Theoretical Study of the Hoogsteen–Watson-Crick Junctions in DNA

Elena Cubero,* F. Javier Luque,† and Modesto Orozco*‡
*Molecular Modelling & Bioinformatic Unit, Institut de Recerca Biomèdica-Parc Científic de Barcelona, Barcelona 08028, Spain; and †Departament de Fisicoquímica, Facultat de Farmacia, and ‡Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Barcelona 08028, Spain

ABSTRACT A series of d $(AT)_n$ oligonucleotides containing mixtures of normal B-type Watson-Crick and antiparallel Hoogsteen helices have been studied using molecular dynamics simulation techniques to analyze the structural and thermodynamic impact of the junction between Watson-Crick and antiparallel Hoogsteen structures. Analysis of molecular dynamics simulations strongly suggests that for all oligonucleotides studied the antiparallel Hoogsteen appears as a reasonable conformation, only slightly less stable than the canonical B-type Watson-Crick one. The junctions between the Watson-Crick and Hoogsteen structures introduces a priori a sharp discontinuity in the helix, because the properties of each type of conformation are very well preserved in the corresponding fragments. However, and quite counterintuitively, junctions do not largely distort the duplex in structural, dynamics or energetic terms. Our results strongly support the possibility that small fragments of antiparallel Hoogsteen duplex might be embedded into large fragments of B-type Watson-Crick helices, making possible protein-DNA interactions that are specific of the antiparallel Hoogsteen conformation.

INTRODUCTION

DNA is very polymorphic and can adapt its structure depending on the sequence, the presence of ligands, or changes in the environment (1–3). Repetitive sequences have an especially rich conformational space (1–4) showing a marked tendency to appear in left-handed duplexes, triplexes, and tetraplexes (4–8), or even duplexes with parallel arrangements of the complementary strands (2,9,10). Particularly, d (AT) sequences have been experimentally detected in: i), partly or largely distorted B-type conformation (11–13), ii), in the right-handed C (14) and D (15) forms, and iii), even left-handed models have been suggested (16). Recently, Subirana and co-workers (17,18) solved crystals of the d (AT) polymer where the duplex appeared as an antiparallel right-handed (apH) double helix with Hoogsteen d (A-T) pairings and all the adenosines in the *syn* conformation. The structure has not been yet directly verified in solution by NMR experiments, but recent molecular dynamics (MD) simulations (19) suggested that the apH conformation is not much with respect to the canonical Watson-Crick B-form (B helix) in aqueous solution. In fact, MD and molecular mechanics-Poisson Boltzmann/surface area (MM-PB/SA) calculations suggested that both B and apH duplexes are almost isoenergetic and only the greater rigidity of the latter explain the larger population of B helix in physiological environments (19).

Hoogsteen d (A-T) pairings are known to stabilize triads in parallel triplexes (3,4,20,21), as well as in parallel duplexes (9). Furthermore, Hoogsteen A-T pairings are commonly found inserted in the middle of B-type Watson-Crick duplexes when the DNA interacts with some drugs or proteins (22–26). Because d $(AT)_n$ regions have a key role in the control of gene expression (24,27,28), the formation of Hoogsteen islands in the middle of tracks of d $(AT)_n$ B-DNA might be involved in the existence of a subtle regulatory mechanism of gene function (19,22,24).

In a previous article (19) we analyzed the properties of pure apH duplexes made with tracts from 2 to 8 d (AT), i.e., duplexes from 4 to 16 mer. These calculations can provide evidence for the existence of conformation derived from crystal structures reported by Subirana's group (17,18) in solution. Our calculations also suggested that segments of apH helix might exist in conditions where the flexibility of the DNA is externally restricted. Preliminary calculations indicated that, at least for poly-D (A) sequences, the junctions between B and apH helices are not sterically hindered. In this article, we will analyze systematically the structure of chimeric duplexes formed by portions of B and apH helices in d (AT) tracks. For this purpose we have performed molecular dynamics simulations of 12-, 14-, and 16-mer duplexes containing each 2, 4, or 6 apH steps in the center of the helix and surrounded by segments of normal B helix. Results are compared with those previously obtained for pure B and apH duplexes of the same global length. The analysis of >40 ns of equilibrated MD simulation on the same structural motif allowed us to describe with detail the characteristics of the B/apH chimeras.

METHODS

To analyze the structure and stability of chimeras containing fragments of apH and B helices for d $(AT)_n$ sequences, we built starting structures for duplexes d $(AT)_n$ (n = 12, 14, and 16) using standard Arnott's parameters (29) for B-DNA (see a summary of simulations in Table 1). In the center of each structure, apH helices with length 2, 4, or 6 were inserted using geometrical data from crystal structure (17). The junction steps were

Submitted January 14, 2005, and accepted for publication October 20, 2005.

Address reprint requests to M. Orozco, E-mail: modesto@mmb.pcb.ub.es or F. J. Luque, E-mail: javier@far1.far.ub.es.

0006-3495/06/02/1000/09 $2.00

doi: 10.1529/biophysj.105.059535

TABLE 1 Scheme of MD simulations of pure helices (B and apH), and chimeras containing both types of helices

Helix	Sequence	Recognition mode	Length simulation	T (K)
B12	d $(AT)_6$	WC	10 ns	298
B12	d $(AT)_6$	WC	2 ns	310
B14*	d $(AT)_7$	WC	2 ns	298
B16*	d $(AT)_8$	WC	2 ns	298
apH2B10	d $(AT)_6$	WC-2H-WC	2 ns	298
apH2B12	d $(AT)_7$	WC-2H-WC	2 ns	298
apH2B14	d $(AT)_8$	WC-2H-WC	2 ns	298
apH4B8	d $(AT)_6$	WC-4H-WC	2 ns	298
apH4B10	d $(AT)_7$	WC-4H-WC	2 ns	298
apH4B12	d $(AT)_8$	WC-4H-WC	2 ns	298
apH6B6	d $(AT)_6$	WC-6H-WC	10 ns	298
			2 ns	310
apH6B8	d $(AT)_7$	WC-6H-WC	5 ns	298
apH6B10	d $(AT)_8$	WC-6H-WC	5 ns	298
apH12	d $(AT)_6$	H	10 ns	298
apH12	d $(AT)_6$	H	2 ns	310
apH14*	d $(AT)_7$	H	2 ns	298
apH16*	d $(AT)_8$	H	2 ns	298

All oligonucleotides are antiparallel duplexes and they are labeled by the sequence of one of the strands.
*Simulations taken from Cubero et al. (19).

optimized by restricted MD to avoid unreliable starting geometries. The nine different starting structures were neutralized by adding a suitable number of Na^+ ions and immersed in rectangular boxes containing between 3000 and 3600 water molecules. All of the systems were optimized, thermalized (298 K), and equilibrated for 200 ps using our standard multistage protocol (30–32). The equilibrated structures were then subjected to 2 and 5 ns (after the referee's suggestions some simulations were extended to 10 ns for control purposes; see Table 1) of MD simulation at constant temperature (298 K; after the referee's suggestions we performed three additional simulations at $T = 310$; see Table 1) and pressure (1 atm). Periodic boundary conditions were used to simulate a diluted environment, and the particle mesh Ewald method (33). 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 (35,36) and TIP3P (37) force fields were used to describe DNA and water. All MD simulations were performed using the AMBER6.0 suite of programs (38). This simulation protocol is identical to that used in our previous simulations (19), which allowed us to include in our comparison the B and apH duplexes ($n = 12$, 14, and 16) described there.

The free energy of every structure was estimated from Eq. 1 (9,19,39,40), where averages were performed for snapshots collected after the first nanosecond of the corresponding trajectory. The intramolecular term (E_{intra}) was computed using the standard AMBER-98 force field (35,36). The solvation contribution (G_{sol}) was determined using: i), finite difference Poisson-Boltzmann (PB) calculations as implemented in the MEAD program (41,42) for initial and final grid spacing of 1 and 0.4 Å, with external and internal dielectric constants of 80 and 1 (GB/SA solvation calculations (43) provide the same trends and accordingly results are not shown here ((19,39,44) for discussion). Intramolecular DNA entropy was determined using Schlitter's and Andricioaei-Karplus methods (45–47). As done in previous studies, end basepairs were removed from the analysis (9,19,39) to reduce noise in the calculations.

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

The free energy computed from Eq. 1 does not provide a direct measure of the stability of a structure, but for two oligonucleotides of the same sequence the difference in free energies gives a direct measure of the relative stability (see Eq. 2). Equivalent equations can be formulated for the entropic and effective energy contributions.

$$\Delta G^{X-Y}_{stab} = G^X - G^Y. \quad (2)$$

For a regular helix there is a linear relationship between the free energy (as computed in Eq. 1) and the total length of the oligonucleotide (see Eq. 3; (9,19,39)). This allowed us to obtain accurate estimates of the relative stability of two helical conformations by averaging data obtained from series of different MD simulations, thus reducing the statistical noise intrinsic to the use of Eq. 1.

$$G^X_n = nG^X_{elon} + G^X_{nuc}, \quad (3)$$

where G_{elon} and G_{nuc} represent free-energy contributions associated for the elongation and nucleation of the nucleotide, respectively, and n is the length of the oligonucleotide in conformation X.

Equation 3 cannot be directly used in these systems, because in principle the free energy of the chimera depends not only on the length of the oligonucleotide, but also on the relative content of each type of structure (B, apH) and the number of junctions. A reasonable extension of Eq. 3 can be derived assuming that the thermodynamics of each constituent structure of the chimera is independent of the other. This yields to Eq. 4, where the total oligonucleotide (containing p steps) is divided in n steps of X and m steps of Y conformations, and z junctions linking the two types of structures. In a first approximation, the effective nucleation free energy of the chimera ($\alpha_{x/y}$) can be interpolated from those of the pure oligonucleotides using Eq. 5. We should note that the nucleation free energies of B and apH helices are very similar (19), reducing the uncertainties derived from the use of Eq. 5. We should also notice that Eqs. 4 and 5 can be easily rewritten in terms of effective energies.

$$G^{X/Y}_p = nG^X_{elon} + mG^Y_{elon} + zG^{X/Y}_{jun} + \alpha_{X/Y} \quad (4)$$

$$\alpha_{X/Y} = \frac{nG^X_{nuc} + mG^Y_{nuc}}{p}. \quad (5)$$

Because the elongation and nucleation free energies of the B and apH helices are known from previous work (19) the free energy of junction can be easily derived using Eqs. 4 and 5 from the analysis of the free energies of the nine oligonucleotides of different lengths and compositions studied here.

Interaction properties of the structures were obtained from classical molecular interaction potential (cMIP) calculations for an O^+ probe molecule (30–32,48). Solvation was represented by integrating the water population along the trajectory (30–32). The molecular flexibility was analyzed using principal component analysis method (49). The eigenvalues $\{\lambda_i\}$ were manipulated (50) to derive information on the strength of the harmonic constant associated to the essential movements (Eq. 6). The eigenvectors (ν_i) were used to describe the nature of the essential deformation modes and to compare the pattern of flexibility of two trajectories using our standard algorithms ((39,50–52); Eqs. 7 and 8), or the recently developed metrics (53), which introduces a Boltzmann-weighting scheme in the absolute (Eq. 9) and relative (Eq. 10) comparison indexes.

$$K_i = kT/\lambda_i, \quad (6)$$

where k is Boltzmann's constant, T is the absolute temperature, and λ_i is the eigenvalue describing the essential movement i in $Å^2$.

$$\gamma_{AB} = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{n}(\nu^A_i \cdot \nu^B_j)^2, \quad (7)$$

where A and B stand for two different trajectories of equal length, ν stands for the eigenvectors, and n is the number of motions that account for a given value of structural variance in the trajectories (typically only 10 motions account for ~80% of the variance). Only backbone atoms (up to C1′) were considered in the comparison.

$$\kappa_{AB} = 2\frac{\gamma_{AB}}{(\gamma^T_{AA} + \gamma^T_{BB})}, \quad (8)$$

where self-similarity indexes for A and B trajectories (γ^{T}_{AA} and γ^{T}_{BB}) are 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 sample orthogonal movements.

$$\xi_{AB} = \frac{2\sum_{i=1}^{i=n}\sum_{j=1}^{j=n}\left[(\nu_i^A \cdot \nu_j^B)\frac{\exp\left\{-\frac{(\Delta x)^2}{\lambda_i^A}-\frac{(\Delta x)^2}{\lambda_j^B}\right\}}{\sum_{i=1}^{i=n}\exp\left\{-\frac{(\Delta x)^2}{\lambda_i^A}\right\}\sum_{j=1}^{j=n}\exp\left\{-\frac{(\Delta x)^2}{\lambda_j^B}\right\}}\right]^2}{\sum_{i=1}^{i=n}\left(\frac{\exp\left\{-2\frac{(\Delta x)^2}{\lambda_i^A}\right\}}{\left(\sum_{i=1}^{i=n}\exp\left\{-\frac{(\Delta x)^2}{\lambda_i^A}\right\}\right)^2}\right)^2 + \sum_{j=1}^{j=n}\left(\frac{\exp\left\{-2\frac{(\Delta x)^2}{\lambda_j^B}\right\}}{\left(\sum_{j=1}^{j=n}\exp\left\{-\frac{(\Delta x)^2}{\lambda_j^B}\right\}\right)^2}\right)^2} \quad (9)$$

$$\delta_{AB} = 2\frac{\xi_{AB}}{(\xi^{T}_{AA} + \xi^{T}_{BB})}, \quad (10)$$

where λ_i is the eigenvalue (in Å^2) associated to eigenvector ν_i and Δx is a displacement (in Å) common to all the modes, that is selected (J. R. Blas, A. Pérez, F. J. Luque, and M. Orozco, unpublished data) as the minimum one that makes negligible to the calculation of ξ (Eq. 9) of eigenvectors (i, j) associated to high frequencies (those that are not needed to explain 99% of the variance). For coherence with the other indexes the sums are extended to the same active space as in Eqs. 9 and 10. Note that when $\Delta x = 0$, Eqs. 9 and 10 converge to Eqs. 7 and 8.

Geometrical analysis was performed using the *ptraj* module in AMBER6.0 (38) and in-house programs, and helical analysis was carried out using Olson's X3DNA program (54) and Curves5.3 (55). Basepairs at the ends of the duplex were removed in all the cases.

RESULTS AND DISCUSSION

Stability of trajectories of B/apH chimeras

The MD trajectories for the nine chimeras studied here (Table 1) are stable in the simulation time (2–5 ns) considered and the same is found for selected simulations extended to 10 ns or performed at 310 K. In all cases the root mean-square deviations (RMSDs) determined with regard to the MD-averaged structures are small, with values intermediate between those of B and apH helices (see Table 2), thus indicating that even the 2-ns trajectories are well converged. Interestingly, the content of B and apH duplex in the chimeras does not introduce major alterations in the RMSD with respect to the MD-averaged conformation (Table 2). The RMSD with respect to the starting structures (built from experimental data of the constituent fragments; see Table 2) are small (~2 Å in all the cases), and intermediate between those found for B and apH helices. In summary, the RMSD data suggest that the structures of chimeras seem to be very stable under these simulation conditions. Moreover, such structural stability is not affected by the length of the oligonucleotide nor the ratio between B and apH helices. The presence of junction, therefore, does not seem to introduce any dramatic effect on the structure of the duplex.

The hydrogen-bond (H-bond) pattern for Watson-Crick and Hoogsteen pairings is well preserved for all the chimeras. The percentage of total hydrogen bonding is >93% in all cases (Table 3), as also found in B and apH helices (19), thus indicating that mixing of the two helices in the same structure does not alter the hydrogen bonding pattern. The hydrogen bonding at the junctions is also well preserved, the percentage of hydrogen bonds being similar to that found for the rest of the duplex. As found in our previous study (19), partial breathing movements were detected in the subnanosecond timescale for both Watson-Crick and Hoogsteen pairings. In general these movements last for ~0.5 ns and imply interchanges of hydrogen-bond donors and acceptors between adenine and thymine. For one of the trajectories (apH6B10; see Table 3), the ''partial opening'' movement lasts >2 ns and implies that at one Hoogsteen basepair of the junction the N6 (A) → O4 (T) and N3 (T) → N7 (A) interactions are lost while one N6 (A) → O2 (T) hydrogen bond is formed leading to an overall reduction in the percentage of H-bonding detected during the entire trajectory (see Table 3). These ''partial opening'' events, nevertheless, do not introduce any major alteration in the general geometry of the helix.

In summary, all the chimeras seem to be as stable as the parent B and apH helices. No major alterations in the global structure nor in the specific pattern of basepair interactions are detected in the trajectories related to the existence of chimeras of two different helices in a common duplex.

TABLE 2 All-atom root mean square deviations (Å) for duplexes containing B or apH helices (data from Cubero et al. (19)) or B/apH chimeras with respect to the MD-averaged and starting structures

Structure, 298 K	Average*	Starting structure
apH2B10	1.4 ± 0.3	1.7 ± 0.3†
apH2B12	1.6 ± 0.3	1.9 ± 0.3†
apH2B14	1.9 ± 0.6	2.1 ± 0.5†
apH4B8	1.4 ± 0.3	1.8 ± 0.4†
apH4B10	1.6 ± 0.3	1.9 ± 0.3†
apH4B12	2.0 ± 0.7	2.2 ± 0.5†
apH6B6	1.6 ± 0.4	2.1 ± 0.5†
apH6B8	1.8 ± 0.3	2.2 ± 0.4†
apH6B10	1.8 ± 0.4	2.0 ± 0.4†
B12	1.7 ± 0.4	3.4 ± 0.7‡
B14	1.7 ± 0.3	3.4 ± 0.6‡
B16	2.1 ± 0.6	3.9 ± 0.9‡
apH12	1.3 ± 0.3	1.4 ± 0.3§
apH14	1.5 ± 0.4	1.5 ± 0.3§
apH16	1.3 ± 0.3	1.6 ± 0.4§
Structure, 310 K	**Average***	**Starting structure**
apH6B6	1.6 ± 0.3	1.8 ± 0.4
B12	1.8 ± 0.3	3.6 ± 0.7
apH12	1.3 ± 0.2	1.4 ± 0.2

Standard deviations (mean ± SD) in the averages (Å) are also displayed.
*MD-averaged structures obtained using the last 1 or 4 ns of trajectories.
†Structures built by inserting apH sequences in the central steps of B-DNA followed by restricted minimization.
‡Arnott's structures. See Arnott and Hukins (29).
§Crystal structure built by joining two 5-mer structures in Protein Data Bank entry 1gpu followed by restricted minimization.

TABLE 3 Occurrence (in % with respect to the maximum number of hydrogen-bond interactions) of canonical Watson-Crick and Hoogsteen hydrogen bonds for the different trajectories

Structure	HB Watson-Crick	HB Hoogsteen	HB total
apH2B10	96.8 / *96.3*	– / *96.3*	96.6
apH2B12	97.1 / *96.8*	– / *96.0*	96.9
apH2B14	96.9 / *97.3*	– / *96.8*	96.9
apH4B8	97.0 / *97.5*	97.0 / *97.5*	97.2
apH4B10	97.5 / *97.3*	93.8 / *98.0*	96.9
apH4B12	91.1 / *97.3*	96.8 / *98.0*	93.8
apH6B6	98.0 / *97.3*	96.4 / *97.5*	97.1
apH6B8	96.8 / *97.8*	96.8 / *97.0*	97.0
apH6B10	96.3 / *97.8*	96.6 / *71.0*	93.0
B12	97.1 / –	–	97.1
B14	97.0 / –	–	97.0
B16	96.0 / –	–	96.0
apH12	–	95.5	95.5
apH14	–	96.8	96.8
apH16	–	97.3	97.3

Values in Roman correspond to all nucleobases except those at the ends of the duplex. Values in italics correspond to nucleobases at junctions (i.e., those placed in the boundaries between B and apH fragments). Data from B or apH helices are taken from Cubero et al. (19).

Structural description of the chimeras

The general shape of the chimeras is very similar to that of B and apH helices (see Fig. 1). The different ratio of B and apH helices or the presence of B/apH junctions does not alter the general shape nor the helical parameters of the duplex. The periodicity and rise of the apH helix is similar to that of B-DNA (17–19) and do not change much in the different chimeras (see Table 4). The fact that phase angles for sugars in apH helices are typically smaller (5–7°) than those found in B-DNA is also visible in chimeras, because the average phase angle generally decreases as the content of apH helix increases. However, all the sugars are found in the South-Southeast region. All the structures (see Fig. 1 and Table 4) have a groove topology similar to that of B-DNA, but the average size of the minor groove decreases as the percentage of apH helix increases, as expected from the fact that apH helices show in general smaller widths of the minor groove than B-DNAs.

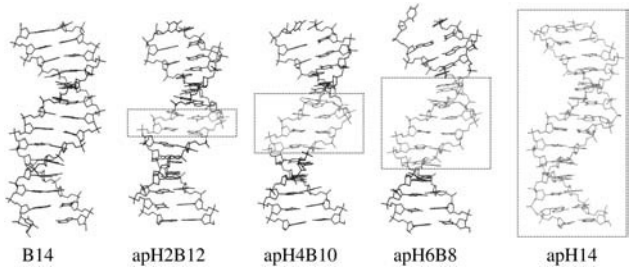

FIGURE 1 MD-averaged structure of 14-mer DNA duplexes containing (*from left to right*) none, two, four, six, and 14 antiparallel Hoogsteen (apH) steps. The B fragment always appears in black and the apH one in gray, and framed by the dotted box.

Very interestingly, the helical parameters describing B and apH helices are well maintained in the respective fragments (see Table 4). For example, the RMSD between the central 6-mer apH fragment in apH12, apH6B6, apH6B8, and apH6B10 varies between 1.0 and 1.1 Å. The maintenance of the structural characteristics of B and apH helices in the chimeric duplexes is especially clear in the shortest intrastrand C1′–C1′ distance, which is the best single parameter to discriminate between B and apH helices. Thus, this parameter adopts the values expected for the canonical B and apH helices in the corresponding chimeric fragments. The junctions are characterized by a reduced twist and a slightly increased rise with respect to B and apH helices (Table 4). Apart from these changes, they do not appear to be structurally stressed nor produce any noticeable change in the global helical parameters of the duplex.

The three-dimensional distribution of classical molecular interaction potential (CMIP (48)) in chimeras is intermediate between those found in both B and apH helices (see Fig. 2). For B-DNA there is a strong region of favorable interaction with small cations along the minor groove. The size of such a region becomes progressively diminished as the content of apH helix in the duplex increases. All the helices are well solvated by 25–27 waters per step (water molecules at <3.5 Å from any heteroatom of DNA), with no major difference for B, apH, or chimeric duplexes. As described elsewhere (19), despite the different orientation of the adenine, the characteristic spine of hydration in the minor groove of B-DNA (56) is also clear in the apH helix and in all the chimeras (see Fig. 2). Furthermore, the specific hydration pattern of apH helices in the major groove (17–19) is found also here even for the smallest apH helical fragments (see Fig. 3 and compare with Fig. 5 in Cubero et al. (19)). In summary, there is a strong memory of the intrinsic interaction properties of B and apH helices in the chimeras. Accordingly, we suggest that even small apH fragments inserted in large B-DNA duplexes will display recognition properties very close to those of an apH helix, thus opening the possibility to expand (for a common sequence) the possibilities for specific DNA-protein recognition in the cell.

Dynamics properties of chimeras

The essential dynamics of B and apH duplexes and all the chimeras is dominated by bending and twisting deformation of the helices. Analysis of the 5-ns trajectories collected for 12-mer duplexes shows a large degree of similarity between pure duplexes and chimeras, as noted in relative similarity indexes (κ and δ; see Eqs. 8 and 10) around 90% (see Table 5). Comparison of the nature of the essential movements for the 14- and 16-mer oligonucleotides requires some caution, because deformation modes might not be completely converged after 2-ns trajectories. With this caution in mind, we must notice that the results support the occurrence of similar similarity essential movements in B, apH, and chimeric

TABLE 4 Average values for selected helical parameters obtained in MD simulations of B and apH duplexes as well as for different chimeras of the 14-mer oligonucleotides

	B14	apH2B12	apH4B10	apH6B8	apH14
Minimum C1′–C1′ distance	10.6	10.3 10.5 / *9.1*	9.9 10.5 / *8.6*	9.6 10.5 / *8.7*	*8.6*
Displacement	4.3	4.9 4.8 / *5.4*	5.5 5.4 / *5.7*	5.3 5.1 / *5.6*	*5.8*
Inclination	4.7	5.5 6.3 / *1.8*	5.3 6.9 / *2.2*	5.3 7.00 / *3.6*	*3.3*
Helical twist	33.5	32.7 31.9 / *37.7* / **33.6**	31.7 31.2 / *35.8* / **27.0**	32.20 32.5 / *33.9* / **27.3**	*31.4*
Helical rise	3.2	3.2 3.2 / *3.1* / **3.0**	3.2 3.2 / *3.1* / **3.3**	3.3 3.2 / *3.1* / **3.6**	*3.3*
Phase	139.6	137.1 139.2 / *126.7*	132.2 141.9 / *112.7*	131.8 141.6 / *122.4*	*125.8*
Amplitude	38.7	38.9 38.6 / *40.6*	39.4 38.5 / *41.1*	40.0 38.6 / *41.5*	*41.8*
Minor groove	11.8	11.5	10.8	10.4	*10.0*
Major groove	19.4	18.6	21.3	21.4	*21.5*

For chimeras, values corresponding to the average for all the helix (top) and for each fragment ((bottom) B, Roman; apH, italic; junction, bold) are shown. Distances are in angstroms and angles in degrees.

duplexes (see data in Table 5 for the 16-mer duplexes). At this point, it is worth noting the similarity between normal (κ) and Boltzmann's-weighted (δ) relative similarity indexes, which indicates that not only the nature of the essential movements, but also their relative contribution to the structural variance, is retained in the different duplexes. Overall, the presence of junctions and apH pairs does not seem to induce major changes in the natural deformation modes of the duplex.

The essential movements of B and apH duplexes and chimeras are associated to very weak ($<$10 cal/mol Å^2) harmonic force constants (see Eq. 6 and Fig. 4), which allow large deformations in the helices under room temperature. Interestingly, the increase in the percentage of apH steps leads to a parallel increase in all the harmonic constants (see Fig. 5). This finding suggests that the nature of the easiest deformation modes is similar for apH and B helices, but that deformation along these modes is more difficult for apH than for B fragments. Thus, B/apH chimeras will display an intermediate deformation pattern, but they will become more rigid as the apH content increases.

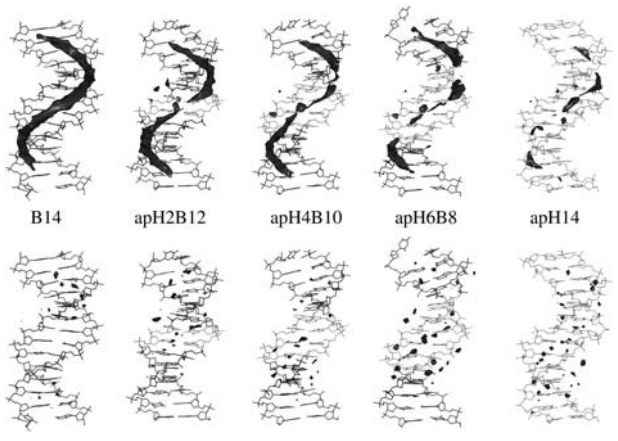

FIGURE 2 (*Top*) Classical molecular interaction potential (isocontour at −5 kcal/mol) using an O^+ probe particle (see Methods). (*Bottom*) Solvation maps (isocontour for an apparent water density of 2.5 g/cm^3) for the different 14-mer duplexes studied here.

Stability of B/apH chimeras

The effective energy (Eq. 1) of the chimeras is mostly dominated by the total length of the duplex and neither by the B/apH ratio nor the position of the junctions (see Table 6). The effective energies, as well as their intramolecular and solvation free-energy components, for B and apH helices and for the different chimeras are very similar (the same result was obtained when GB/SA was used to compute the solvation term).

The statistical analysis (see Methods and Eqs. 4 and 5) shows that the effective energy penalty due to the existence

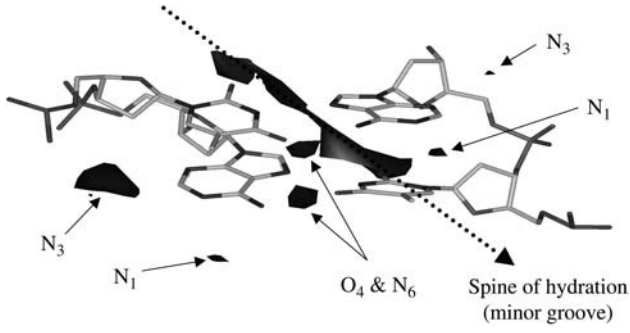

FIGURE 3 Detail of the solvation along the grooves in the apH fragment of the apH2B12 duplex. The isocontour corresponds to an apparent water density of 2.5 g/cm^3.

TABLE 5 Comparison indexes for 2-ns trajectories for 16-mer duplexes and 5-ns trajectories for 12-mer duplexes

2-ns trajectories for 16-mer duplexes	apH2B14	apH4B12	apH6B10	apH16
B16	0.63 / 0.95 0.77 / 0.97	0.60 / 0.89 0.77 / 0.95	0.62 / 0.87 0.76 / 0.93	0.61 / 0.85 0.75 / 0.89
apH2B14		0.65 / 0.93 0.79 / 0.96	0.65 / 0.88 0.78 / 0.95	0.64 / 0.86 0.77 / 0.91
apH4B12			0.68 / 0.91 0.81 / 0.96	0.66 / 0.87 0.80 / 0.92
apH6B10				0.71 / 0.89 0.81 / 0.93

5-ns trajectories for 12-mer duplexes	apH6B6	apH12
B12	0.57 / 0.81 0.68 / 0.89	0.63 / 0.93 0.75 / 0.94
apH6B6		0.59 / 0.84 0.71 / 0.88

Each cell contains (top) the normal absolute and relative similarity indexes (γ / κ) and (bottom) the Boltzmann's-weighted indexes (ξ / δ; see text for details).

of two junctions in our B/apH helices (see Eq. 5) is only 3.4 ± 1.0 kcal/mol, that is, 1.7 kcal/mol per junction, which is <0.05% the total effective energy of the duplexes. Such a penalty is even reduced (0.7 kcal/mol × junction) when solvation free energies determined from GB/SA calculations are considered. In other words, the small destabilization (in terms of effective energy) required to insert an apH fragment into a B duplex appears to have little impact on the whole stability of a DNA duplex under physiological conditions.

The presence of *syn*-adenosines destabilizes the apH helix (~0.7 kcal/mol × step; see Cubero et al. (19) and Stolarski et al. (57), but this is mostly compensated by hydrogen bonding, which is expected to be more stable in the Hoogsteen pairing than in the Watson-Crick scheme (19,58,59). Thus, the analysis of all these trajectories confirms that on average the Hoogsteen H-bonds are around −1.2 ± 0.01 kcal/mol (see Table 7) more stable than the Watson-Crick ones, in good agreement with previous estimates (19). The Watson-Crick or Hoogsteen hydrogen bonds are equally stable in the junction or in the middle of the fragments (data not shown), and

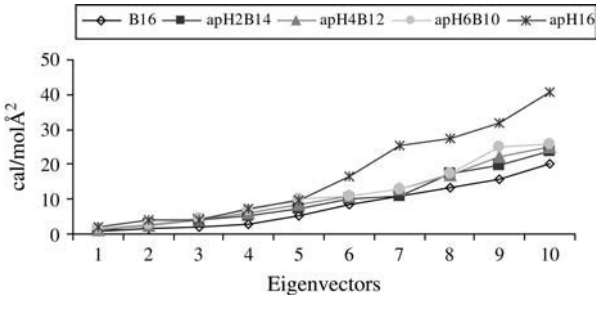

FIGURE 4 Harmonic force constants associated to deformations along the 10 first essential movements of B and apH duplexes and different chimeras (for 16-mer duplexes).

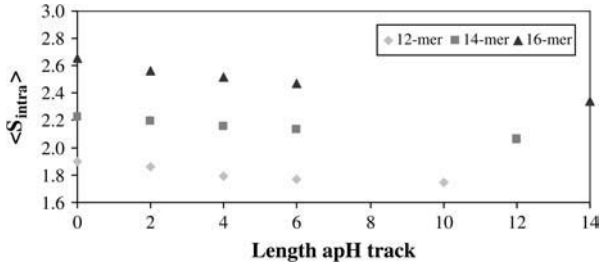

FIGURE 5 Dependence of the intramolecular entropy with the length of the apH segment in different duplexes. Entropy values (in cal/Kmol) were obtained using the Andricioaei-Karplus method, but Schlitter's approach yields equivalent results.

accordingly, the hydrogen-bond energy associated to a given junction is just the average between those of Watson-Crick and Hoogsteen pairs. The stacking energy of (A·T)-(T·A) dimers is the same irrespective of whether the hydrogen bond (the ''·'' here) pattern is Watson-Crick or Hoogsteen (Table 7). On the contrary, the stacking (T·A)-(A·T) is on average 3.8 kcal/mol more stable if the two dimers are hydrogen-bonded following the Hoogsteen scheme (Table 7). Despite the low twist values (Table 4), the stacking in the junction is quite favorable. Thus, for the (A·T)-(T·A) dimers there is only a loss of 0.5 kcal/mol with respect to B and apH stackings, whereas for the (T·A)-(A·T) dimers there is a loss of stability (1.3 kcal/mol) relative to apH stacking, but a gain (2.5 kcal/mol) with regard to B stacking. In summary, the d $(AT)_n$ DNA seems so flexible that the distortions needed to accommodate a junction between two different types of helices do not introduce any major destabilization in the structure. As noted in Supplementary Material, the conclusions on stability derived here from 2- to 5-ns trajectories remain valid when tested in 10-ns trajectories, or when more physiological temperatures (310 K) are considered.

Previous calculations (19) suggested that whereas in terms of effective energy B and apH helices were equally stable, the former was favored by entropic reasons. As noted elsewhere (39,60,61), 2–5-ns simulations are in general too short to provide reliable measures of relative intramolecular entropy of different helices, which precludes the calculation of the total free energy using Eq. 1 and derived equations. However, for each oligonucleotide, comparison of the five estimates (B, apH, and the three chimeras) of the intramolecular entropy can be useful to identify general trends about the impact of chimeras on the entropy of the helix (19). It is also worth noting that entropy values obtained for selected cases from 10-ns trajectories are obviously larger, but in relative terms (apH, B, junctions) no significant change is found, confirming the validity of our conclusions (see Table 8). Chimeras have intramolecular entropies intermediate between those of B and apH helices (Table 8), and the increase in the content of apH helix in duplexes of constant length produces a linear (r^2 ranging between 1.00 and 0.96) decrease in the entropy of the system, as shown in

TABLE 6 Effective energy ($\langle E_{effec} \rangle$; bold), their solvation ($\langle G_{sol} \rangle$; Roman), and intramolecular ($\langle E_{intra} \rangle$; italics) contributions for the B, apH, and chimeric duplexes

Length-2	B	apH2	apH4	apH6	apH
10	**−3548.9 ± 3.0**	**−3541.6 ± 2.5**	**−3546.4 ± 2.9**	**−3541.0 ± 3.0**	**−3548.9 ± 2.6**
	−5681.4 ± 2.9	−5674.8 ± 2.3	−5699.3 ± 2.7	−5717.8 ± 2.9	−5681.5 ± 2.4
	2132.5 ± 1.0	*2133.2 ± 0.8*	*2152.9 ± 0.9*	*2176.8 ± 0.9*	*2132.6 ± 0.8*
12	**−4269.9 ± 3.3**	**−4268.9 ± 4.2**	**−4271.1 ± 3.5**	**−4267.5 ± 4.1**	**−4276.5 ±3.4**
	−7453.7 ± 3.2	−7464.7 ± 4.0	−7416.6 ± 3.3	−7447.1 ± 3.9	−7387.8 ± 4.1
	3183.8 ± 1.0	*3195.8 ± 1.3*	*3145.5 ± 1.1*	*3179.6 ± 1.3*	*3111.3 ± 1.3*
14	**−4999.7 ± 4.7**	**−4992.9 ± 6.1**	**−4995.2 ± 4.4**	**−4987.7 ± 4.5**	**−5001.5 ± 3.6**
	−9350.3 ± 4.5	−9328.0 ± 5.8	−9340.1 ± 4.2	−9307.1 ± 4.3	−9297.1 ± 3.4
	4350.6 ± 1.4	*4335.1 ± 1.8*	*4344.9 ± 1.4*	*4319.5 ± 1.4*	*4295.6 ± 1.1*

Values and their mean ± SE are in kcal/mol. Energies are always computed removing bases at the ends, which means that the effective length of the duplex is obtained by subtracting 2 to the total number of base steps (length-2).

Fig. 5, thus reflecting the larger rigidity of apH helices (19). The fact that entropies of chimeras and pure helices (B and apH) can be fitted in the same regression line (see Fig. 5) demonstrates that the junctions does not introduce any major alteration in the global intramolecular entropy of the helix, and that no entropy penalty can be specifically assigned to the junctions.

CONCLUSIONS

B/apH chimeras appear stable under simulation conditions. The global structural, dynamical, and molecular recognition properties of chimeras depend linearly on the percentage of apH and B helices. The structural transition from B to apH helices (and vice versa) is sharp (limited to one step) and happens without any dramatic alteration in the structure of the helical fragments. A certain local unwinding is the only remarkable conformational change detected at the junctions.

The fragments of B and apH helices in the chimera have a very strong memory of the structure, flexibility, and molecular interaction properties of the corresponding pure helices. Interestingly, such a memory does not depend on the length of the fragment, and even a very short apH fragment suffices to define local properties close to those of an apH helix.

B and apH helices possess similar effective energy. The B form is however entropically favored. The stability of the chimeras is very similar to that of the constituting helical fragments, without any dramatic lost of stability related to the presence of the junctions.

Overall, our simulations strongly suggest that, beside kinetic factors related to the *anti-syn* rotation of adenosine, the formation of short apH helices in the middle of long B-type duplexes might be possible if external factors (i.e., by interactions with proteins or drugs) reduce the flexibility of the duplex, thus compensating the entropic preference for the B form. This suggestion, in conjunction with the ability of small apH fragments to maintain a memory of their structural, dynamical, and reactive properties, support the possibility that apH fragments might play a greater than expected role in the control of gene expression.

TABLE 7 Average hydrogen-bond energy (kcal/mol) in Hoogsteen and Watson-Crick pairings in chimeras, B, and apH helices; stacking energy (kcal/mol) for the different steps and its intra- and intermolecular components

Hydrogen-bond energy			$\langle E_{HBond} \rangle$
Watson-Crick			−12.1
Hoogsteen			−13.3
Stacking energy	$\langle E_{Stack,\ intra} \rangle$	$\langle E_{Stack,\ inter} \rangle$	$\langle E_{Stack,\ TOT} \rangle$
T*·*A			
A·T	−17.9	−1.7	−19.6
T*·*A			
A*·*T	−21.3	0.4	−20.9
T·A			
A·T	−15.0	−2.2	−17.1
A*·*T			
T·A	−14.5	−1.0	−15.5
A*·*T			
T*·*A	−11.9	−4.1	−16.0
A·T			
T·A	−16.5	0.5	−16.0

Roman characters mean Watson-Crick, and italics and bold characters mean Hoogsteen pairs. The dot means H-bonded pairs. The total intrastrand contribution to stacking stands for stacking interactions between the nucleobases in the same strand, whereas the interstrand contribution represent the cross-stacking interactions between nucleobases in the two strands. Values were obtained by averaging data from the entire sets of trajectories and the mean ± SE associated to averages is always <0.1 kcal/mol.

TABLE 8 Entropy estimates (cal/Kmol) computed using Andricioaei & Karplus (in Roman) and Schlitter's (italics) methods for 2-ns simulations (values obtained for 10-ns trajectories are displayed for reference when available)

Length-2	B	apH2	apH4	apH6	apH
10	1.897 (2.072)	1.864	1.790	1.772 (1.979)	1.745 (1.885)
	2.052 (2.263)	*2.017*	*1.943*	*1.924(2.168)*	*1.898 (2.062)*
12	2.221	2.195	2.157	2.127	2.065
	2.395	*2.368*	*2.330*	*2.302*	*2.237*
14	2.652	2.558	2.513	2.473	2.338
	2.842	*2.747*	*2.702*	*2.663*	*2.525*

Intramolecular entropy values for B and apH helices are taken from Cubero et al. (19).

Supplementary material is available upon request to the corresponding authors.

We thank Prof. J. Subirana for many helpful discussions. 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.

We acknowledge the Centre de Supercomputació de Catalunya (CESCA) and the Spanish Ministry of Science and Technology (SAF2002-4282 and BIO2003-06848), and the Fundación BBVA for financial support.

REFERENCES

1. Saenger, W. 1984. Principles of Nucleic Acid Structure. Springer-Verlag, New York, NY.
2. Bloomfield, V. A., D. N. Crothers, and I. Tinoco. 2000. Nucleic Acids: Structures, Properties and Functions. V. A. Bloomfield, D. N. Crothers, and I. Tinoco, editors. University Science Books, Sausalito, CA.
3. Blackburn, G. M., and M. J. Gait. 1990. Nucleic Acids in Chemistry and Biology. G. M. Blackburn and M. J. Gait, editors. IRL Press, Oxford, UK.
4. Robles, J., A. Grandas, E. Pedroso, F. J. Luque, R. Eritja, and M. Orozco. 2002. Nucleic acid triple helices: stability effects of nucleobase modifications. *Curr. Org. Chem.* 6:1333–1368.
5. Gehring, K., J. L. Leroy, and M. Guéron. 1993. A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature*. 363: 561–565.
6. Gray, D. M., M. Vaughn, R. L. Ratliff, and F. N. Hayes. 1980. Circular dichroism spectra show that repeating dinucleotide DNAs may form helices in which every other base is looped out. *Nucleic Acids Res*. 8: 3695–3707.
7. Laughlan, G., A. I. Murchie, M. H. Norman, P. C. Moody, D. M. Lilley, and B. Luisi. 1994. The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science*. 265:520–524.
8. Wang, A. H., G. J. Quiqkey, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, and A. Rich. 1979. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature*. 282:680–686.
9. Cubero, E., F. J. Luque, and M. Orozco. 2001. Theoretical studies of d (A:T)-based parallel-stranded DNA duplexes. *J. Am. Chem. Soc.* 123: 12018–12025.
10. Cubero, E., A. Aviñó, B. G. de la Torre, M. Frieden, R. Eritja, F. J. Luque, C. González, and M. Orozco. 2002. Hoogsteen-based parallel-stranded duplexes of DNA. Effect of 8-amino-purine derivatives. *J. Am. Chem. Soc*. 124:3133–3142.
11. Guzikevich-Guerstein, G., and Z. Shakked. 1996. A novel form of the DNA double helix imposed on the TATA-box by the TATA-binding protein. *Nat. Struct. Biol*. 3:32–37.
12. Viswamitra, M. A., P. Kennard, P. G. Jones, G. M. Sheldrick, S. Salisbury, L. Falvello, and Z. Shakked. 1978. DNA double helical fragment at atomic resolution. *Nature*. 273:687–688.
13. Viswamitra, M. A., Z. Shakked, P. G. Jones, G. M. Sheldrick, S. A. Salisbury, and O. Kennard. 1982. Structure of the deoxytetranucleotide d-pApTpApT and a sequence-dependent model for poly (dA-dT). *Biopolymers*. 21:513–533.
14. Brahms, S., J. Brahms, and K. E. Van Holde. 1976. Nature of conformational changes in poly [d (A-T) · d (A-T)] in the premelting region. *Proc. Natl. Acad. Sci. USA*. 73:3453–3457.
15. Davis, D. R., and R. L. Baldwin. 1963. X-ray studies on two synthetic DNA copolymers. *J. Mol. Biol*. 6:251–255.
16. Radwan, M. M., and H. R. Wilson. 1982. Fibre and molecular structure of thymidylyl-3′, 5′-deoxyadenosine. *Int. J. Biol. Macromol*. 4:145–149.
17. Abrescia, N. G. A., A. Thompson, T. H. Dinh, and J. A. Subirana. 2002. Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing. *Proc. Natl. Acad. Sci. USA*. 99:2806–2811.
18. Abrescia, N. G. A., C. González, C. Gouyette, and J. A. Subirana. 2004. X-ray and NMR studies of the DNA oligomer d (ATATAT): Hoogsteen base pairing in duplex DNA. *Biochemistry*. 43:4092–4100.
19. Cubero, E., N. G. A. Abrescia, J. A. Subirana, F. J. Luque, and M. Orozco. 2003. Theoretical study of a new DNA structure: the anti-parallel Hoogsteen duplex. *J. Am. Chem. Soc*. 125:14603–14612.
20. Felsenfeld, G., D. R. Davis, and A. Rich. 1957. Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc*. 79:2023–2024.
21. Pauling, L., and R. B. Corey. 1953. A proposed structure for the nucleic acids. *Proc. Natl. Acad. Sci. USA*. 39:84–97.
22. Aishima, J., R. K. Gitti, J. E. Noah, H. H. Gan, T. Schlick, and C. Wolberger. 2002. A Hoogsteen base pair embedded in undistorted B-DNA. *Nucleic Acids Res*. 30:5244–5252.
23. Gilbert, D. E., G. A. van der Marel, J. H. van Boom, and H. Feigon. 1989. Unstable Hoogsteen base pairs adjacent to echinomycin binding sites within a DNA duplex. *Proc. Natl. Acad. Sci. USA*. 86:3006–3010.
24. Patikoglou, G. A., J. L. Kim, L. Sun, S.-H. Yang, T. Kodadek, and S. K. Burley. 1999. TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev*. 13: 3217–3230.
25. Rice, P. A., S. Yang, K. Mizuuchi, and H. A. Nash. 1996. Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell*. 87:1295–1306.
26. Wang, A. H., G. Ughetto, G. J. Quigley, T. Hakoshima, G. A. van der Marel, J. H. van Boom, and A. Rich. 1984. The molecular structure of a DNA-triostin A complex. *Science*. 225:1115–1121.
27. Moreau, J., L. Marcaud, F. Maschat, J. Kejzlarova-Lepesant, J. A. Lepesant, and K. Scherer. 1982. A + T-rich linkers define functional domains in eukaryotic DNA. *Nature*. 295:260–262.
28. Oosumi, T., B. Garlick, and W. R. Belknap. 1995. Identification and characterization of putative transposable DNA elements in solanaceous plants and caenorhabditis elegans. *Proc. Natl. Acad. Sci. USA*. 92: 8886–8890.
29. Arnott, S., and D. W. L. Hukins. 1972. Optimized parameters for A-DNA and B-DNA. *Biochem. Biophys. Res. Commun*. 47:1504–1509.
30. Shields, G. C., C. A. Laughton, and M. Orozco. 1997. Molecular dynamics simulations of the d (T·A·T) triple helix. *J. Am. Chem. Soc*. 119:7463–7469.
31. Shields, G. C., C. A. Laughton, and M. Orozco. 1998. Molecular dynamics simulation of a PNA·DNA·PNA triple helix in aqueous solution. *J. Am. Chem. Soc*. 120:5895–5904.
32. Soliva, R., C. A. Laughton, F. J. Luque, and M. Orozco. 1998. Molecular dynamics simulations in aqueous solution of triple helices containing d (G·C·C) trios. *J. Am. Chem. Soc*. 120:11226–11233.
33. Darden, T. A., D. M. York, and L. G. Pedersen. 1993. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys*. 98:10089–10092.
34. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys*. 23: 327–341.
35. Cheatham, T. E., P. Cieplak, and P. A. Kollman. 1999. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn*. 16:845–862.
36. Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Fergurson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc*. 117: 5179–5197.
37. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. 79:926–935.
38. Case, D. A., D. A. Pearlman, J. W. Caldwell, T. E. Cheatham, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, et al. 1999. AMBER 6.0. University of California, San Francisco, CA.

39. Orozco, M., A. Perez, A. Noy, and F. J. Luque. 2003. Theoretical methods for the simulation of nucleic acids. *Chem. Soc. Rev.* 32: 350–364.

40. Srinivasan, J., T. E. Cheatham, P. Cieplak, P. A. Kollman, and D. A. Case. 1998. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* 120: 9401–9409.

41. Bashford, D., and K. Gerwert. 1992. Electrostatic calculations of the pKa values of ionizable groups in bacteriorhodopsin. *J. Mol. Biol.* 224: 473–486.

42. Bashford, D. 1997. An object-oriented programming suite for electrostatic effects in biological molecules. *In* Scientific Computing in Object-Oriented Parallel Environments. Y. Ishikawa, R. Rodney, R. R. Oldehoeft, J. V. M. Reynders, and M. Tholburn, editors. Lecture Notes in Computer Science. Springer, Berlin, Germany. 1343:233–240.

43. Still, W. C., A. Tempcyck, R. C. Hawley, and T. Hendrickson. 1990. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 112:6127–6129.

44. Orozco, M., and F. J. Luque. 2000. Theoretical methods for the description of the solvent effect in biomolecular systems. *Chem. Rev.* 100:4187–4226.

45. Andricioaei, I., and M. Karplus. 2001. On the calculation of entropy from covariance matrices of the atomic fluctuations. *J. Chem. Phys.* 115:6289–6292.

46. Schlitter, J. 1993. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem. Phys. Lett.* 215: 617–621.

47. Harris, S. A., E. Gavathiotis, M. S. Searle, M. Orozco, and C. A. Laughton. 2001. Cooperativity in drug-DNA recognition: a molecular dynamics study. *J. Am. Chem. Soc.* 123:12658–12663.

48. Gelpí, J. L., S. G. Kalko, X. de la Cruz, X. Barril, J. Cirera, F. J. Luque, and M. Orozco. 2001. Classical molecular interaction potentials: improved setup procedure in molecular dynamics simulations of proteins. *Proteins.* 45:428–437.

49. Sherer, E., S. A. Harris, R. Soliva, M. Orozco, and C. A. Laughton. 1999. Molecular dynamics studies of DNA A-tract structure and flexibility. *J. Am. Chem. Soc.* 121:5981–5991.

50. Noy, A., A. Pérez, F. Lankas, F. J. Luque, and M. Orozco. 2004. Relative flexibility of DNA and RNA: a molecular dynamics study. *J. Mol. Biol.* 343:627–638.

51. Pérez, A., A. Noy, F. Lankas, F. J. Luque, and M. Orozco. 2004. The relative flexibility of B-DNA and A-RNA duplexes: database analysis. *Nucleic Acids Res.* 32:6144–6151.

52. Rueda, M., S. G. Kalko, F. J. Luque, and M. Orozco. 2003. The structure and dynamics of DNA in the gas phase. *J. Am. Chem. Soc.* 125:8007–8014.

53. Pérez, A., J. R. Blas, M. Rueda, J. M. López-Bes, X. de la Cruz, and M. Orozco. 2005. Exploring the essential dynamics of B-DNA. *J. Chem. Theory and Comput.* 1:790–800.

54. Lu, X.-J., and W. K. Olson. 2003. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* 31:5108–5121.

55. Lavery, R., and J. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature of irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.

56. Drew, H. R., R. M. Wing, T. Tanako, C. Broka, S. Tanaka, K. Itakura, and R. Dickerson. 1981. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. USA.* 78:2179–2183.

57. Stolarski, R., A. Pohorille, L. Dudycz, and D. Shugar. 1980. Comparison of theoretical and experimental approaches to determination of conformation of nucleosides about the glycosidic bond. *Biochim. Biophys. Acta.* 610:1–19.

58. Hoogsteen, K. 1959. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* 12:822–823.

59. Sponer, J., J. Leszczynski, and P. Hobza. 1996. Structures and energies of hydrogen-bonded DNA base pairs. A nonempirical study with inclusion of electron correlation. *J. Phys. Chem.* 100:1965–1974.

60. Ponomarev, S. Y., K. M. Thayer, and D. L. Beveridge. 2004. Ion motions in molecular dynamics simulations on DNA. *Proc. Natl. Acad. Sci. USA.* 101:14771–14775.

61. Rueda, M., E. Cubero, C. A. Laughton, and M. Orozco. 2004. Exploring the counterion atmosphere around DNA: what can be learned from molecular dynamics simulations? *Biophys. J.* 87:800–811.