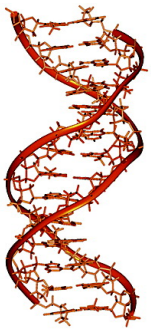
Article

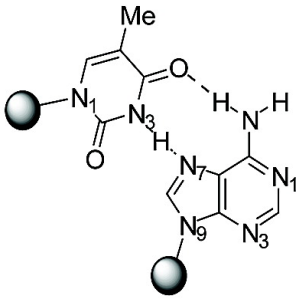
Theoretical Study of a New DNA Structure: The Antiparallel Hoogsteen Duplex

Elena Cubero, Nicola G. A. Abrescia, Juan A. Subirana, F. Javier Luque, and Modesto Orozco

J. Am. Chem. Soc., **2003**, 125 (47), 14603-14612• DOI: 10.1021/ja035918f • Publication Date (Web): 04 November 2003



Antiparallel Hoogsteen Duplex



Hoogsteen A:T base pair

More About This Article

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

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- 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

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 11/04/2003

Theoretical Study of a New DNA Structure: The Antiparallel Hoogsteen Duplex

Elena Cubero,[†,⊥] Nicola G. A. Abrescia,[‡] Juan A. Subirana,[§] F. Javier Luque,*[,∥] and Modesto Orozco*[,†,⊥]

Contribution from the Molecular Modeling & Bioinformatic Unit, Institut de Recerca Biomèdica, Parc Científic de Barcelona, Josep Samitier 1-5, Barcelona 08028, Spain, Division of Structural Biology, Wellcome Trust Centre for Human Genetics, Oxford University, Oxford OX3 7BN, United Kingdom, Departament d'Enginyeria Química, Universitat Politécnica de Catalunya, Avgda Diagonal 647, Barcelona 08028, Spain, Departament de Fisicoquímica, Facultat de Farmacia, Universitat de Barcelona, Avgda Diagonal 643, Barcelona 08028, Spain, and Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain

Received May 2, 2003; E-mail: modesto@mmb.pcb.ub.es

Abstract: The structure of a new form of duplex DNA, the antiparallel Hoogsteen duplex, is studied in polyd(AT) sequences by means of state-of-the-art molecular dynamics simulations in aqueous solution. The structure, which was found to be stable in all of the simulations, has many similarities with the standard Watson−Crick duplex in terms of general structure, flexibility, and molecular recognition patterns. Accurate MM-PB/SA (and MM-GB/SA) analysis shows that the new structure has an effective energy similar to that of the B-type duplex, while it is slightly disfavored by intramolecular entropic considerations. Overall, MD simulations strongly suggest that the antiparallel Hoogsteen duplex is an accessible structure for a polyd(AT) sequence, which might compete under proper experimental conditions with normal B-DNA. MD simulations also suggest that chimeras containing Watson−Crick duplex and Hoogsteen antiparallel helices might coexist in a common structure, but with the differential characteristics of both type of structures preserved.

Introduction

Under physiological conditions, DNA is mostly found as a right-handed double helix that shows a conformation, named B-DNA, not far from that suggested by Watson and Crick.[1] However, 50 years after the discovery of the double helix, it is clear that DNA is very polymorphic and that its equilibrium conformation can change depending on sequence and environment.[2−4] Repetitive sequences are especially rich in terms of accessible helical structures (for reviews, see refs 2, 4, and 7). For example, poly(GC), poly(G), and poly(GA) sequences are known to form easily left-handed Z-DNA,[5] four-stranded helical structures,[6] and triple helices,[7] respectively. Other examples are sequences containing poly(C) or poly(CT) tracks, which can generate the i form of DNA,[8,9] and sequences rich in polyd(A) tracks, which can yield parallel duplexes showing reverse Watson−Crick pairings, parallel duplexes with Hoogsteen pairings,[10] or parallel duplexes involving $d(A{\cdot}A)^+$ dimers.[2] The poly(AT) sequences are probably those exhibiting the widest range of accessible structures.[11−21] Thus, crystal structures of $d(AT)_2$ show a distorted B-type conformation.[11,12] Other crystal

† Institut de Recerca Biomèdica.
‡ Oxford University.
§ Universitat Politécnica de Catalunya.
∥ Departament de Fisicoquímica, Universitat de Barcelona.
⊥ Departament de Bioquímica i Biologia Molecular, Universitat de Barcelona.

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.
(2) (a) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984. (b) Ghosh, A.; Manju, B. *Acta Crystallogr.* **2003**, *D59*, 620.
(3) Bloomfield, V. A., Crothers, D. N., Tinoco, I., Eds. *Nucleic Acids: Structures, Properties and Functions*; University Science Books: Sausalito, 2000.
(4) Blackburn, G. M., Gait, M. J., Eds. *Nucleic Acids in Chemistry and Biology*; IRL Press: Oxford, 1990.
(5) Wang, A. H.-J.; Quiqkey, G. J.; Kolpak, F. J.; Crawford, J. L.; van Boom, J. H.; van der Marel, G.; Rich, A. *Nature* **1979**, *282*, 680.
(6) Laughlan, G.; Murchie, A. I. H.; Norman, D. G.; Moore, M. H.; Moody, P. C. E.; Lilley, D. M.; Luisi, B. *Science* **1994**, *265*, 520.
(7) Robles, J.; Grandas, A.; Pedroso, E.; Luque, F. J.; Eritja, R.; Orozco, M. *Curr. Org. Chem.* **2002**, *6*, 1333 and references therein.
(8) Gehring, K.; Leroy, J. L.; Guéron, M. *Nature* **1993**, *363*, 561.
(9) Gray, D. M.; Vaughn, M.; Ratliff, R. L.; Hayes, F. N. *Nucleic Acids Res.* **1980**, *8*, 3695.
(10) (a) Cubero, E.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2001**, *123*, 12018. (b) Cubero, E.; Aviñó, A.; de la Torre, B. G.; Frieden, M.; Eritja, R.; Luque, F. J.; González, C.; Orozco, M. *J. Am. Chem. Soc.* **2002**, *124*, 3133.
(11) Viswamitra, M. A.; Kennard, P.; Jones, P. G.; Sheldrick, G. M.; Salisbury, S.; Falvello, L.; Shakked, Z. *Nature* **1978**, *273*, 687.
(12) Viswamitra, M. A.; Shakked, Z.; Sheldrick, J. G. M.; Salisbury, S. A.; Kennard, O. *Biopolymers* **1982**, *21*, 513.
(13) Guzikevich-Guerstein, G.; Shakked, Z. *Nat. Struct. Biol.* **1996**, *3*, 32.
(14) Radwan, M. M.; Wilson, H. R. *Int. J. Biol. Macromol.* **1982**, *4*, 145.
(15) Davis, D. R.; Baldwin, R. L. *J. Mol. Biol.* **1963**, *6*, 251.
(16) Brahms, S.; Brahms, J.; Van Holde, K. E. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 3453.
(17) (a) Abrescia, N. G. A.; Thompson, A.; Dinh, T. H.; Subirana, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2806. (b) Abrescia, N. G. A.; Subirana, J. A. *Acta Crystallogr.* **2002**, *D58*, 2205.
(18) Aishima, J.; Gitti, R. K.; Noah, J. E.; Gan, H. H.; Schlick, T.; Wolberger, C. *Nucleic Acids Res.* **2002**, *30*, 5244.
(19) Lefèvre, J. F.; Lane, A. N.; Jardetzky, O. *Biochemistry* **1988**, *27*, 1086.
(20) Schmitz, U.; Sethson, I.; Egan, W. M.; James, T. L. *J. Mol. Biol.* **1992**, *227*, 510.

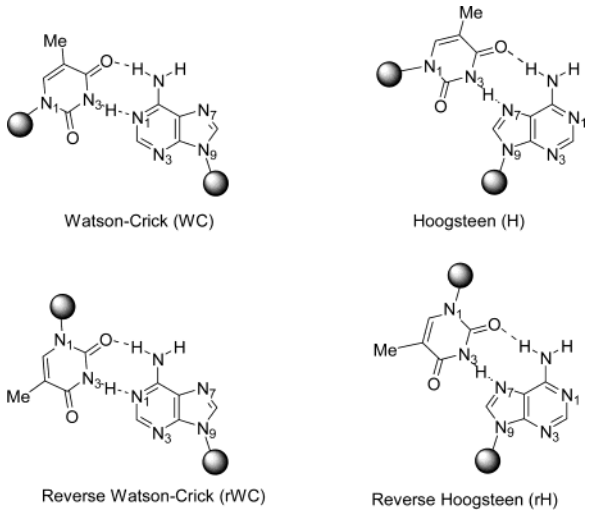


Figure 1. Schematic representation of the possible recognition patterns between neutral adenine and thymine.

structures even show a larger distortion, suggesting the existence of a very special type of B DNA named TA-DNA.[13] Early fiber diffraction data on polyd(AT) sequences suggested left-handed models,[14] but this has not been confirmed by higher resolution techniques. On the other hand, polyd(AT) fibers generated under low hydration conditions show a D form of DNA,[15] while at low temperature the C form is obtained.[16] Very recently, Subirana and co-workers[17] have reported the high-resolution structure of the d(ATATAT) duplex, which corresponds to a new structural motif of DNA consisting of an antiparallel right-handed double helix with adenosines in the syn conformation and unexpected Hoogsteen d(A·T) pairings. To our knowledge, the existence of this form in aqueous solution has not been confirmed by NMR experiments.

The possibility of different H-bond patterns for the A·T pair (see Figure 1) has been known since the 1960s.[22] Accurate theoretical calculations[23] showed that in fact the Watson−Crick pairing is not the best recognition mode for the A·T pair and that all of the four recognition patterns shown in Figure 1 are accessible, at least in the gas phase. Molecular dynamics simulations have also shown that stable helices can be built for both duplexes[10a] and triplexes[24] using Hoogsteen A·T pairings. Experimentally, the Hoogsteen recognition mode is found in different structures of DNA and RNA,[25] including the parallel triplexes,[7,26] where it stabilizes both d(A−T·T) and d(G−C·C) triads. Very interestingly, Hoogsteen pairs are common in complexes between duplex DNA and drugs or proteins.[18,27−29] In conjunction with the biological role of d(AT)$_n$ regions,[27,30] these findings strongly suggest an important role for Hoogsteen pairings in the modulation of gene expression.[18,27] In summary, recent theoretical and experimental information points out that the apparently exotic Hoogsteen interaction might be common and have a large biological significance.

In this paper, we present a wide and systematic molecular dynamics (MD) study of the poly(AT) duplex in aqueous solution at low ionic strength. Both the normal Watson−Crick (B form) and the antiparallel Hoogsteen (apH) structures[17] are studied using large MD trajectories for 4-, 6-, 8-, 10-, 12-, 14-, and 16-mer duplexes (more than 40 ns of trajectories are presented here). The structure, dynamics, and recognition properties of the apH helix in water are determined and compared to those of the standard B-DNA duplexes. The relative stability of apH and B-type duplexes is discussed using information derived from the MD simulations. Finally, trajectories of chimeras containing fragments of apH and B-DNAs are generated to analyze the impact of Hoogsteen pairings in the structure of a long canonical piece of DNA. The results obtained here complement previous X-ray data by Abrescia et al.[17] and provide a complete picture of a new and intriguing structure of the DNA duplex.

Methods

To analyze with good statistical accuracy[10a] the structure and stability of apH and B duplexes for poly(AT) sequences, we built starting structures for complementary duplexes d(AT)$_{n/2}$ for $n = 4$, 6, 8, 10, 12, 14, and 16 using standard Arnott's parameters for B-DNA[31] and the crystal structure data for the apH-DNA.[17] The 14 different structures were neutralized by adding a suitable number of Na^+ ions (10a) and immersing them in rectangular boxes containing between 1000 and 3500 water molecules. All of the systems were optimized, thermalized (298 K), and equilibrated using our standard multistage protocol.[24,32] The equilibrated structures were then subjected to 2 ns (4-, 6-, 8-, 14-, and 16-mer) and 5 ns (10- and 12-mer) of MD simulation at constant temperature (298 K) and pressure (1 atm) with standard relaxation times of 0.2 ps. Periodic boundary conditions were used to simulate a diluted environment, and the Particle Mesh Ewald method was used to account for long-range electrostatic effects.[33] Ewald tolerance of 5×10^{-6} was used in conjunction with a grid spacing of 1 Å and a four-order interpolation scheme. SHAKE[34] was used to constrain all of the bonds at their equilibrium positions, which allowed us to use a 2 fs time step for integration of Newton equations. AMBER-98[35] and TIP3P[36] forcefields were used to describe DNA and water. All of the MD simulations were performed using the AMBER6.0 suite of programs.[37]

(21) McAtter, K.; Ellis, P. D.; Kennedy, M. A. *Nucleic Acids Res.* **1995**, *23*, 3962.
(22) Hoogsteen, K. *Acta Crystallogr*. **1959**, *12*, 822.
(23) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1969.
(24) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1997**, *119*, 7563.
(25) (a) Voet, D.; Rich, A. *Prog. Nucleic Acid Res. Mol. Biol.* **1970**, *10*, 183. (b) Leontis, N. B.; Westhof, E. *Q. Rev. Biophys.* **1998**, *31*, 399. (c) Nolten, G. M. J.; Sijtema, N. M.; Otto, C. *Biochemistry* **1997**, *36*, 13241.
(26) (a) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1953**, *39*, 84. (b) Felsenfeld, G.; Davis, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023.
(27) Patikoglou, G. A.; Kim, J. L.; Sun, L.; Yang, S.-H.; Kodadek, T.; Burley, S. K. *Genes Dev.* **1999**, *13*, 3217.
(28) Rice, P. A.; Yang, S.; Mizuuchi, K.; Nash, H. A. *Cell* **1996**, *87*, 1295.
(29) (a) Wang, A. H.; Ughetto, G.; Quigley, G. J.; Hakoshima, T.; van der Marel, G. A.; van Boom, J. H.; Rich, A. *Science* **1984**, *225*, 1115. (b) Gilbert, D. E.; van der Marel, G. A.; van Boom, J. H.; Feigon, J. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 3006.
(30) (a) Moreau, J.; Maschat, M. F.; Kejzlarova-Lepesant, J.; Lepesant, J.-A.; Scherrer, K. *Nature* **1982**, *295*, 260. (b) Oosumi, T.; Garlick, B.; Belknap, W. R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 8886.
(31) Arnott, S.; Hukins, D. W. L. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1504.
(32) (a) Soliva, R.; Laughton, C. A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 11226. (b) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 5895.
(33) Darden, T. A.; York, D. M.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089.
(34) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(35) (a) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Fergurson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179. (b) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845.
(36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(37) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. L.; Marz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER6*; University of California: San Francisco, CA, 1999.

The MD simulations of apH and B duplexes were analyzed to evaluate their relative stability. For this purpose, the free energy of every structure was computed as shown in eq 1.[38] MD averages ($\langle\ \rangle$) for each contribution in eq 1 were obtained after dropping the first nanosecond of every trajectory. E_{intra} was computed using the standard AMBER-98 force-field.[35] The solvation contribution (G_{sol}) was determined using two different approaches (for a discussion of the two techniques, see ref 39): (i) finite difference Poisson−Boltzmann (PB) calculations as implemented in the MEAD program[40] for initial and final grid spacings of 1 and 0.4 Å, and (ii) the Generalized-Born (GB) method[41] as implemented in AMBER6.0. A dielectric constant of 80 was used to represent water, while dielectric constants of 1 or 2 were used to represent the solute. Finally, the contribution to the free energy of the system due to the intramolecular DNA entropy was determined by using Schliter's method[42] as described by Harris et al.[43] We should note that test calculations using the Andricioae and Karplus method[44] did not lead to any significant difference with respect to Schliter-type calculations and were then not included. To avoid artifacts in the averaging of the different terms, the end-base pairs were removed from the analysis.

$$G \approx \langle E_{intra}\rangle + \langle G_{sol}\rangle - T\Delta S_{intra} = \langle E_{effec}\rangle - TS_{intra} \qquad (1)$$

Calculations noted in eq 1 were performed for the different apH and B-type structures. Taking advantage of the linear relationship between the length of the oligonucleotide and the free energy (see ref 10a and below), we can then derive estimates of the relative stability of apH and B structures with very good statistical accuracy (see Results and Discussion).

A description of the interaction properties of apH (and B) structures was obtained from cMIP calculations for an O^+ probe molecule.[24,32,45] Solvation around the structures was represented by integrating the water population along the trajectory as explained elsewhere.[24,32] The analysis of molecular flexibility was performed using the principal component analysis (PCA) method as described by Sherer et al.[46] Principal components were obtained by diagonalization of the covariance matrix obtained after recentering all of the collected snapshots. The result of this procedure is a set of eigenvalues and the corresponding eigenvectors. Eigenvalues can be manipulated to obtain vibrational frequencies and provide useful information on the magnitude of the flexibility of DNA along their essential movements. The eigenvectors projected into the Cartesian space illustrate the nature of the essential movements of the DNA molecule. As shown previously,[47] we can compare the similarity between two unitary eigenvectors using the dot product, obtaining then a quantitative measure of the similarity between two essential movements. The strategy can be expanded (see eq 2 and ref 47) to quantitatively compute the similarity between a reduced set of eigenvectors, which are known to explain more than a certain amount of variance in the trajectory.[47] Note that due to the orthogonality between principal components, two identical trajectories will yield a similarity index γ of 1, while due to the limited number of eigenvectors considered two unrelated trajectories will have γ equal to 0.

$$\gamma_{AB} = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{n}(\nu_i^{A}\cdot\nu_j^{B})^2 \qquad (2)$$

where A and B stand for two different trajectories of equal length (for the same or different structures), ν stands for a unit eigenvector, and n is taken as 10, a number of modes which accounts for around 80% of the variance for the 10- and 12-mer oligonucleotides considered here. To reduce errors, only backbone atoms (up to C1′) were considered in comparison between B and apH helices.

A relative similarity index (κ), which accounts for the uncertainties intrinsic to any time-limited MD trajectory, can be obtained by normalizing the absolute index γ_{AB} by the self-similarities obtained in the A and B trajectories, $\gamma_{AA}{}^{T}$ and $\gamma_{BB}{}^{T}$ (see eq 3). Self-similarities are easily obtained by comparing the first and last halves of the same trajectory. Both γ and κ are 1 for two identical trajectories and 0 when they are orthogonal. As noted in ref 47, this index runs also from 1 (identical) to 0 (unrelated) and has the advantage of being less simulation-sensitive than the absolute similarity index (γ).

$$\kappa_{AB} = 2\frac{\gamma_{AB}}{(\gamma_{AA}{}^{T} + \gamma_{BB}{}^{T})} \qquad (3)$$

Geometrical analysis was performed by using the PTRAJ module of the AMBER6.0 program, as well as by using in-house programs. Helical analysis was carried out using Olson's X3DNA program[48] and Curves5.3.[49] Base pairs at the ends of the duplex were removed in all of the cases. All calculations were performed at the Centre de Supercomputació de Catalunya (CESCA) and in workstations in our laboratories.

Results and Discussion

Stability of Trajectories of polyd(AT)$_n$ in B and apH Forms. All of the trajectories of B and apH structures show stable helical conformations (see Figure 2), as noted in the small root-mean square deviations (rmsd) with respect to the MD-averaged structures (see Table 1). In general, lower rmsd values are found for the apH duplexes than for the B-like ones, suggesting that the former sample a narrower region of the configurational space, probably due to a greater rigidity of the apH structure (see below). Comparison of 2 and 5 ns trajectories demonstrates that the trajectories converge well in the 2 ns time scale and that apparently no relevant transitions are expected for longer simulation times. The all-atom rmsd values of the apH trajectories with respect to the crystal structure (PDB entry 1gqu[17]) are small (<2 Å in all of the cases), showing that MD simulations in aqueous solution sample regions of the configurational space very close to the crystal structure. This agreement supports the suitability of the MD simulation for the study of the apH duplex, and the ability of X-ray data to represent a reasonable conformation of the polyd(AT)$_n$ duplex in dilute solution.

The B-like trajectories are closer to the B than to the A form, but the rmsd values with respect to the canonical B form are larger than those typically found for simulations of B-DNA. For example, for the duplex d(CGCGAATTCGCG), the rmsd with respect to the canonical B form found in 10−20 ns

(38) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889.
(39) Orozco, M.; Luque, F. J. *Chem. Rev.* **2000**, *100*, 4187.
(40) (a) Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992**, *224*, 473. (b) Bashford, D. In *Scientific Computing in Object-Oriented Parallel Environments*; Reynders, J. V. M., Tholburn, M., Eds.; Springer: Berlin, 1997; pp 233−240.
(41) Still, W. C.; Tempcyck, A.; Hawley, R. C.; Hendrickson. T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
(42) Schliter, J. *Chem. Phys. Lett.* **1993**, *215*, 617.
(43) Harris, S.; Gavathiotis, E.; Searle, M. S.; Orozco, M.; Laughton, C. A. *J. Am. Chem. Soc.* **2001**, *123*, 12658.
(44) Andricioaei, I.; Karplus, M. *J. Chem. Phys.* **2001**, *115*, 6289.
(45) Gelpí, J. L.; Kalko, S.; de la Cruz, X.; Barril, X.; Cirera, J.; Luque, F. J.; Orozco, M. *Proteins* **2001**, *45*, 428.
(46) Sherer, E.; Harris, S. A.; Soliva, R.; Orozco, M.; Laughton, C. A. *J. Am. Chem. Soc.* **1999**, *121*, 5981.
(47) (a) Rueda, M.; Kalko, S. G.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2003**, *125*, 8007. (b) Orozco, M.; Noy, A.; Pérez, A.; Luque, F. J. *Chem. Soc. Rev.* **2003**, 32.
(48) Lu, X,-J.; Olson, W. K. X3DNA Program, Rutgers University, 2001.
(49) Lavery, R.; Sklenar, J. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63.

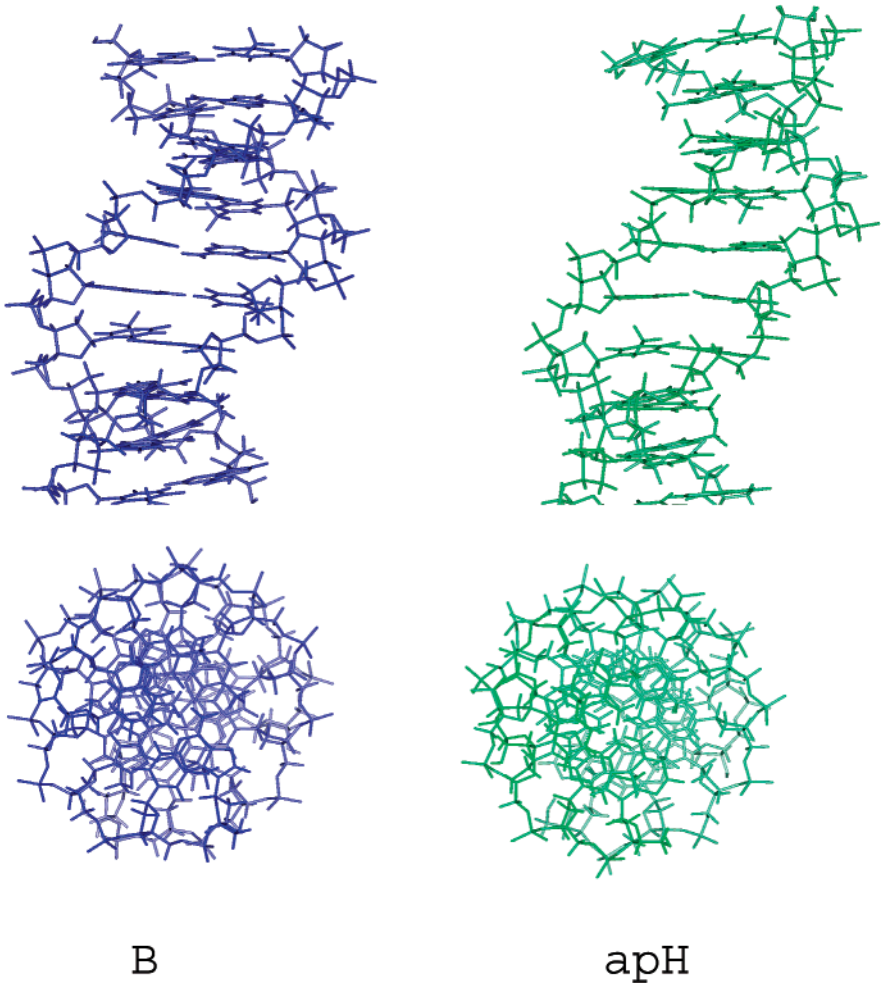


Figure 2. Representation of the MD-averaged structure of the $d(AT)_{n/2}$ for $n = 10$ in the B (left) and antiparallel Hoogsteen conformations (right). See text for details.

Table 1. All-Atom Root-Mean-Square Deviations (rmsd; Å) of B-DNA and apH-DNA Trajectories with Respect to Different Reference Structures (Standard Deviations (SD) in the Averages Are Also Displayed)

B-DNA

structure	average[b]	B-DNA[c]	A-DNA
B4	1.0 ± 0.2	1.4 ± 0.2	1.9 ± 0.2
B6	1.1 ± 0.2	2.1 ± 0.3	2.3 ± 0.2
B8	1.4 ± 0.3	2.5 ± 0.5	3.4 ± 0.6
B10[a]	1.5 ± 0.2	2.9 ± 0.5	3.7 ± 0.5
B12[a]	1.7 ± 0.4	3.3 ± 0.7	4.4 ± 0.7
B14	1.7 ± 0.3	3.4 ± 0.6	5.0 ± 0.6
B16	2.1 ± 0.6	3.9 ± 0.9	5.0 ± 0.8

apH-DNA

structure	average[b]	crystal structure[d]
apH4	1.5 ± 0.6	1.5 ± 0.2
apH6	1.0 ± 0.2	1.0 ± 0.2
apH8	0.8 ± 0.2	1.1 ± 0.2
apH10[a]	1.3 ± 0.3	1.9 ± 0.5
apH12[a]	1.3 ± 0.3	1.5 ± 0.3
apH14	1.5 ± 0.4	1.5 ± 0.3
apH16	1.3 ± 0.3	1.6 ± 0.3

[a] 5 ns trajectories, the rest are only 2 ns long. [b] MD-averaged structures obtained using the last 1 or 4 ns of trajectories. [c] Arnott's structures (see ref 31). [d] Structures built by joining two 5-mer structures in PDB entry 1gpu followed by restricted minimization.

Table 2. Occurrence (in % with Respect to the Maximum Number of Hydrogen-Bond Interactions) of Canonical Watson−Crick (WC) and Hoogsteen (H) Hydrogen Bonds for the Different Trajectories

structure	WC H-bonds	structure	H H-bonds
B4	96.6	apH4	73.8
B6	70.4	apH6	97.3
B8	88.6	apH8	97.2
B10	94.4	apH10	91.9
B12	91.8	apH12	94.8
B14	92.7	apH14	96.5
B16	93.3	apH16	97.1

trajectories in water,[50] using the same force-field and simulation protocol, is around 2.3 Å (1.9 Å if the crystal structure is used as reference), that is, around 1 Å lower than the rmsd found here for oligonucleotides of the same length. As previously noted by experimentalists (see Introduction), this deviation suggests that polyd$(AT)_n$ duplexes do not adopt a pure-canonical B structure in solution. However, for the sake of clarity, we will continue denoting as "B form" those structures obtained in the B trajectories of polyd$(AT)_n$ duplexes.

The hydrogen-bond (H-bond) patterns defining the Watson−Crick and Hoogsteen pairings are well preserved for B and apH trajectories (see Table 2). However, the total percentage of H-bonds in the apH trajectories is typically larger than that found for the B trajectories. Very interestingly, the general stability

(50) Rueda, M.; Cubero, E.; Laughton, C. A.; Luque, F. J.; Orozco, M., to be published.

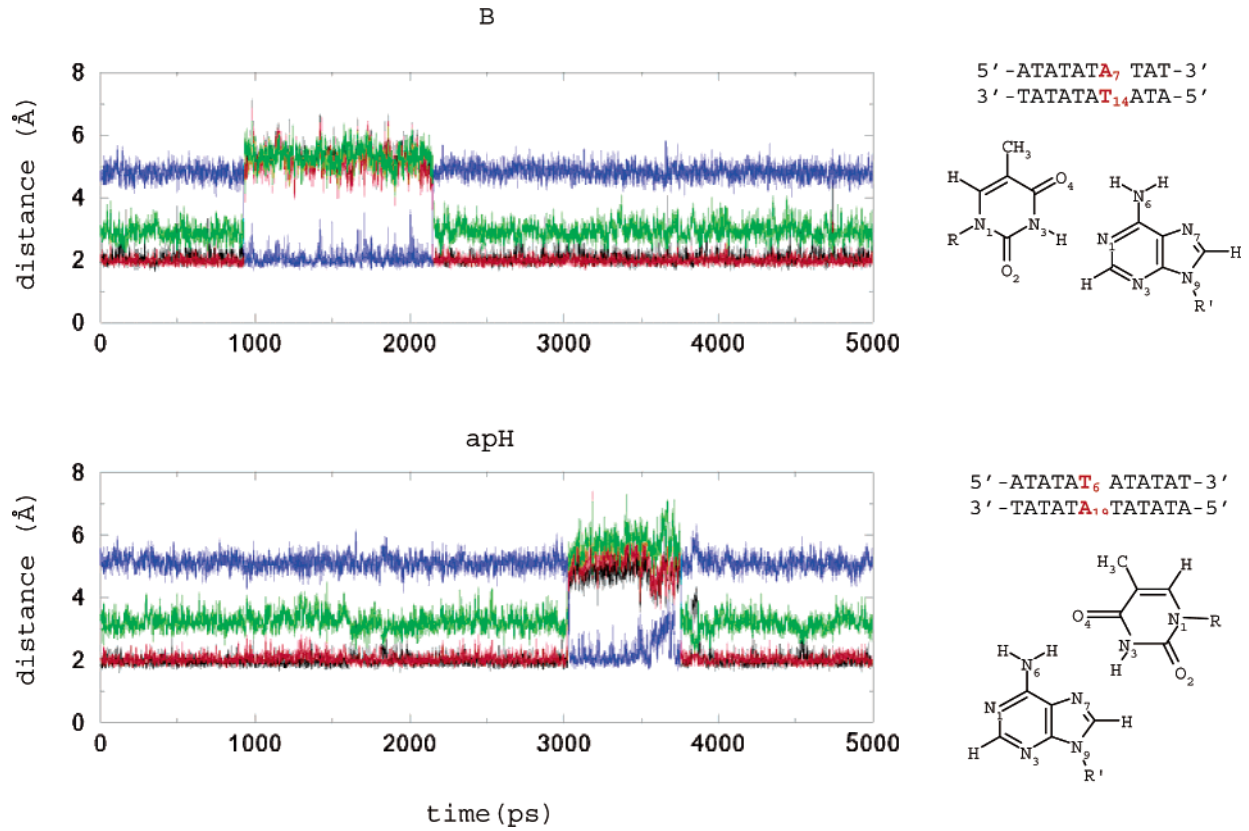


Figure 3. Evolution of key H-bond distances (in Å) in B (10-mer) and apH (12-mer) trajectories. The existence of partial breathing movements is seen in a temporary loss of canonical Watson−Crick or Hoogsteen hydrogen bonds, and the simultaneous gain of one additional noncanonical hydrogen bond. The color code is as follows. Top: black, H6(A_7)−O4(T_{14}); red, N1(A_7)−H3(T_{14}); green, H2(A_7)−O2(T_{14}); and blue, H6(A_7)−O2(T_{14}). Bottom: black, O4-(T_6)−H6(A_{19}); red, H3(T_6)−N7(A_{19}); green, O2(T_6)−H8(A_{19}); and blue, O2(T_6)−H6(A_{19}).

of Watson−Crick and Hoogsteen H-bonds does not preclude the existence of reversible, partial breathing movements, like those shown in Figure 3. These movements occur in the nanosecond time scale and do not lead to a complete opening of the bases, which always retain at least one nonstandard H-bond (see Figure 3). They are, in general, related to movements of thymine toward the major groove and are more common in the B trajectories.

Description of the apH Structure, Interaction Properties, and Dynamics. The analysis of the 5 ns trajectories of $d(AT)_5$ and $d(AT)_6$ apH helices provides a complete structural picture of medium-sized apH helices in aqueous solution, complementing the X-ray information derived by Abrescia et al.[17] The general shape of the apH helix is similar to that of a normal B-type duplex (see Figure 2), a quite surprising finding considering the dramatic differences in terms of nucleotide conformation and H-bond recognition pattern between the two helices. The periodicity of the apH helix is similar to that of B-DNA (see Table 3 and ref 51), the rise is only slightly larger (around 0.1 Å), the inclination is small and very similar to that found in B-trajectories, and the average phase angles suggest south and east-south puckerings, as is usual in crystal structures of B-DNA.[52] The minor groove in apH trajectories is around 2 Å narrower than that found in B-like trajectories, while the major groove of the apH structures is around 3 Å wider (see also Figure 2). There are differences between the two types of structures in the shortest C1′−C1′ distances (around 2 Å shorter for apH helices) and in the χ angles, which arise from the existence of the *syn*-2′-deoxyadenosines in the apH helices (the average χ values for apH 2′-deoxyadenosines are around 50°, in contrast to standard anti values of around −110°). Finally, it is worth noting the similarity between the X-ray structure of the apH helix and that found in MD simulations in solution (see Tables 1 and 3). In fact, the only noticeable discrepancy is found for the twist, which is 3° smaller in MD simulations, reproducing a well-known tendency of the AMBER force-field (see Table 3 and refs 47b and 53).

Table 3. Key Helical Parameters for the MD-Averaged B and apH Helices (for 10- and 12-mer Oligonucleotides)[a]

parameter	B10	B12	apH10	apH12	B-DNA	X-ray[c]
C1′−C1′[b]	10.4	10.3	7.9	8.3	10.9	8.2
inclination	4.5	6.4	3.2	3.4	−5.0	2.2
helical twist	33.2	32.3	32.0	32.0	36.0	35.3
helical rise	3.1	3.1	3.3	3.2	3.4	3.3
phase	132.0	128.4	118.9	121.1	191.7	155.1
amplitude	36.5	35.3	42.3	40.9	35.7	35.6
m-groove[d]	12.9	13.0	10.6	10.4	11.7	9.3−11
M-groove[d]	18.7	18.8	21.9	21.6	16.9	∼20

[a] The helical parameters of a standard B-DNA and of the structure generated from crystal data in ref 17 are also displayed. Distances are in Å, and angles are in deg. [b] Minimum C1′−C1′ distance. [c] Crystal structure from 1gpu (see text). [d] Determined from the shortest P−P distances across the groove.

(51) The existence of syn adenines makes the helical analysis of apH helices very difficult, even when the very flexible X3DNA program is used. The only parameters that can be safely derived from helical analysis are those displayed in Table 3.

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

(53) Cheatham, T. E.; Young, M. A. *Biopolymers* **2001**, *56*, 232.

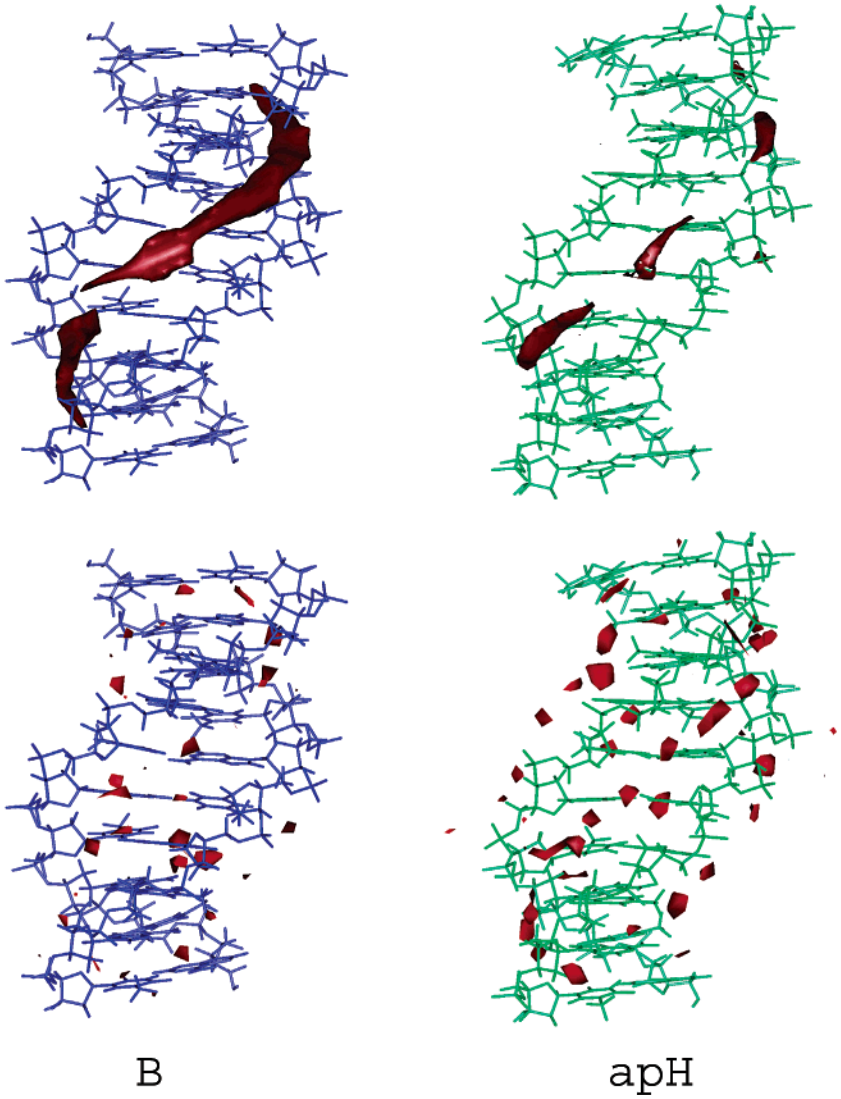


Figure 4. Molecular interaction potential contours (−5 kcal/mol; top) and solvation maps (2.5 g/mL; bottom) for the B (right) and apH (left) helices. The data for the 10-mer oligonucleotides are shown in all cases (similar profiles are obtained for the other sequences).

Despite the similar shape of B and apH helices in solution, it remains to be determined whether they exhibit similar interaction profiles. MIP maps (Figure 4) show that, as found for the B-DNA, the minor groove is the region with a stronger ability to interact with small cationic probes (smaller favorable regions are found in the major groove) in the apH structure. The shapes of MIP (isocontours of −5 kcal/mol) maps in apH and B helices are quite similar (see Figure 4). However, while the MIP isocontours are continuous for the B helix, they are not for the apH structures. This indicates that the minor groove in apH helices is slightly less charged than that of B helices (i.e., slightly less polar). However, the differences found in the MIP maps are less relevant than those expected a priori from the dramatic changes in the distribution of polar groups at the bottom of the grooves in B and apH helices (see Figure 1).

Both B and apH helices are very well solvated in our simulations. On average (for all of the sequences and excluding the terminal base pairs), there are 26 ± 7 (B) and 27 ± 12 (apH) water molecules per step in direct contact (distance less than 3.5 Å) to any heteroatom of DNA, indicating a very similar hydration. For the B trajectories (Figure 4), small amounts of water are located in the major groove, and most of the regions of large water concentration (>2.5 g/cm^3) are in the minor groove, tracing clearly Dickerson's spine of hydration.[52] For the apH structure, the spine of hydration along the minor groove is also clearly defined, with waters located in the vicinities of O2 (thymines). Interestingly, a secondary spine of hydration (see details in Figure 5) is obtained along the major groove, with regions of intense hydration around N6 and N1 of adenines and O4 of thymines. A very dense hydration region is always located around N3 of adenines, bridging this atom to the phosphate group. The preferred regions of hydration found in MD simulation perfectly agree with highly ordered crystal waters located by Abrescia et al.[17] in the major groove. Interestingly, the minor groove which is occupied by nucleobases in the crystal structure[17] is occupied by water in solution (see Figure 4). All of these regions of large hydration correspond also to regions where waters with long residence times (up to 200 ps) are located (see Figure 5).

PCA provides a general view of the essential motions of the duplex in B and apH conformations. The first frequencies are very low (around 10−20 cm^{-1}) and correspond to bending and

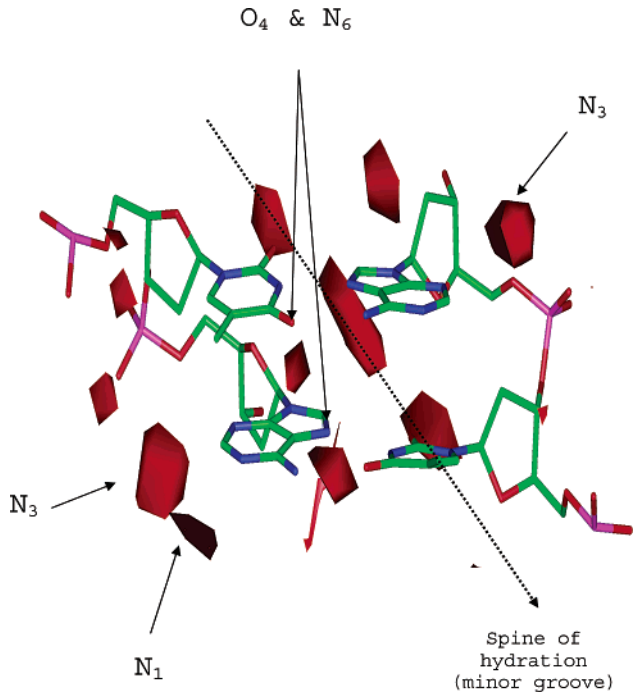


Figure 5. Details of regions of specific hydration in the apH helix (2.5 g/mL contours) along the major groove. Data correspond to the 10-mer structure. Water molecules with a long residence time are located in the areas. Residence times for waters bound to O_2 and O_4 are very large (150−200 ps), waters bound to N_3 show also large residence times (around 50−100 ps), waters around N_1 show residence times of 30−50 ps, and those around N_6 have residence times around 20 ps.

Table 4. Similarity Indexes γ Obtained Using Eq 1 and the 10 First Principal Components for the 10- and 12-mer Duplexes in B and apH Structures[a]

	B	apH
	10-mer	
B	0.7	0.5
apH		0.8
	12-mer	
B	0.7	0.6
apH		0.7

[a] Similarity indexes $\gamma_{B/apH}$ were computed using the last 4 ns of both trajectories. Self-similarity indexes $\gamma_{B/B}$ and $\gamma_{apH/apH}$ were determined comparing the first (1 → 3 ns) and second (3 → 5 ns) halves of the trajectories.

untwisting movements for the two structures. In all cases, only very small differences are found between the frequencies of the first essential motions of B and apH structures. For example, the first three frequencies of the 10-mer sequence are (in cm^{-1}) 15, 22, and 24 for the B helix, and 13, 16, and 27 for the apH structure. A quantitative comparison between the essential motions of B and apH helices by the absolute and relative similarity indexes $\gamma_{B/apH}$ and $\kappa_{B/apH}$ (eqs 2 and 3; Table 4) indicates a close similarity in the dynamics of the two helices.

In summary, MD simulations point out that many of the structural, reactive, and dynamic properties of the apH helix are similar to those of the B helix, suggesting that proteins which evolved to recognize the general shape of B DNA could also recognize apH DNA. The subtle differences at the bottom of grooves might then alter the specificity of the interactions, opening a wide range of possible biological roles for the apH helix.

Table 5. Effective Energy ($\langle E_{effec} \rangle$ in Bold), and Their Intramolecular and Solvation Contributions ($\langle G_{sol} \rangle$ in Roman/$\langle E_{intra} \rangle$ in *Italics*) Computed for the Different Oligonucleotides in B and apH Conformations[a]

length	B	apH
2	**−637.5 ± 0.4**	**−635.4 ± 0.4**
	−544.0 ± 0.4/*−93.5 ± 0.2*	−559.9 ± 0.4/*−75.6 ± 0.2*
4	**−1363.0 ± 0.7**	**−1364.9 ± 0.6**
	−1461.3 ± 0.6/*98.2 ± 0.3*	−1484.3 ± 0.5/*119.4 ± 0.2*
6	**−2093.3 ± 1.6**	**−2095.2 ± 1.1**
	−2647.0 ± 1.5/*553.7 ± 0.5*	−2682.3 ± 1.0/*587.1 ± 0.4*
8	**−2819.7 ± 1.9**	**−2821.9 ± 1.9**
	−4075.7 ± 1.8/*1256.1 ± 0.6*	−4083.8 ± 1.9/*1261.9 ± 0.6*
10	**−3548.9 ± 3.0**	**−3548.9 ± 2.6**
	−5681.4 ± 2.9/*2132.5 ± 1.0*	−5681.5 ± 2.4/*2132.6 ± 0.8*
12	**−4269.9 ± 3.3**	**−4276.5 ± 3.4**
	−7453.7 ± 3.2/*3183.8 ± 1.0*	−7387.8 ± 4.1/*3111.3 ± 1.3*
14	**−4999.7 ± 4.7**	**−5001.5 ± 3.6**
	−9350.3 ± 4.4/*4350.6 ± 1.4*	−9297.1 ± 3.4/*4295.6 ± 1.1*

[a] Values determined using PB/SA estimates of solvation free energy for internal and external dielectric constants of 1 and 80. Values and their standard errors (in parentheses) are in kcal/mol.

Relative Stability of B and apH Helices. The preceding discussion points out that the crystal structure found by Abrescia et al.[17] is not a lattice-artifact, but an intrinsically stable structure of polyd(AT)$_n$ in aqueous solution. However, to determine the possible biological role of apH helices, it is necessary to determine the relative stability of the apH helix versus the standard B-like conformation. For this purpose, we computed the free energy associated with each conformation using eq 1 (see Methods). It is worth noting that with 2−5 ns trajectories the average effective energy, $\langle E_{effec} \rangle = \langle E_{intra} \rangle + \langle G_{sol} \rangle$, can be estimated more accurately than the entropic term, which should need longer trajectories to converge.[10a] Thus, we have performed two separate analyses: one quantitative for $\langle E_{effec} \rangle$ and another qualitative for ΔS_{intra}.

The effective energies for the B and apH helices are shown in Table 5 (PB/SA values were used here to compute the solvation free energy; similar results were obtained with the GB/SA method; see below). The similarity in the $\langle E_{effec} \rangle$ values determined for the two conformations is evident. Thus, for the 10- and 12-mer oligonucleotides (those for which the longest trajectories are available), the $\langle E_{effec} \rangle$ values differ by only 0 and 3 kcal/mol, that is, less than 0.06% of the total energy value, and lie within the estimated error of the averages (see Table 5).

A more quantitative estimate of the relative effective energy for B and apH helices can be obtained,[10a] taking advantage of the linear dependence between $\langle E_{effec} \rangle$ and the length (n) of the oligonucleotide ($r^2 = 1.0000$ for the regression line $\langle E_{effect} \rangle = a \times (n - 2) + b$ in all cases). Typical errors (see Figure 6) in the slope (a: helix-growth stabilization factor) and intercept (b: nucleation energy) are only around 0.2−0.5 and 2−4 kcal/mol, respectively. The regression equations (see Figure 6) suggest that there is not a significant preference for either B or apH conformations in aqueous solution in terms of $\langle E_{effec} \rangle$ (the same trends were found when solvation was determined using the GB/SA model, and when PB/SA calculations were repeated considering the internal dielectric constant equal to 2; data not shown). This supports the equivalence in effective energy between B and apH structures and suggests that the population of the two helices in solution might depend on the existence of cofactors, specific hydration waters, and entropic considerations. Due to the similar degree of hydration in the first hydration

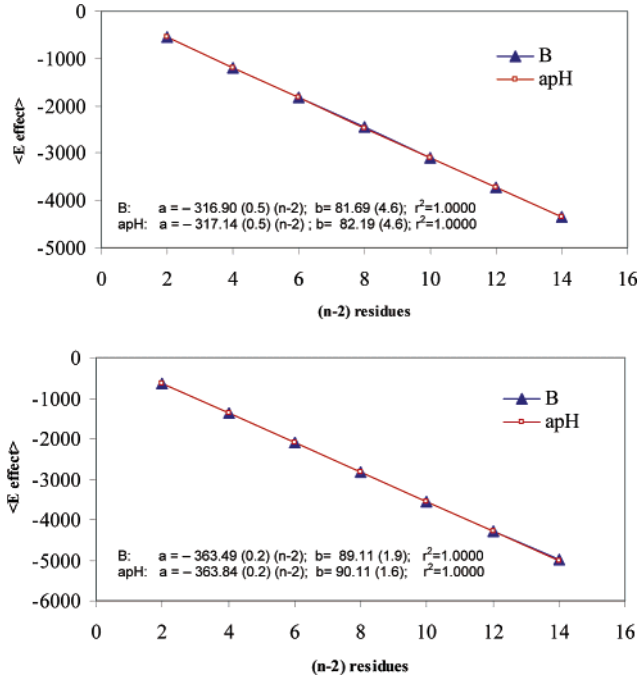


Figure 6. Representation of the variation of the effective energy $\langle E_{effec} \rangle = \langle E_{intra} \rangle + \langle G_{sol} \rangle$ as a function of the length of the oligonucleotide for B and apH conformations. Top: Solvation computed using the GB/SA method. Bottom: Solvation computed using the PB/SA approach. In both cases, interior and exterior dielectric constants are 1 and 80. All values are in kcal/mol.

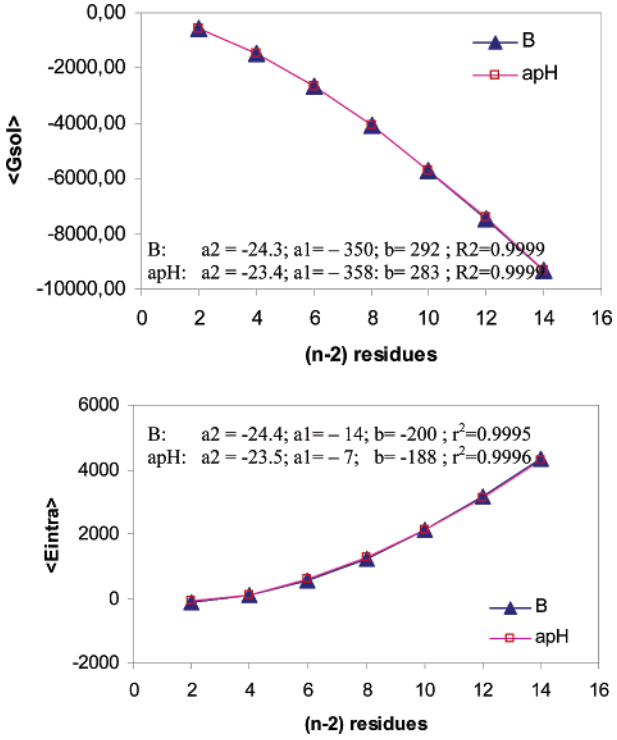


Figure 7. Representation of the variation of the solvation free energy (top: $\langle G_{sol} \rangle$) and intramolecular energy (bottom: $\langle E_{intra} \rangle$) as a function of the length of the oligonucleotides for B and apH conformations (all values are in kcal/mol). Solvation values displayed here correspond to the PB/SA calculations with dielectric constants of 1 (interior) and 80 (exterior). Identical profiles where obtained using other solvation models (see text and Table 5). Parabolic fits correspond to the equation: $y = a_2x^2 + a_1x + b$.

shell (see above), no large differences in stability are expected to arise from interaction with specific water molecules. The impact of entropic effects and cofactors will be discussed below.

To gain further insight into the similar effective energies of B and apH structures, we examined their intramolecular and solvation contributions (see eq 1 and Table 5). As found in previous work,[10a] there is a second-order polynomial relationship between the size of the duplex and its solvation and intramolecular energy components. The linear relationship found for E_{effec} then stems from the cancellation of the second-order coefficients for the solvation and intramolecular terms (see Figure 7). In general, for medium to long oligonucleotides, the B form is slightly favored by solvation and slightly disfavored by intramolecular interactions with respect to the apH conformation. However, the differences are small (see Table 5 and Figure 7), and we can conclude that not only the B and apH structures are similar in terms of the effective energies, but also in their solvation and intramolecular contributions.

The similar solvation free energy of apH and B helices is not surprising considering the solvation maps (Figures 4 and 5) and the number of water molecules in the first hydration shell (see above). However, the similar intramolecular energy is surprising considering that apH duplexes imply the existence of *syn*-adenosines, which are expected to be less stable than the anti conformation. However, NMR data in DMSO and water[54] show that the free energy difference between syn and anti conformations is small (0.4−0.6 kcal/mol for adenosine). A similar value (0.7 kcal/mol) is also found by restricted AMBER-99 optimizations of syn and anti adenosines at the MD-averaged χ torsions found in our B and apH simulations. This small energy penalty and others related to unfavorable backbone−backbone interactions in the apH helix are compensated by the better stacking and H-bond interactions in Hoogsteen A−T pairs. Thus, AMBER calculations in the gas phase using the MD-averaged structures of the Watson−Crick and Hoogsteen pairs show interaction energies of −12.7 (WC) and −13.2 (H) kcal/mol; that is, the Hoogsteen H-bond pattern is around 0.6 kcal/mol more stable than the Watson−Crick scheme.[55] Linear regression lines (see Figure 8) between the H-bonding energy and the length of the oligonucleotides suggest that, in the DNA environment, the Hoogsteen scheme is around 1.3 kcal/(mol step) more stable than the Watson−Crick pattern. Similarly (Figure 8), the stacking energy of the Hoogsteen pair is on average around 1.9 kcal/(mol step) more stable than that of the Watson−Crick pair in the DNA environment.

In summary, the MM/PB-SA (and MM/GB-SA) analysis, despite its shortcomings, strongly suggests similar stability for the B and apH structures in aqueous solution in terms of intramolecular energy and solvation free energy. The relative population of the two helices might then be determined by entropic considerations. The structural and dynamical analysis described above points out that the B form is more flexible than the apH form, the former being then entropically favored. To obtain more detailed information, entropic calculations were

(54) Stolarski, R.; Pohorille, A.; Dudycz, L.; Shugar, D. *Biochim. Biophys. Acta* **1980**, *610*, 1.

(55) These values agree very well with our own B3LYP/6-31G(d) estimates (−12.1 and −12.6 kcal/mol, respectively), and with MP2/6-31G*(0.25) estimates by Sponer and Hobza (−11.8 (WC) and −12.7 kcal/mol, respectively, from ref 23).

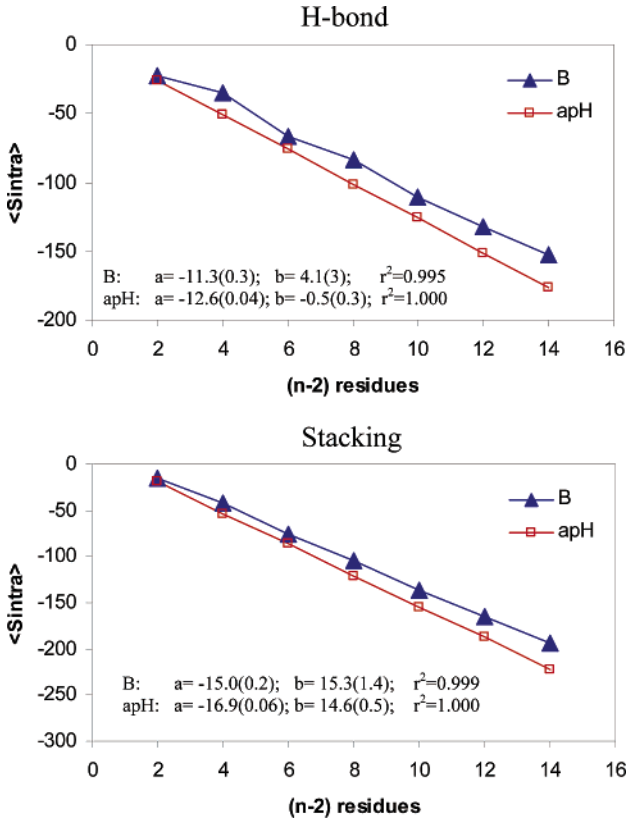


Figure 8. Representation of the variation of the hydrogen bond (top) and stacking energies (bottom) as a function of the length of the oligonucleotides for B and apH conformations (all values are in kcal/mol).

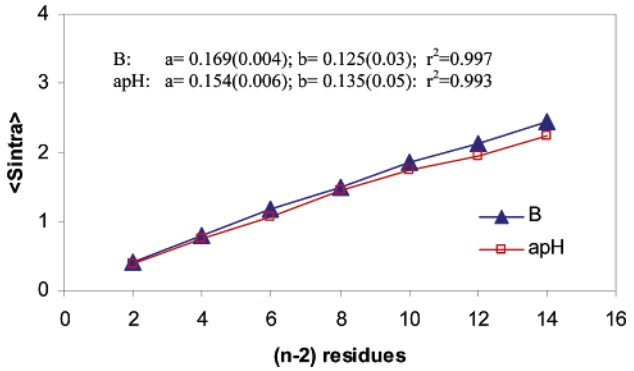


Figure 9. Representation of the variation of the entropy (in kcal/(mol K)) as a function of the length of the oligonucleotides for B and apH conformations (all values are in kcal/mol). All values were obtained from 1 ns samplings (see text).

performed using Schlitter's method, as described in the Methods section. Although accurate entropic calculations would need longer trajectories than those performed here, the analysis of 2−5 ns trajectories can still be useful to obtain relative entropy values.

Entropy calculations were then performed for all of the trajectories in the 1−2 ns period, finding linear relationships with the length of the oligonucleotides ($r^2 > 0.993$; see Figure 9). The regression lines confirm that the B form is entropically favored as compared to the apH helix in all cases and that the difference slightly increases with the length of the oligonucleotide. For example, the regression equations indicate that for the 10- and 12-mer oligonucleotides the B structure is favored by 0.11 and 0.14 kcal/(mol K) with respect to the apH helix. Similar values are obtained when entropies are calculated using the 5 ns trajectories and extrapolated to infinite simulation time using Harris's method:[43] 0.15 (10-mer) and 0.21 (12-mer) kcal/(mol K). These values suggest that at $T = 298$ K the entropic effect will be around 45 (10-mer) or 63 (12-mer) kcal/mol, a sizable difference considering the small differences in effective energies, which would lead to a predominance of the B form in aqueous solution. We can expect that when the flexibility of the DNA is restricted by the presence of drugs, proteins, or lattice constraints, the intrinsic flexibility of the DNA is drastically altered, modifying then the entropy balance between B and apH helices. In the latter cases, the apH helix could become competitive with the B structure, explaining the existence of apH structures in the crystal[17] and the presence of apH recognition modes in complexes of DNA with either intercalators or proteins (see Introduction).

Chimeras of B and apH Helices. In long DNA fragments, the apH structure may occur in short fragments surrounded by the B-structure, and often the fragments will contain both d(AT) and d(AA) steps. Accordingly, whether an apH track can be inserted into a B canonical structure without inducing large distortions, and whether it can exist in polyd(A) sequences, are key issues in guessing the biological importance of apH helices. Test calculations of poly(dA)$_{10}$ (data not shown, but available upon request) show that the apH helix can also be stable: the average effective energy (eq 1) found in 2 ns trajectories of the apH helix (−2663 kcal/mol) was almost identical to that found for a WC duplex of the same sequence (−2662 kcal/mol; ref 10a) and was slightly more negative than that found for parallel stranded duplexes with reverse Watson−Crick (−2642 kcal/mol; ref 10a) and Hoogsteen pairings (−2657 kcal/mol; ref 10a). In summary, our MD simulations suggest that the apH helix can exist also in poly(dA) sequences, opening then the window of structures for which apH helices are possible.

The impact of having apH sequences inserted in B-DNA was studied in two chimeric 11-mer sequences: d(ATAT**AAA**ATAT) and d(**ATAT**AAA**ATAT**), where the steps having the Watson−Crick recognition mode are in plain text, and those with the apH pattern are in bold text. Unrestricted NPT trajectories collected for 2 ns after equilibration are very stable and reveal no strong destabilization, conformational transition, or unfolding (see Figures 10 and 11). The backbone rmsd values with respect to the average structures are around 1.3 Å in the two cases (Figure 11), thus confirming the stability of the trajectories. The apH-rich d(**ATAT**AAA**ATAT**) structure is closer to an apH helix than to a B conformation (backbone rmsd values of 1.9 and 2.4 Å, respectively), while the Watson−Crick-rich d(ATAT**AAA**ATAT) structure is intermediate between apH and B conformations (backbone rmsd values of 1.9 and 1.8 Å, respectively). Very interestingly, the chimeras maintain clearly separable conformations corresponding to the regions with apH or Watson−Crick pairings (see Figure 10). Such a structural difference has a clear impact on the recognition properties of the helix (see Figure 10). Thus, for the d(ATAT**AAA**ATAT) helix, the −5 kcal/mol contour is discontinuous at the central region, as expected for an apH helix, and contiguous in the rest, as in a canonical B-helix (see Figure 4). On the contrary, the reverse trend occurs for the d(**ATAT**AAA**ATAT**) sequence. This finding indicates that not only apH and B helices are

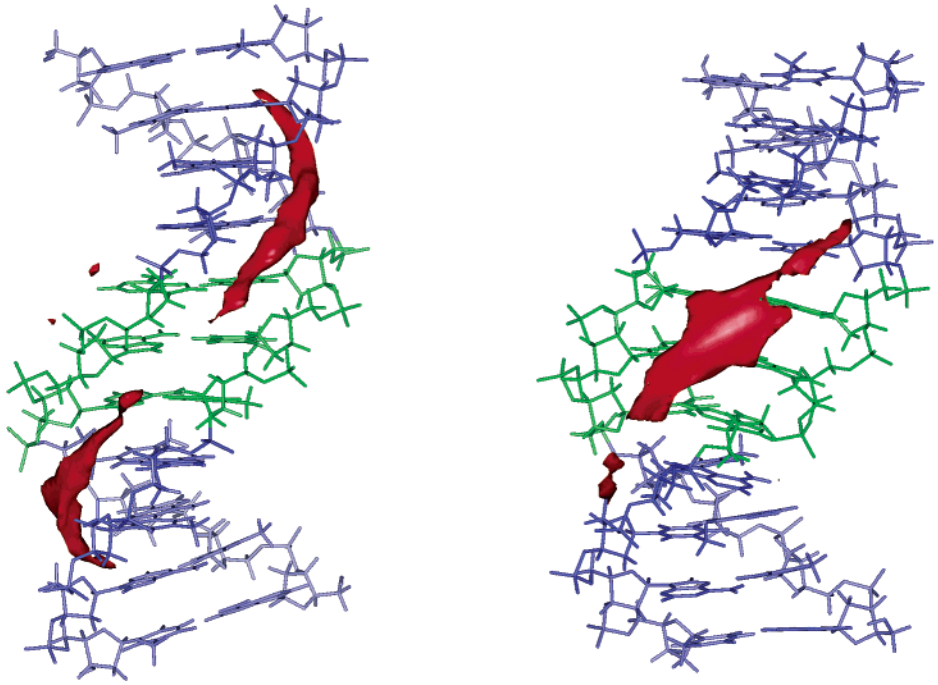


Figure 10. MIP (−5 kcal/mol) contour for the d(ATAT**AAA**ATAT) (left) and d(**ATAT**AAA**ATAT**) (right) structures. The apH portions are colored in green, and the WC portions are in blue. Bold means Hoogsteen pairings.

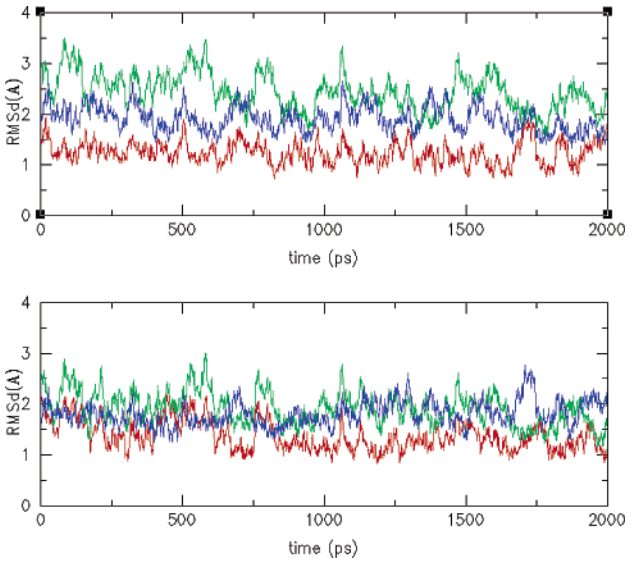


Figure 11. Backbone root-mean-square deviations (rmsd in Å) of the d(**ATAT**AAA**ATAT**) (top) and d(ATAT**AAA**ATAT) (bottom) trajectories with respect to the following: red, the MD-averaged structures; green, the B structure; and blue, the apH structure. Averaged rmsd values with respect to the average, B, and apH structures are 1.24, 2.44, and 1.91 Å for the d(**ATAT**AAA**ATAT**) structure, and 1.37, 1.91, and 1.80 Å for the d(ATAT**AAA**ATAT) structure. Bold means Hoogsteen pairings.

compatible in the same oligonucleotide, but that even when they are contiguous each helix fragment preserves its own structural and reactive integrity.

Conclusions

In summary, the preceding results show that the apH structure is a thermodynamic stable conformation, which might compete with the B conformation if the flexibility of the helix is altered. However, for the sequences studied, MD simulations favor the B form in dilute aqueous solution of lineal DNAs in the absence of cofactors. Interestingly, the structural, dynamical, and reactive properties of the apH helices are not extremely different from those of the B helix. However, the small changes should allow proteins and drugs to distinguish between these two helical models. MD simulations suggest that apH structures can coexist next to B-type structures, demonstrating again the extreme plasticity of DNA. If kinetic factors are not limiting (the syn ↔ anti barrier is around 6.2 kcal/mol[56]), AT-rich regions might adopt B and apH forms, and this possibility might be exploited to tune the interaction with specific proteins.

Acknowledgment. We are indebted to X.-J. Lu and W. Olson for a copy of X3DNA and help in the use of the code for a nonstandard structure. We also thank D. Bashford for a copy of his MEAD program. The Centre de Supercomputació de Catalunya (CESCA) and the Spanish Ministry of Science and Technology (SAF2002-4282, PM99-0046, and BIO2003-06848) are acknowledged for financial support.

JA035918F

(56) Rhodes, L. M.; Schimmel, P. R. *Biochemistry* **1971**, *10*, 4426.