over several helical turns. This makes possible mutual dependence of local conformations in sites separated by considerable DNA stretches and, consequently, an allosteric regulation. The degree of backbone compression is connected with supercoiling, and this theory may shed additional light upon the mechanism by which topoisomerases control activity of gens.[65] One may note also that we have proposed here a unified model which explains intrinsic DNA curvature as well as bending induced by the negative supercoiling in circular plasmids.

Conclusions

The static curvature spontaneously emerges in free MD simulations of an atom-level molecular model of a B-DNA double helix, with the nucleotide sequence as a single structural bias. Convergence to similar statically bent states have been demonstrated in three independent MD trajectories of 10−20 ns. The bending direction and its magnitude are in good agreement with experimental observations. Unexpectedly, the computed similarly bent structures exhibit striking microscopic heterogeneity concerning variations of helicoidal and conformational parameters along the molecule. On the basis of the computational results, as well as the literature experimental data, a new possible mechanism of bending in the double helical DNA is proposed. It postulates that in physiological conditions, the equilibrium specific length of the DNA backbone is larger than that admissible in the regular B-DNA form, which forces it to fold in a wavy trace on the cylindrical surface of the double helix. This results in modulations of the minor groove width, slight asymmetrical destacking of base pairs and, eventually, bending of the DNA molecule.

Supporting Information Available: Illustration of the bending dynamics in the three trajectories and comparison of the last 1-ns average structures. This material is available free of charge via the Internet at http://pubs.acs.org.

JA001605Y

(59) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(60) Marini, J. C.; Effron, P. N.; Goodman, T. C.; Singleton, C. K.; Wells, R. D.; Wartell, R. M.; Englund, P. T. *J. Biol. Chem.* **1984**, *259*, 8974−8979.
(61) Zacharias, D. S. W.; Wood, Z. A.; Harvey, S. C. *Nucleic Acids Res.* **1995**, *23*, 1816−1821.
(62) Yanagi, K.; Privé, G. G.; Dickerson, R. E. *J. Mol. Biol.* **1991**, *217*, 201−214.
(63) Møllegaard, N. E.; Bailly, C.; Waring, M. J.; Nielsen, P. E. *Nucleic Acids Res.* **1997**, *25*, 3497−3502.
(64) Zakrzewska, K.; Poltev, V. I.; Oguey, C.; Lavery, R. *J. Mol. Struct.* **1993**, *286*, 219−229.
(65) Kornberg, A.; Baker, T. A. *DNA Replication*; Freeman: New York, 1992.
(66) Dickerson, R. E.; Bansal, M.; Calladine, C. R.; Diekmann, S.; Hunter, W. N.; Kennard, O.; Lavery, R.; Nelson, H. C. M.; Olson, W. K.; Saenger, W.; Shakked, Z.; Sklenar, H.; Soumpasis, D. M.; Tung, C.-S.; von Kitzing, E.; Wang, A. H.-J.; Zhurkin, V. B. *J. Mol. Biol.* **1989**, *205*, 787−791.
(67) Mazur, A. K. *J. Mol. Biol.* **1999**, *290*, 373−377.
(68) Calladine, C. R.; Drew, H. R. *Understanding DNA: The Molecule & How it Works*; Academic Press: London, 1992.