DNA Unfolding

DOI: 10.1002/anie.201000593

Real-Time Atomistic Description of DNA Unfolding**

*Alberto Perez and Modesto Orozco**

Despite recent efforts, the folding/unfolding of DNA (understood as a dramatic conformational change from the native conformation in a significantly large portion of the duplex) is still poorly described. A range of classical physical studies led to the assumption that the unfolding of short DNA fragments is reversible and follows a two-state mechanism, starting at d(A·T) pairs.[1–3] Nevertheless, this traditional view has been recently challenged[4–9] by ultrafast techniques, which suggested a more complex scenario where compact intermediates are detected during the microsecond-long unfolding process.[4] Unfortunately, these experiments were not able to provide atomistic-detailed information on the process, thus making necessary the use of simulation techniques (mainly molecular dynamics, MD) as complementary tools. For computational reasons MD simulations of unfolding have typically followed indirect approaches, such as the use of multiple short trajectories,[10,11] replica exchange,[12,13] or moderately large (100 ns) simulations under nonphysical denaturing conditions ($T = 400$ K).[14,15] These simulations provided clear evidence of the complexity of DNA unfolding but were unable to define a mechanistic view, which would require multiple very large unbiased trajectories. Herein, we present a full atomistic description[16] of the unfolding of a full turn of DNA under realistic denaturing conditions. The study, a real "tour de force" for MD, provides for the first time a detailed atomistic picture of DNA unfolding in the microsecond timescale.

All simulations were performed using Dickerson's dodecamer[17] (DDD; Protein Data Bank (PDB) code 1BNA), a well-studied model of a short nonhairpin duplex. To guarantee unfolding on the microsecond timescale,[10,11] we simulated strong (but realistic) denaturing conditions by adding a high concentration of a chemical denaturant (pyridine; Pyr) and increasing the temperature to nearly the boiling point of water (these simulations are coded as PHT). Control simulations were performed that considered: 1) water at low temperature (WLT), where we expected no signal of unfolding on the microsecond timescale; 2) a mixture of water and Pyr at low temperature (PLT), where we expected the beginning of unfolding; and 3) simulations in water at high temperature (WHT), where we expected partial unfolding. WLT and PLT microsecond simulations were performed at $T = 300$ K considering pure water (5000 water molecules) and a 1:1 (by mass) Pyr/water mixture (3923 water and 760 Pyr molecules), respectively. Simulations in hot solvent (WHT and PHT) were performed at $T = 368$ K using the same simulation boxes as before (for simulation details, see the Supporting Information). To gain better statistics, ten replicas were performed under PHT conditions for at least 1 μs. Simulations were performed using state-of-the-art simulation conditions (see the Supporting Information), the parmbsc0/parm99[18] force field for DNA, TIP3P[16] for water, and a new force field for Pyr derived from quantum mechanics calculations and condensed-phase simulations (see the Supporting Information for details).

As expected,[16,19] WLT trajectories sample conformations of DNA close to the experimental structure. The addition of large amounts of Pyr (1:1 ratio; simulation PLT) does not lead to complete unfolding of the DNA in 1 μs (Figure S1, Supporting Information), but the increase in the mobility of the duplex, and the lack of convergence in the trajectory, seems evident even from global properties such as the root-mean-square deviation (RMSD; Figure S1, Supporting Information), thus illustrating the denaturing properties of Pyr. Many irreversible unfolding microevents are detected on the submicrosecond timescale, related in most cases to the disruption of the two ends, which generates conformations where the opened nucleobases move along the major groove towards the center of the duplex in a "fraying–peeling"-like movement (Figure S2, Supporting Information). Looking at the trajectories it is clear that it is only a question of time before the "fraying–peeling" process detected here will be repeated again, thus pushing the structure to the unfolded state.

The heating of the DNA/water system to 368 K (WHT) leads to a large increase in the mobility of the entire duplex. Some distorted conformations are quite distant to the crystal (RMSD (central 10-mer) > 6 Å; see Figure S1, Supporting Information), thus indicating the existence of several unfolding microevents, which are assigned to the disruption of up to three base steps (following the fraying–peeling mechanism noted above) in one of the ends of the duplex (two base pairs lost in the central 10-mer; see Figure S2, Supporting Information). These unfolding microevents are common and appear as soon as 50 ns after equilibration, but they do not progress towards DNA unfolding since they are reversible on the nanosecond timescale. Cluster analysis[20] of 10^6 snapshots of the central 10-mer duplex reveals that a full canonical duplex (32%) and a duplex with nine canonical base pairs

[*] A. Perez, M. Orozco
Joint IRB–BSC Research Program in Computational Biology
Institute for Research in Biomedicine (IRB)
Josep Samitier 1–5, Barcelona 08028 (Spain)
and
Barcelona Supercomputing Centre (BSC)
Jordi Girona 29, Barcelona 08034 (Spain)
E-mail: modesto@mmb.pcb.ub.es
Homepage: http://mmb.pcb.ub.es

[**] We are indebted to Prof. Luque for helpful comments. This work was supported by the Spanish Ministry of Science (BIO2009-10964, Combiomed ISCIII, Consolider E-Science) and Fundacion Marcelino Botín.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201000593.

(64%) are the dominant conformers. Again, it is clear that these structural distortions should lead to the unfolding of the duplex, but this is not observed on the microsecond timescale.

Ten microsecond-long simulations in Pyr/water at 368 K (PHT1–10) were performed to determine the unfolding routes under strong denaturing conditions (see the Supporting Information). Unfolding was detected in six out of the ten simulations (see Figure 1), whereas only partial unfolding was

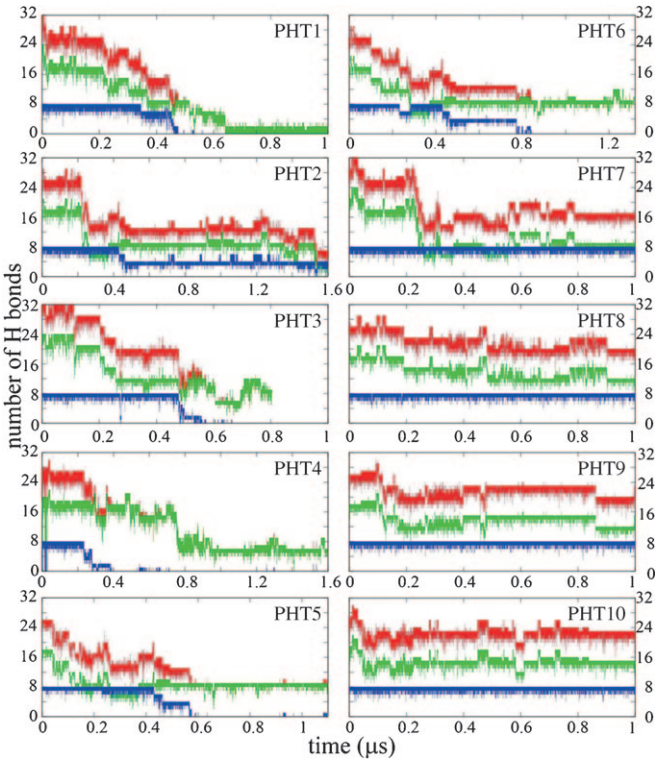

Figure 1. Time evolution of the total number of canonical hydrogen bonds for the 12-mer in ten independent PHT simulations. Red: canonical hydrogen bonds; green: hydrogen bonds resulting from C·G pairings; blue: hydrogen bonds resulting from A·T pairings.

observed in the remaining four (see Figure S1 and S3, Supporting Information), which demonstrates the stochastic character of unfolding. Interestingly, the first unfolding microevents in PHT happen quite late, around 200–400 ns after equilibration (that is, later than in the WHT simulation; see Figure S1, Supporting Information). However, contrary to the situation in water these events are now irreversible, thus leading to trajectories ending with disruption of the helix in six of the ten cases in 0.5–0.8 μs.

Analysis of the number of native hydrogen bonds in the central 10-mer duplex reveals the presence of transient states corresponding to partially unfolded structures with quite long residence times (100 ns or more; see Figure 1). Five of the six unfolding trajectories follow a common "zipper unfolding model" involving opening from the ends. Cluster analysis shows that (near) canonical B-DNA forms are detected (state 1; Figure 2) for 200–400 ns, then evolving towards more opened forms (state 2; Figure 2) because of the disruption of terminal G·C pairs, which are then wrapped along the groove

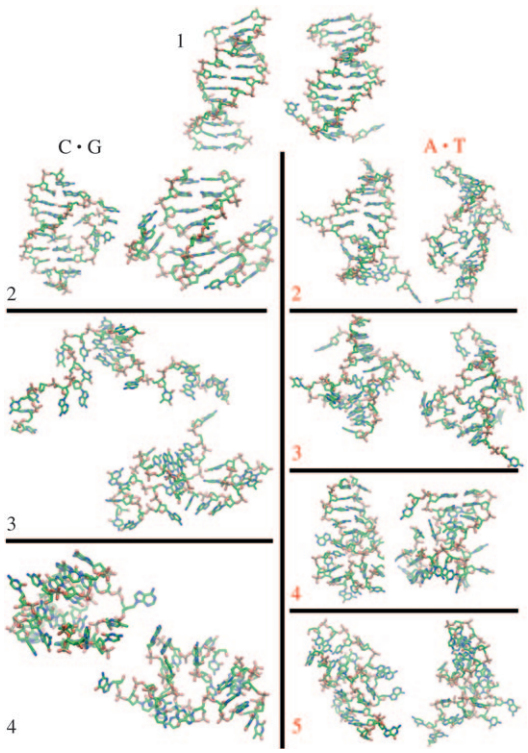

Figure 2. Representative clusters in the two pathways detected in PHT simulations. Starting from a canonical state (state 1), two pathways are detected: left: C·G-mediated unfolding (intermediates 2–4) and right: A·T-mediated pathway (states 2–5).

(as envisaged from WHT and PLT simulations). Distortions found in these two clusters are quickly reversible if the DNA is reimmersed in water (Figure S4, Supporting Information), which indicates that the trajectories are not far from the canonical state.

After 100–400 ns, state 2 evolves towards a very open and elongated structural family (state 3; Figure 2), in which the central d(AATT) core is lost in a single step (Figure 1) through the propagation of the distortion wave. The loss of canonical hydrogen bonding is not fully compensated by noncanonical inter- or intrastrand interactions, and the structure becomes very mobile losing any helical memory (Figure S4 and S5, Supporting Information). Not surprisingly, attempts to refold structures in state 3 by reimmersion in water were unsuccessful (on the 0.1 μs timescale). State 3 displays lifetimes around 50–250 ns and evolves in all cases into a myriad of diverse compact quasi-globular structural families (state 4), which have quite long lifetimes (in many cases above 200 ns; Figure 2) and are entropically restricted (Figure S6 A, Supporting Information).

Interestingly, one of the unfolding trajectories (PHT4) shows a completely different unfolding pathway, as the perturbation is generated in the central AATT track (Figures 1 and 2), which is fully disrupted in a very fast process (100 ns) after 0.4 μs from the beginning of the trajectory (Figure 1). The distortion wave is then transmitted to non-terminal d(G·C) pairs 0.3–0.5 μs later and after 1.6 μs of trajectory the unfolded structures remain quite compact and globular. This resembles the situation found in the dominant

unfolding routes defined by the "fraying–peeling" mechanism.

Our simulation results clearly demonstrate the stochastic nature of the unfolding process and the fact that different routes with different fluxes and residence times coexist, which means that concepts such as "unfolding pathway" are meaningless. We strongly support recent experimental models derived from quenching measurements,[21] which point to the existence of many states for the unfolding of small duplexes and, with ultrafast spectrometric measures of the unfolding of small DNA/RNA hairpins,[4] suggest a compact nature of the unfolded state. The bulk of the results presented argue against a classical two-state model as the dominant mechanism for folding/unfolding of small duplexes like the one considered here, since many compact intermediates are expected to be sampled in multiple unfolding pathways. Note that even though our results do not support a two-state model, our trajectories predict chromicity changes (Figure S6 B,C, Supporting Information) that could be erroneously interpreted using two-state models. Clearly, caution is necessary when employing macroscopic data to derive microscopic models.

The present results reveal important information on the factors determining the denaturing properties of Pyr. Analysis of the solvent (Pyr and water) environment around DNA demonstrates that DNA generally prefers to be solvated by water, as shown in the radial distribution functions for the first ("native-like") state (Figures 3 and 4). As the unfolding

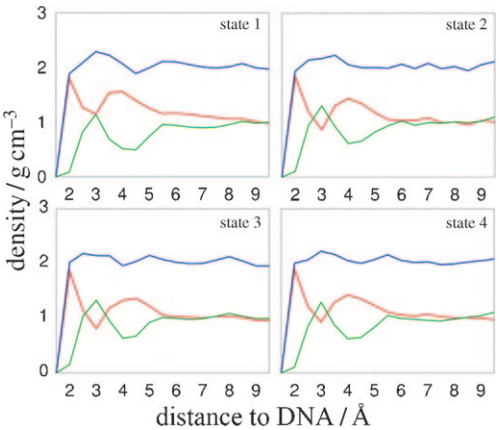

Figure 3. Solvent radial distribution density functions in PHT simulations. Values are always referred to the respective solvent density. Blue: total solvent; red: water; green: Pyr.

progresses, many new nucleobase–Pyr stacking interactions are created (Table S1 in the Supporting Information and Figure 4), which leads to a spectacular gain in the number of Pyr molecules in contact with the DNA (see Figures 3 and 4). Clearly Pyr stabilizes open forms by offering extrahelical nucleobases hydrophobic contacts above and below the aromatic rings, which do not affect their ability to establish in-the-plane polar contacts with water. The results presented here show that Pyr itself does not accelerate DNA unfolding microscopically, which is actually triggered by thermal energy and water molecules, but acts by "freezing" the unfolded

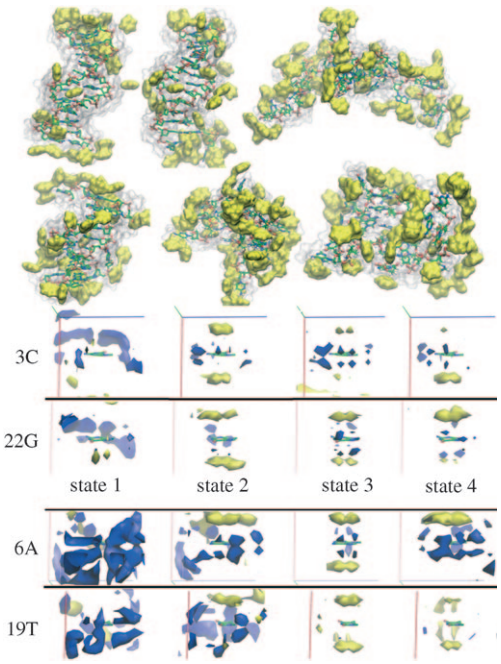

Figure 4. Solvent effect on DNA (PHT1) simulation. Top: DNA with surrounding water (transparent white) and pyridine (yellow) represented explicitly. Stacking of pyridine units to ends in the initial steps and opened bases during unfolding is observed. Bottom: Solvent density maps of water (blue, 0.9 g cm^{-3}) and pyridine (yellow, 0.6 g cm^{-3}) around selected bases (C·G and A·T pairs) in different unfolding states obtained in the PHT1 simulation (see Figure 2).

microstates through stacking with the solvent-exposed nucleobases when unfolding microevents occur. This prevents refolding, which then increases the macroscopic unfolding rate by making unfolding microevents more productive.

Overall, the present simulations demonstrate the complexity of the unfolding process of even small fragments of DNA and the unique capability of current MD simulation protocols to describe it at the atomic level. Much caution is needed, however, before extrapolating these results to much larger oligomers and especially to in vivo chromatin, in which the DNA is attached to many proteins (including histones) that certainly modulate its physical properties. The extrapolation of unfolding results to folding should also be made with caution, but our results seem to suggest that folding will be limited by the encountering of the two strands in a productive way, since a bad encounter can yield to the formation of abortive compact intermediates, the disruption of which might be very slow. We can speculate that duplex folding will not follow a single pathway and that formation of d(G·C) pairs will probably be among the first events happening in productive routes.

Received: February 1, 2010
Revised: March 8, 2010
Published online: May 17, 2010

Keywords: conformation analysis · DNA · molecular dynamics · unfolding

[1] L. M. Ying, M. I. Wallace, D. Klenerman, *Chem. Phys. Lett.* **2001**, *334*, 145.
[2] S. V. Kuznetsov, Y. Q. Shen, A. S. Benight, A. Ansari, *Biophys. J.* **2001**, *81*, 2864.
[3] G. Bonnet, O. Krichevsky, A. Libchaber, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8602.
[4] H. R. Ma, C. Z. Wan, A. G. Wu, A. H. Zewail, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 712.
[5] N. L. Goddard, G. Bonnet, O. Krichevsky, A. Libchaber, *Phys. Rev. Lett.* **2000**, *85*, 2400.
[6] J. Y. Jung, A. Van Orden, *J. Am. Chem. Soc.* **2006**, *128*, 1240.
[7] H. R. Ma, D. J. Proctor, E. Kierzek, R. Kierzek, P. C. Bevilacqua, M. Gruebele, *J. Am. Chem. Soc.* **2006**, *128*, 1523.
[8] A. Ansari, S. V. Kuznetsov, Y. Q. Shen, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7771.
[9] Y. Q. Shen, S. V. Kuznetsov, A. Ansari, *J. Phys. Chem. B* **2001**, *105*, 12202.
[10] E. J. Sorin, Y. M. Rhee, V. S. Pande, *Biophys. J.* **2005**, *88*, 2516.
[11] E. J. Sorin, Y. M. Rhee, B. J. Nakatani, V. S. Pande, *Biophys. J.* **2003**, *85*, 790.
[12] S. Kannan, M. Zacharias, *Biophys. J.* **2007**, *93*, 3218.
[13] M. M. Lin, L. Meinhold, D. Shorokhov, A. H. Zewail, *Phys. Chem. Chem. Phys.* **2008**, *10*, 4227.
[14] K. Qamhieh, K. Y. Wong, G. C. Lynch, B. M. Pettitt, *Int. J. Numer. Anal. Methods Geomech.* **2009**, *6*, 474.
[15] K. Y. Wong, B. M. Pettitt, *Biophys. J.* **2008**, *95*, 5618.
[16] A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, M. Orozco, *Biophys. J.* **2007**, *92*, 3817.
[17] H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 2179.
[18] T. E. Cheatham, P. Cieplak, P. A. Kollman, *J. Biomol. Struct. Dyn.* **1999**, *16*, 845.
[19] A. Pérez, F. J. Luque, M. Orozco, *J. Am. Chem. Soc.* **2007**, *129*, 14739.
[20] J. Shao, S. W. Tanner, N. Thompson, I. Cheatham, *J. Chem. Theory Comput.* **2007**, *3*, 2312.
[21] A. Montrichok, G. Gruner, G. Zocchi, *Europhys. Lett.* **2003**, *62*, 452.

