JACS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Subscriber access provided by ISTANBUL TEKNIK UNIV

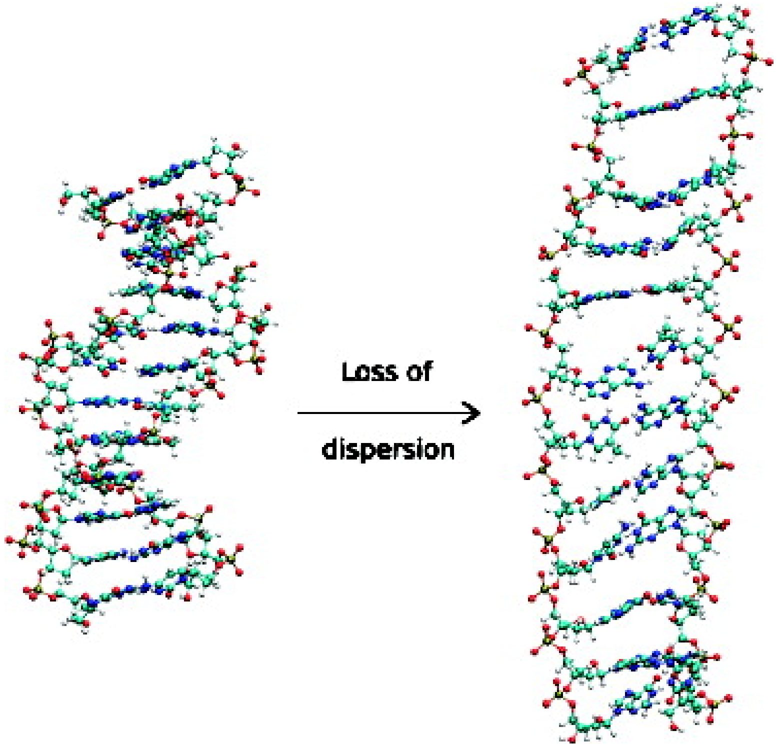
Article

Double-Helical # Ladder Structural Transition in the B-DNA is Induced by a Loss of Dispersion Energy

Jir#i# C#erny#, Martin Kabela#c#, and Pavel Hobza

J. Am. Chem. Soc., **2008**, 130 (47), 16055-16059 • DOI: 10.1021/ja805428q • Publication Date (Web): 01 November 2008

Downloaded from http://pubs.acs.org on February 8, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

ACS Publications High quality. High impact.

Journal of the American Chemical Society is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Subscriber access provided by ISTANBUL TEKNIK UNIV

View the Full Text HTML

Journal of the American Chemical Society is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

J|A|C|S
ARTICLES

Published on Web 11/01/2008

Double-Helical → Ladder Structural Transition in the B-DNA is Induced by a Loss of Dispersion Energy

Jiří Černý, Martin Kabeláč, and Pavel Hobza*,†

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and Center for Biomolecules and Complex Systems, 166 10 Prague 6, Czech Republic

Received July 14, 2008; E-mail: pavel.hobza@uochb.cas.cz

Abstract: The role of the dispersion energy and electrostatic energy on the geometry and stability of the B-DNA helix was investigated. Both molecular dynamics simulations with empirical force field and hybrid quantum mechanical/molecular mechanics molecular dynamics simulations, where the dispersion or electrostatics term is suppressed/increased, on the one hand and an ab initio minimization procedure on the other have shown that the lack of the dispersion term leads to an increase of the vertical separation of the bases as well as to a loss of helicity, thus resulting in a ladder-like structure. A decrease of the electrostatic term produces a separation of the DNA strands. The biological consequences of both electrostatic and dispersion forces in DNA are enormous, and without either of them, DNA would become unstable and unable to provide the storage and transfer of genetic information.

Introduction

Being responsible for the structure of biomacromolecules, and indirectly also for their function, noncovalent interactions play a unique role in biology. The double-helical structure of DNA, responsible for storing and transferring genetic information, is determined by a subtle balance of noncovalent interactions among the DNA building blocks. The most prominent role is played by the interactions between the DNA bases, where two binding motifs can be recognized: planar hydrogen bonding and vertical stacking. In DNA, unlike in RNA, two exclusive types of H-bonded arrangement exist, that is, the Watson−Crick (WC) arrangement of guanine (G) and cytosine (C) and the WC arrangement of adenine (A) and thymine (T), whereas there are 10 different types of stacked arrangements of the four bases. It was believed that H-bonding was more important than stacking and the stability of DNA was thus governed by the presence of H-bonds. The role of stacking had not been known for a long time, which was partially caused by the fact that the absolute as well as relative stability of both binding motifs in the crystal geometry can be determined experimentally only with great difficulty. Furthermore, a theoretical description of the stacking of DNA bases, where London dispersion energy plays a key role, is more complicated than that for H-bonding. The stacking was underestimated by the first correlated ab initio quantum chemical calculations[1,2] (MP2 with small and medium basis sets), performed in the 1990s, and definitive information about the relative strength of both binding motifs was only provided by recent, highly accurate calculations performed at the CCSD(T)/complete basis set (CBS) level.[3] The gas-phase optimized and DNA crystal structures look similar for the H-bonded GC WC and AT WC structures, and their CCSD(T)/CBS stabilization energies amount to about 29 and 15 kcal/mol, respectively.[4] A description of the stacked systems is more complicated, because the gas-phase optimized and crystal structures differ; the stabilization energies for GC stacked pairs in both geometries are nevertheless very large and reach approximately 17 and 14 kcal/mol, respectively. The stacked crystal structures of other combinations of bases are less stable but still reach high values of ca. 10 kcal/mol.[4] An important conclusion can be drawn immediately—the stacking of G and C (without any H-bond and governed mainly by London dispersion energy) is comparably stable to the H-bonding of A and T having two strong H-bonds.

The transfer of genetic information is triggered by the opening (unwinding) of DNA followed by a synthesis of two daughter strands. The unwinding of DNA is proportional to the stability of the DNA, and several empirical models have been suggested to deduce the stability from the sequence of the DNA bases.[5] We have shown recently that unwinding free energy can be correlated with the H-bonded and stacked interaction energies, and since the process occurs in a water environment, solvation free energy should also be considered.[6] The main (and surprising) conclusion of the study was that H-bonding contributes less to the stability of DNA than stacking, because H-bonding (unlike stacking) is penalized by a large desolvation energy.

The role of stacking can be proven in two completely independently ways—namely by performing geometry optimization or running molecular dynamics simulations with a modified energy function. As has been shown, the stacking energy is governed mainly by London dispersion energy (see also refs 4 and 7). When the dispersion energy term is reduced to zero (or conversely when it is enlarged), the structure of DNA should

---

[†] Second address: Department of Physical Chemistry, Palacký University, Olomouc, tř. Svobody 26, 771 46, Olomouc, Czech Republic.

(1) Aida, M.; Nagata, C. *Int. J. Quantum Chem.* **1986**, *29*, 1253–1261.
(2) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590–5596.
(3) Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2003**, *125*, 15608–15613.
(4) Sedlák, R.; Jurečka, P.; Hobza, P. *J. Chem. Phys.* **2007**, *127*, 075104.
(5) Doktycz, M. J.; Morris, M. D.; Dormady, S. J.; Beattie, K. L. *J. Biol. Chem.* **1995**, *270*, 8439–8445.
(6) Řezáč, J.; Hobza, P. *Chem.−Eur. J.* **2007**, *13*, 2983–2989.

10.1021/ja805428q CCC: $40.75 © 2008 American Chemical Society

be affected. A similar experiment can be carried out with H-bonding, which is stabilized mainly by electrostatic interactions. Once these interactions are reduced, the strength of H-bonding is also suppressed.

In the paper being presented, the stability of DNA in the water environment is investigated, with the study being performed along two directions. First, the quantum mechanical (QM) optimization of a small fragment of DNA in the water environment (and, for the sake of comparison, also in a vacuum) was performed. A similar study has already been published,[8] but the information that it provides is limited. It is, however, evident than when dispersion energy is not covered, the structure of DNA is modified. The authors[8] mentioned that "it is not surprising that fragments of DNA are structurally unstable when computed with self-consistent charges density-fitting tight binding (SCC-DFTB)" (i.e., with a method that does not cover dispersion energy). We are aware of the weakness of this approach, as optimizations cannot describe the dynamic character of DNA at real temperature even if performed in water. Second, we hence performed various molecular dynamics (MD) simulations in explicit water based either fully on the empirical potential or on more accurate (and thus also more reliable) QM/MM MD simulations. Each type of MD simulations was performed first with standard (unmodified) energy functions and, subsequently, with modified energy functions. In the case of dispersion energy, its weakening as well as strengthening was considered while in the case of electrostatic energy only weakening was taken into account. In all the steps, the structural changes were detected and compared with the data of the unmodified empirical potential and the quantum mechanical/molecular mechanics simulations (QM/MM).

Methods

The QM optimization was performed for the central, tetrameric part of a crystal structure of a B-DNA 12-mer (PDB code 1BNA).[9] To compensate for the negative charge of the phosphate groups, we neutralized them with six sodium cations. To mimic the presence of a solvent, the implicit solvent COSMO model[10] was employed. Such a size of the system is still computationally feasible at the DFT level of theory employing a medium size basis set provided that the resolution of identity approximation (RI-) is used. The system is large enough to avoid the terminal effects and thus able to describe at least the base pairs situated in the center of the tetramer correctly. The current theoretical procedure can easily be used for studying the role of dispersion energy, because the empirical dispersion energy[11] is added to the DFT energy, which does not cover dispersion.[12,13] The optimizations were performed with dispersion energy either fully taken into account or not considered at all. During the fully unconstrained minimization procedure, the TPSS functional[14] and the SV(P) basis set were used. The same functional but with a larger TZVP basis set was employed for the calculation of the interaction energy of two strands. All of the calculations were performed using the TurboMole 5.8 suite of programs.[15]

The entire crystal structure of the above-mentioned B-DNA 12-mer with the sequence: 5′-CGCGAATTCGCG-3′ was taken as an initial geometry for our MD simulations. The molecule was oriented along the principal axes and placed into the cubic box with a distance of 10 Å from the molecule to the box edges. The box contained 3373 TIP3P water molecules and 22 sodium cations so that system electroneutrality would be achieved. When the ϵ parameter was scaled to 0.01 (see below), the DNA molecule extended significantly and started approaching the edge of the box. We hence performed another simulation, elongating the box in the direction of the long DNA axis. The resulting box contained 6273 water molecules.

The molecular dynamics simulations were performed with an AMBER parm99[16] empirical force field and the following modifications were introduced in the nonbonded part, which describes the potential energy of the system (see eq 1) and is divided into the electrostatic and Lennard-Jones terms. The former term is modeled by the Coulomb interaction of atomic point-charges, whereas the latter describes the dispersion and repulsion energies.

$$V(r) = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} + 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \qquad (1)$$

Modifications of the dispersion energy were introduced by scaling the parameter ϵ. While we are aware that scaling this parameter affects both dispersion and repulsion parts of the potential, no other simple way to modify only the dispersion energy term exists. We found, however, that smaller repulsion (a byproduct of dispersion energy reduction), for which the r^{-12} scaling is selected rather arbitrarily to accelerate the calculation, does not introduce any problems during the simulations. The following values of the scaling factor for ϵ were considered: 0.01, 0.5, 1.0 and 2.0, the first of which almost completely removes the dispersion energy, the second yields its 50% reduction and the last magnifies the dispersion energy by a factor of 2. All the modifications were performed for all the atoms of the DNA, maintaining the original parameters for the water molecules (Any modification of the water parameters resulted in significant artifacts in the solvent structure.).

The modification of electrostatic term was performed via changes of the atomic charges directly involved in the X-H$\cdots$Y hydrogen bonds between the nucleic acid bases. Simulations were performed with these charges scaled by 0.1, 0.5 and 1.0: In the first case the H-bonding was practically suppressed while in the second case it was reduced to 25%. Minor modifications of the sodium counterion charge were introduced (by 0.0017 and 0.0031 electrons, respectively) to maintain the overall neutrality of the entire system. None of the other charges in the DNA and waters were changed. The Gromacs 3.3 MD package[17] was used for all the simulations.

The above-mentioned problems with the scaling of the energy terms led us to use more accurate QM/MM MD simulations, where the QM calculations were performed with the self-consistent charge density-functional tight-binding (SCC-DFTB) method empirically augmented with dispersion energy.[18] The MM part was described by the Amber forcefield parm99. Using the QM/MM method for the presently employed 12-mer, where all the base pairs are taken in the QM part, is impractical, and thus a smaller model was adopted. Only the four central AT base pairs were described quantum mechanically by the SCC-DFTB-D procedure, whereas

---

(7) Hesselmann, A.; Jansen, G.; Schütz, M. *J. Am. Chem. Soc.* **2006**, *128*, 11730–11731.

(8) Elstner, M.; Frauenheim, T.; Suhai, S. *J. Mol. Struct.* **2003**, *632*, 29–41.

(9) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179–2183.

(10) Klamt, A.; Schuurmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

(11) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555–569.

(12) Hobza, P.; Šponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315–1325.

(13) Kristyan, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175–180.

(14) Tao, J. M.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(15) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.

(16) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–1074.

(17) Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(18) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149–5155.

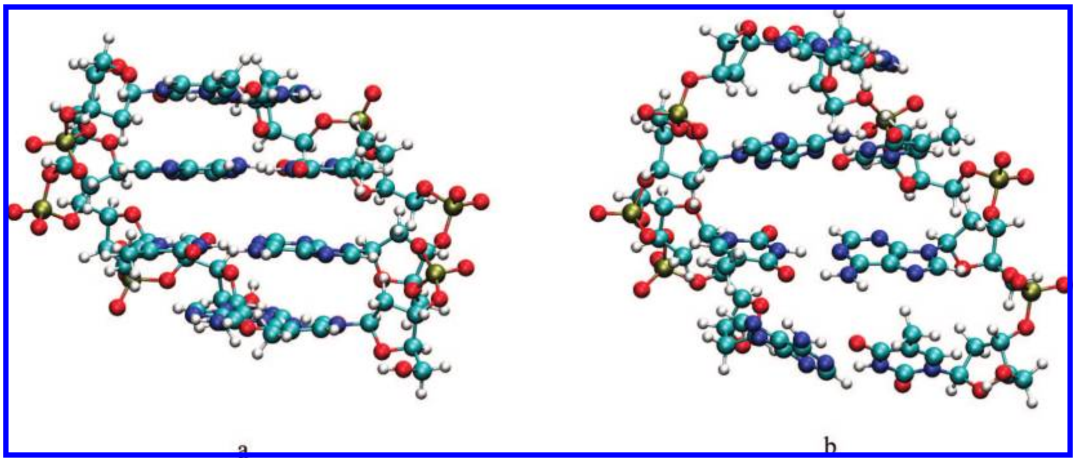


Figure 1. Optimized structures of the 5′-TATA-3′ tetramer obtained using (a) the RI-DFT-D method including the empirical dispersion term and (b) the RI-DFT method (i.e., with the dispersion energy not being taken into account).

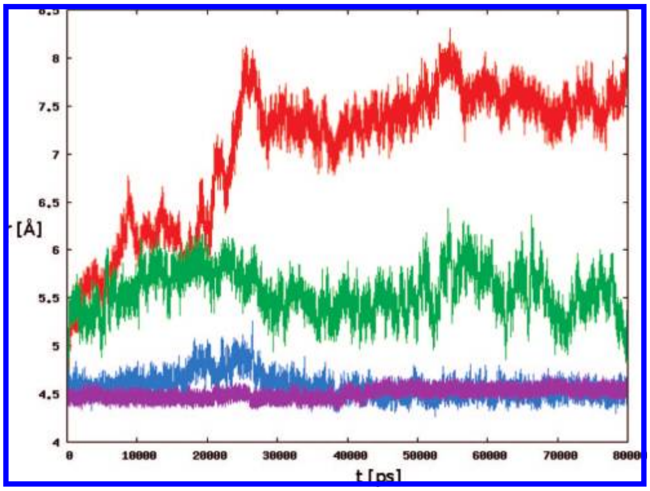


Figure 2. Mean distances (Å) between the nitrogen atoms (connected to the sugar moiety) of the subsequent NA bases along the DNA strands. The results involving the ϵ parameter scaled by 0.01, 0.5, 1.0 and 2.0 (red, green, blue and violet lines, respectively) are presented for the entire 80 ns force field simulation.

their sugar−phosphate backbones and all the water molecules, as well as the rest of the DNA and ions were treated by an unmodified empirical force field. The standard AMBER embedding technique was adopted.[19] For the QM/MM calculations, the AMBER10 package was employed.[19]

There is an important difference between the QM/MM and the previously mentioned fully empirical MD approaches. The QM/MM procedure makes it possible simply to disable the dispersion correction without affecting the repulsion part of the energy function.

All the simulations were performed under constant temperature and pressure conditions, beginning with the parameters obtained after a 1 ns equilibration run. The parmBSC0 correction[20] of the force field (important for the long simulation times) was also included. The length of the MD simulations was 80 ns. For the (much more time-consuming) QM/MM simulations, the trajectory length was approximately 0.9 ns.

Changes of the DNA structure were analyzed in terms of the mean distances between the nitrogen atoms connected to the sugar moiety in all the stacked base pairs in the DNA investigated and the mean distances between the nitrogens in the N−H···N hydrogen bonds in the GC WC and AT WC pairs. Moreover standard parameters defining the DNA structure (rise and twist) were also utilized. The latter parameters were obtained by averaging the values obtained from analysis of the final 5 ns of the trajectories by the 3DNA package of programs.[21]

Results and Discussion

QM Optimizations. The minimization of the 5′-TATA-3′ tetramer in the RI-DFT-D[11] method in the presence of implicit solvent led to a structure which differs only slightly from the initial one (cf. Figure 1a). On the other hand, when optimization was performed using the RI-DFT method (i.e., with the dispersion energy not being taken into account), it yielded a significantly distorted structure (cf. Figure 1b), which exhibits a huge enlargement of the vertical distance of the bases. Furthermore, the central bases are enormously twisted, which is a result of their effort to create unnatural H-bonded contacts with the peripheral bases. This certainly originates from the compensation of the missing dispersion forces stabilization by electrostatic forces. When the dispersion is included, the total calculated stabilization energy between the two strands in the complex of optimized tetramer is equal to 71 kcal/mol, 22 kcal/mol of which were gained from the dispersion and the remaining 49 kcal/mol have their origin in the other terms. The net stabilization energy of the complex, where dispersion is not covered, is only 52 kcal/mol, which is however 3 kcal/mol more than for the nondispersive contribution to the total stability of the first complex.

MD Simulations. Scaling of Dispersion Energy. By varying the parameters which influence the dispersion energy, mainly the stacking interactions are modified, with H-bonding being affected considerably less (see below). Figure 2 shows the mean distances between the nitrogen atoms connected to the sugar moiety in all the stacked base pairs in the DNA investigated. The blue curve corresponds to the simulation with the standard (unmodified) potential and the mean distance between the nitrogens is ca. 4.5 Å. This distance remains practically unchanged during the simulation. Only between 20 and 25 ns of simulation time is apparent a small enlargement of this distance. By reducing the dispersion energy to 50% (represented by the green curve), we obtained a less folded double helical structure, where the distance between the nitrogens increased

---

(19) Case, D. A.; et al. AMBER10; University of California: San Francisco, 2008.

(20) Perez, A.; Marchan, I.; Svozil, D.; Šponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817–3829.

(21) Lu, X. J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5108–5121.

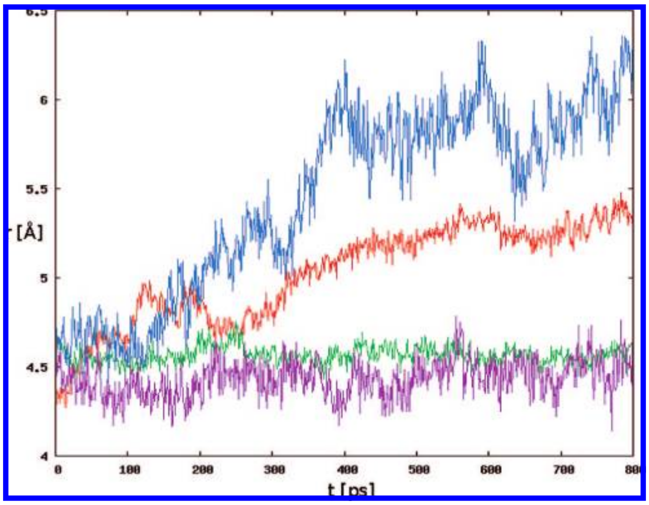


Figure 3. Mean distance (in Å) between the nitrogen atoms (connected to the sugar moiety) of the subsequent NA bases along the DNA strands. A comparison of the QM/MM method with (violet line) and without (blue line) the dispersion correction term and the force field results for ϵ scaled by 0.01 (red line) and 1.0 (green line) is shown. Only the common parts of the simulations (up to 0.8 ns) are presented for clarity.

by approximately 1 Å. A reduction of the dispersion energy to almost zero (shown by the red curve) leads to a rather dramatic change of the double helical structure. The folded structure is almost fully transferred to the ladder-like structure, characterized by a considerably larger N•••N distance (of roughly 7.5 Å). It is important to mention that the distance between the H-bonded pairs remains practically constant throughout the simulation. A very fast unfolding is further evident from Figures 2 and 3. During the first 1000 ps, the N•••N distance increased by approximately 1 Å, and further increase is apparent in the next few ns. The final N•••N distance of ca. 7.5 Å is attained within 20 ns (when the water molecules are entering the space between the stacked bases). When, on the other hand, the strength of the dispersion is increased (by a factor of 2, the violet curve in Figure 2) the folded structure is more rigid, but the average distance differs only negligibly from that obtained from the simulation with the original parameters. Qualitatively very similar results can be obtained when for example the distance of the center of mass of the respective DNA bases or the distance between the C_1 carbon atoms of the deoxyribose moieties is considered. Figure 4 shows a final ladder-like (unfolded) structure, which is practically two-dimensional, and, for the sake of comparison, also the initial folded structure, with the fundamental difference between them being clear.

The QM/MM MD simulations (see Figure 3) fully confirmed the picture discussed above. The blue and violet curves in Figure 3 correspond to QM/MM simulations, where the dispersion energy is either completely missing or fully covered. It should be mentioned here that in this case only the dispersion energy was modified while the other energy terms remained untouched. The blue curve shows the very fast unfolding (with the N•••N distance increased by ca. 1.5 Å within 0.5 ns) while the violet curve suggests a rather rigid folded structure. When comparing the corresponding QM/MM and MM simulations, we found very good agreement. The violet (QM/MM) and green lines (MM) correspond to the simulations with full (nonscaled) dispersion, whereas the blue (QM/MM) and red lines (MM) come from the simulations where the dispersion energy was disabled.

Behind the above shown structure analysis based on the average N•••N distance we also evaluated the standard structural parameters. The values of rise and twist for folded DNA (cf. Figure 4a) amount to 3.4 Å and 31.5°, while upon scaling of the dispersion energy these values change to 5.1 Å and 2.4°,

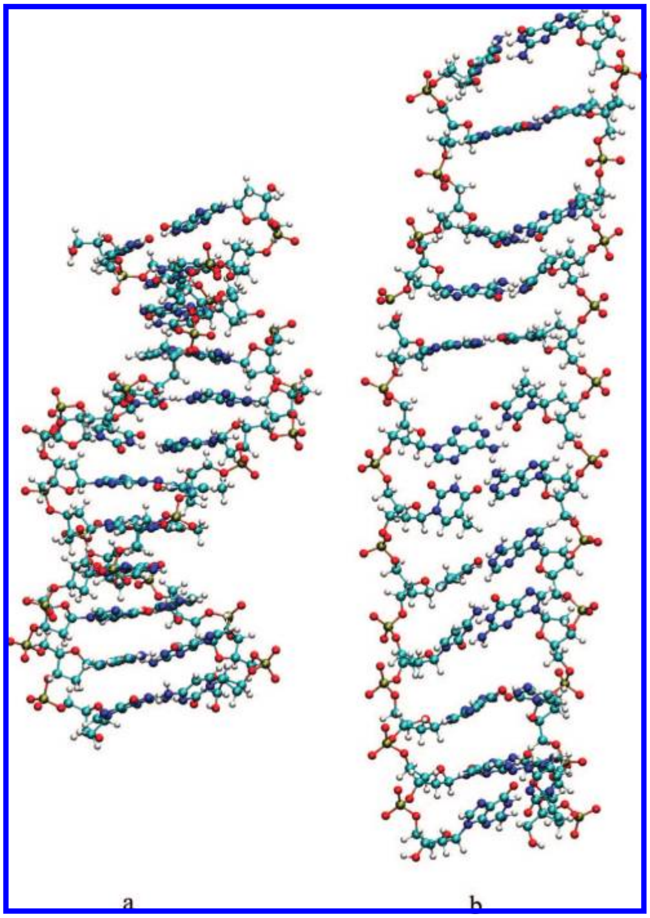


Figure 4. Snapshot figure of (a) the initial crystal structure and (b) the final ladder-like structure of the DNA corresponding to the force field simulation involving the ϵ parameter scaled by 0.01.

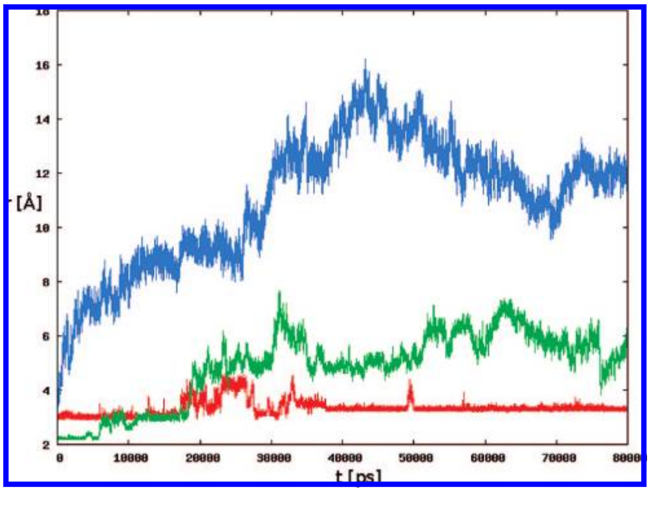


Figure 5. Mean distance (Å) between the nitrogens in the N−H•••N hydrogen bonds in GC WC and AT WC base pairs across the DNA strands. The results for the unmodified empirical force field parameters (ϵ =1) (red line), for the simulations where the dispersion term ϵ is scaled by a factor of 0.01 (green line) and for the simulation involving the charges scaled by 0.1 (blue line) are shown. The fluctuations in the distance for the unmodified force field found around 20 ns represent a temporary breaking of the hydrogen bonds of the central AT pairs.

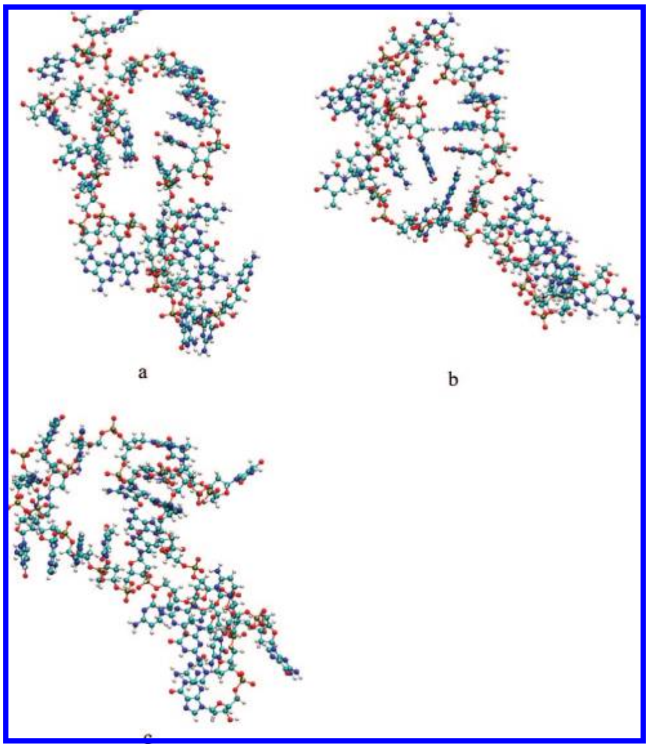


Figure 6. Snapshots of DNA structures obtained during the simulation involving the charges scaled by 0.1. The snapshots were selected to represent the evolution of the structures at approximately 20 (a), 40 (b), and 60 ns (c), respectively.

which clearly corresponds to the completely unfolded structure (cf. Figure 4b).

Scaling of the Electrostatic Energy. The electrostatic energy modification affects mainly the planar H-bonding. The red curve in Figure 5 indicates the mean distance between the nitrogens in the N−H···N hydrogen bonds in the GC WC and AT WC pairs obtained from simulations with the standard (unmodified) empirical potential. The distance remains almost constant in the 80 ns simulation and oscillates around 3.5 Å. When the atomic charges are reduced to 10% (represented by the blue line), the distance between two nitrogens rapidly increases up to 12 Å. This provides evidence of a complete breaking of all the H-bonds, yielding two separate strands. Since stacking is present in each strand, some degree of folding (helicity) is introduced. Despite the fact that H-bonding between complementary pairs no longer exists, both strands attract each other and give rise to various forms of superstructures (fundamentally different from the regular DNA). This is also due to the fact that the dispersion energy contributes non-negligibly to H-bonding as well. Figure 6 contains a few typical snapshots of the structures obtained during MD simulation.

Conclusion

We conclude that both electrostatic and dispersion forces are unambiguously equally essential for the structure and full biological functionality of DNA. The role of the electrostatic force is obvious—the lack of Coulomb attraction would lead to the separation of the two strands with a partial conservation of the helical structure in each strand. The importance of the dispersion forces for stability has been rather unclear until now. Our simulations show that the lack of dispersion leads quickly to the complete unwinding of the strands from the folded double-helical structure to a practically two-dimensional ladder-like arrangement. The vertical distance of the DNA bases (rise) increased from 3.4 Å to about 5.1 Å. Such drastic geometrical changes would have fundamental biological consequences—large base separation in one strand would lead to a loss of replication and transcription activity, proteins like transcription factors would not be able to bind to specific sequences of the bases exposed in the major groove, water can easily penetrate between the bases, etc. Each of these above-mentioned facts would surely be lethal for the cell/living organism.

Acknowledgment. This work was a part of the research project No. Z40550506 of the Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and it was supported by Grants No. LC512 and MSM6198959216 from the Ministry of Education, Youth and Sports of the Czech Republic and No. IAA400550808 (M.K.) from the Czech Science Foundation. The support of Praemium Academiae, Academy of Sciences of the Czech Republic, awarded to P.H. in 2007 is also acknowledged.

Supporting Information Available: Complete ref 19. This material is available free of charge via the Internet at http://pubs.acs.org.

JA805428Q