Molecular Dynamics Simulations of PNA·DNA and PNA·RNA Duplexes in Aqueous Solution

Robert Soliva,[†] Edward Sherer,[‡] F. Javier Luque,[§] Charles A. Laughton,[‡] and Modesto Orozco*,[†]

Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain, School of Pharmaceutical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK, and Departament de Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avgda Diagonal sn, Barcelona 08028, Spain

Received January 24, 2000. Revised Manuscript Received April 20, 2000

Abstract: Molecular dynamics simulations have been used to study the structure and flexibility of a DNA·PNA duplex and a RNA·PNA duplex in aqueous solution. In this study, trajectories have been generated starting from three different conformations of the PNA·DNA and PNA·RNA duplexes: A-like, B-like, and $P_{A/B}$-like. For the DNA·PNA duplex, the three trajectories converge within the nanosecond time scale to give structures resembling closely the P_B model. The RNA·PNA duplex trajectories started from A- and P_A-forms converge to give structures resembling the P_A model, but the trajectory begun from the B-like conformation leads to an unfolded duplex. Despite the similarity between P_A and P_B structures calculations show the existence of important differences in terms of molecular recognition between both conformations. Analysis of the trajectories shows that the PNA backbone is very flexible provided that the backbone movements do not alter the positioning of the bases. It is found that PNA is able to distort the structure of RNA and especially DNA strands during the formation of the PNA·DNA and PNA·RNA hybrids. The impact of these findings in antigene and antisense therapies is discussed.

Introduction

Peptide nucleic acids (PNAs, see Figure 1) are mimics of the nucleic acids where the polymeric unit is a neutral, achiral peptide (based on *N*-2(2-aminoethyl)-glycine units) instead of the usual phospho-ribose/phospho-2′-deoxyribose moieties found in physiological nucleic acids.[1−7] It has been demonstrated that one strand of PNA can recognize another PNA strand, forming PNA duplexes mimicking the behavior of nucleic acids.[8−9] Other studies have shown that PNA can interact with both DNA and RNA single strands, leading to extremely stable hybrid duplexes.[1,3,7,9−11] Recent studies have demonstrated that the PNA is also able to lead to triplex structures when combined with other oligonucleotidic strands,[9,12−14] as well as with PNA duplexes.[15] PNA−DNA interactions can be stronger than DNA−DNA interactions, so that the incubation of a suitable PNA strand with a DNA duplex can result in a process of strand displacement (refs 1 and 7 and references therein). Interestingly, this process is more efficient for supercoiled DNA (the predominant physiological form of DNA) than for relaxed linear DNA,[17] which is a clear advantage for the use of PNAs in antigene therapies.

The nuclease and protease-resistance of PNA,[12,18] combined with the extreme stability of PNA-hybrids,[1,3,7,9−11,19,20] make PNA an excellent drug candidate for antigene and antisense

[†] Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona.

[‡] University of Nottingham.

[§] Departament de Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona.

(1) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497−1500.

(2) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895−1897.

(3) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566−568.

(4) Wittung, P.; Nielsen, P. E.; Buchardt, P.; Egholm, M.; Nordén, B. *Nature* **1994**, *368*, 561.

(5) Bohler, C.; Nielsen, P. E.; Orgel, L. E. *Nature* **1995**, *376*, 578.

(6) Veselkov, A. G.; Demidov, V. V.; Frank-Kamenetiskii, M. D.; Nielsen, P. E. *Nature* **1996**, *379*, 214.

(7) Nielsen, P. E.; Egholm, M. In *Peptide Nucleic Acids. Protocol and Applications*; Nielsen, P. E., Egholm, M., Eds.; Horizon Scientific Press: Wymondham, 1999; pp 1−20.

(8) Rasmussen, H.; Sandholm, J. *Nat. Struct. Biol.* **1997**, *4*, 98−101.

(9) Kurarin, A.; Larsen, H. J.; Nielsen, P. E. *Chem. Biol.* **1998**, *5*, 81−89.

(10) Holmén, A.; Nordén, B. In *Peptide Nucleic Acids. Protocol and Applications*; Nielsen, P. E., Egholm, M., Eds.; Horizon Scientific Press: Wymondham, 1999; pp 87−97 and refs therein.

(11) Jensen, K. K.; Orum, H.; Nielsen, P. E.; Nordén, B. *Biochemistry* **1997**, *36*, 5072−5077.

(12) Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995**, *270*, 1838−1841

(13) Wittung, P.; Nielsen, P.; Nordén, B. *Biochemistry* **1997**, *36*, 7973−7979.

(14) Nielsen, P. E.; Christensen, L. *J. Am. Chem. Soc.* **1996**, *118*, 2287−2288.

(15) Wittung, P.; Nielsen, P.; Nordén, B. *J. Am. Chem. Soc.* **1997**, *119*, 3189−3190.

(16) Ishihara, T.; Corey, D. R. *J. Am. Chem. Soc.* **1999**, *121*, 2012−2020.

(17) Bentin, T.; Nielsen, P. E. *Biochemistry* **1996**, *35*, 8863−8869.

(18) Kuwahara, M.; Arimitsu, M.; Sisido, M. *J. Am. Chem. Soc.* **1999**, *121*, 256−257.

(19) Nielsen, P. E.; Christensen, L. *J. Am. Chem. Soc.* **1996**, *118*, 2287−2288.

(20) Chernym, D. Y.; Belotserkiwskii, B. P.; Frank-Kamenetskii, M.; Egholm, M.; Buchardt, O.; Berg, R. H.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1667−1670.

10.1021/ja000259h CCC: $19.00
Published on Web 06/09/2000

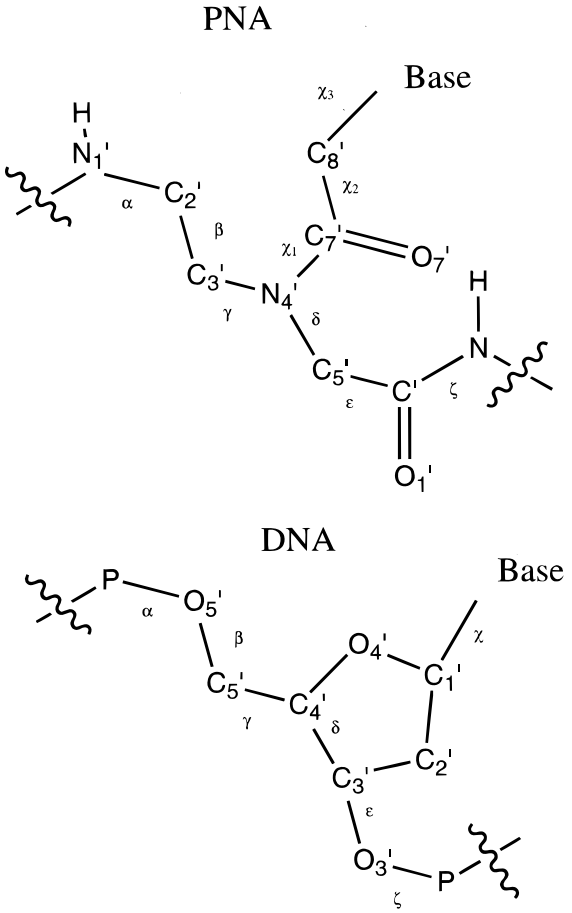

Figure 1. Schematic representation of the PNA and DNA backbones.

strategies (refs 21−28 and references therein), as well as a promising biotechnological tool (refs 29−35 and references therein). This explains the large amount of research effort focused on the structure of PNA and its hybrids.[8,12,15,36−38] A wide repertoire of physical techniques, including high-resolution NMR and X-ray crystallography[8,12,15,36−38] have been used to analyze the structures of complexes of PNA with other nucleic acid polymers. Thus, NMR-refined structures of PNA·DNA[38] and PNA·RNA[36] duplexes have been published, and the structures of a PNA·PNA duplex[8] and a PNA·DNA·PNA triplex[12] have been solved from X-ray diffraction data. Furthermore, evidence of other structures, such as the PNA·PNA·PNA triplex[15] have been obtained from low-resolution CD data. These studies have highlighted important structural features of PNA hybrids. However, fundamental questions related to the dynamic characteristics of these structures, their recognition properties, and the plasticity of the peptide backbone remain yet to be fully elucidated. Knowledge of this information is essential for the successful exploitation of PNA-based therapeutic and biotechnological strategies.

Theoretical calculations, particularly molecular dynamics (MD) simulations, have been proved to be a very powerful complementary technique for the study of nucleic acids, including normal and anomalous structures of duplexes,[39−48] triplexes,[45,49−55] and other forms of DNA.[41,46,53,56,57] Very recently, Sen and Nilsson[46] have reported MD simulations on parallel and antiparallel PNA·DNA duplexes and on the PNA·PNA duplex. Olson and co-workers[52] have reported molecular mechanics calculations of different conformations of the PNA·PNA·PNA triplex, suggesting that it adopts an A-like structure. Recently[53] we reported the first MD simulation of a PNA·DNA·PNA triplex, finding very good agreement between MD and high-resolution X-ray data, even when the simulations started from conformations very different from the experimentally one.[12]

(21) Nielsen. P. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 167−183.
(22) Eriksson, M. E.; Nielsen, P. E. *Q. Rev. Biophys.* **1996**, *29*, 369−394.
(23) Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thompson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481−1488.
(24) Good, L.; Nielsen, P. E. In *Peptide Nucleic Acids. Protocol and Applications*; Nielsen, P. E., Egholm, M., Eds.; Horizon Scientific Press: Wymondham, 1999; pp 213−220.
(25) Larsen, H. J.; Nielsen, P. E. In *Peptide Nucleic Acids. Protocol and Applications*; Nielsen, P. E., Egholm, M., Eds.; Horizon Scientific Press: Wymondham, 1999; pp 221−240.
(26) Raney, K. D.; Hamilton, S. E.; Corey, D. R. In *Peptide Nucleic Acids. Protocol and Applications*; Nielsen, P. E., Egholm, M., Eds.; Horizon Scientific Press: Wymondham, 1999; pp 241−252.
(27) Knudsen, H.; Nielsen, P. E. *Anti-Cancer Drugs* **1997**, *8*, 113−118.
(28) Nielsen, P. E. *Curr. Opin. Biotechnol.* **1999**, *9*, 353−357.
(29) Nielsen, P. E. *Curr. Opin. Biotechnol.* **1999**, *10*, 71−75.
(30) Corey, D. R. *Trends Biotechnol.* **1997**, *15*, 224−229.
(31) Misra, H. S.; Pandey, P. K.; Modak, M. J.; Vinayak, R.; Pandey, V. N. *Biochemistry* **1998**, *37*, 1917−1925.
(32) Orum, H.; Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanley, C. *Nucleic Acids Res.* **1993**, *21*, 5332−5336.
(33) Demidov, V. V.; Frank-Kamenetskii, M. D. In *Peptide Nucleic Acids. Protocol and Applications*; Nielsen, P. E., Egholm, M., Eds.; Horizon Scientific Press: Wymondham, 1999; pp 163−174.
(34) Orum, H. In *Peptide Nucleic Acids. Protocol and Applications*; Nielsen, P. E., Egholm, M., Eds.; Horizon Scientific Press: Wymondham, 1999; pp 185−194.
(35) Liang, X.; Zelphati, O.; Nguyen, C.; Felgner, P. In *Peptide Nucleic Acids. Protocol and Applications*; Nielsen, P. E., Egholm, M., Eds.; Horizon Scientific Press: Wymondham, 1999; pp 201−212.
(36) Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994**, *265*, 777−780.
(37) Leijon, M.; Nielsen, P. E.; Buchardt, O.; Nordén, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry* **1994**, *33*, 9820−9825.
(38) Eriksson, M. and Nielsen. P. E. *Nat. Struct. Biol.* **1996**, *3*, 410−413.
(39) Soliva, R.; Luque, F. J.; Alhambra, C.; Orozco, M. *J. Biomol. Struct. Dyn.* **1999**, *17*, 89−99.
(40) Cheetham, T. E.; Miuller, J. L.; Fox, T.; Darden, T.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.
(41) Cheetham, T. E.; Kollman, P. A. *J. Mol. Biol.* **1996**, *259*, 434−444.
(42) Yang, L.; Pettit, B. M. *J. Phys. Chem.* **1996**, *100*, 2564−2566.
(43) Young, M. A.; Ravishanker, G.; Beveridge, D. L. *Biophys. J.* **1997**, *73*, 2313−2336.
(44) Jayaram, B.; Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Am. Chem. Soc.* **1998**, *120*, 10633.
(45) Güimil Garcia, R.; Ferrer, E.; Macías, M. J.; Eritja, R.; Orozco, M. *Nucleic Acids Res.* **1999**, *27*, 1911−1198.
(46) Sen, S.; Nilsson. L. *J. Am. Chem. Soc.* **1998**, *120*, 619.
(47) Cubero, E.; Sherer, E. C.; Luque, F. J.; Orozco, M.; Laughton, C. A. *J. Am. Chem. Soc.* **1999**, *121*, 8653−8654.
(48) Sherer, E. C.; Harris, A. A.; Soliva, R.; Orozco, M.; Laughton, C. A. *J. Am. Chem. Soc.* **1999**, *121*, 5981−5991.
(49) Weerasingle, S.; Smith, P. E.; Mohan, V.; Cheng, Y. K.; Pettit, M. *J. Am. Chem. Soc.* **1995**, *117*, 2147−2158.
(50) Weerasingle, S.; Smith, P. E.; Pettit, M. *Biochemistry* **1995**, *34*, 16269−16278.
(51) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1997**, *119*, 7463−7469.
(52) Srinavasan, A. R.; Olson, W. K. *J. Am. Chem. Soc.* **1998**, *120*, 492−499.
(53) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 5895−5904.
(54) Soliva, R.; Luque, F. J.; Orozco, M. *Nucleic Acids Res.* **1999**, *27*, 2248−2255.
(55) Soliva, R.; Laughton, C. A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 11226−11233.
(56) Spackova, N.; Berger, I.; Egli, M.; Sponer, J. *J. Am. Chem. Soc.* **1998**, *120*, 6147.
(57) Luo, J.; Bruice, T. C. *J. Am. Chem. Soc.* **1998**, *120*, 115−1123.

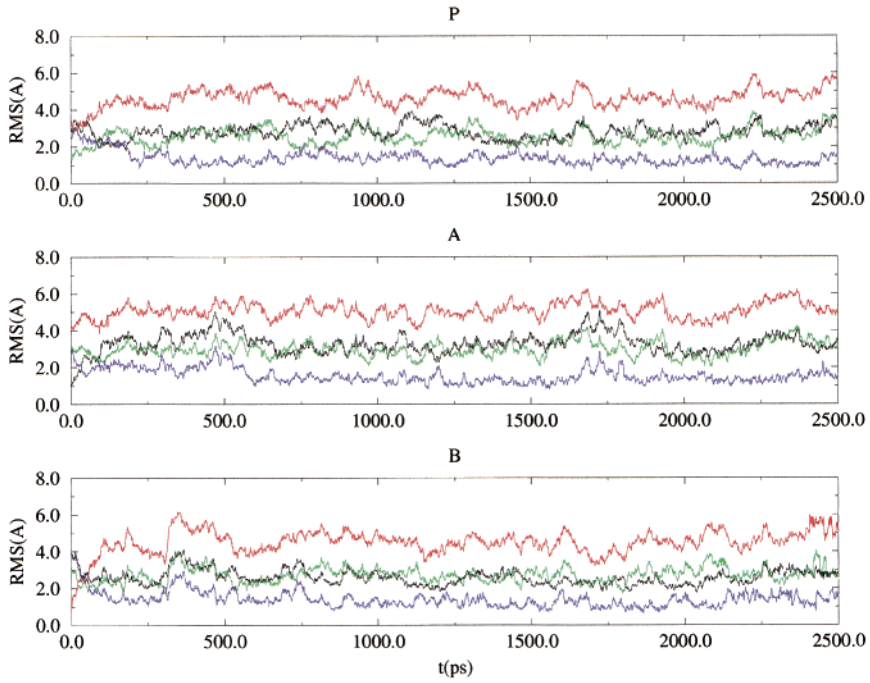

Figure 2. Root-mean-square deviation (rmsd in Å) in the three trajectories of the PNA·DNA duplex. Reference structures are: B-type duplex (red), A-type duplex (black), NMR structure (green), and MD-averaged structures for each trajectory (blue). Simulation times are in picoseconds.

In this paper we report a systematic study of PNA·DNA and PNA·RNA duplexes by MD simulations. The results, which agree with available experimental data derived from NMR experiments,[36−38] provide a complete picture of the conformational flexibility of both duplexes. They also allow us to gain insight into the PNA-induced changes in nucleic acid structures, which might be important to improve the binding characteristics of this polymer to DNA.

Methods

One sequence of PNA·DNA (5′-GACATAGC-3′) and one of PNA·RNA (5′-GAGUUC-3′) were considered in this study. Choice of these sequences was motivated by the availability of NMR data (PDB entries 1PDT,[38] and 176D,[36] respectively), which allowed us to test the quality of our simulations by comparison with the experimental data.

Three different starting structures for MD simulations were considered for each duplex. The first starting structure (referred to as P in the following) was built using the first NMR structure in PDB entries 1PDT and 176D (the remaining structures deposited in both PDB entries were used later to check the convergence of the results). The second starting structure (denoted as B hereinafter) was obtained from restrained energy minimization of the preceding one. For this purpose, a DNA·DNA duplex with the appropriate sequence was constructed using standard B-type parameters as template. In the minimization process the coordinates of all atoms in the DNA or RNA strand and of the bases in the PNA strand were restrained to fit the positions of the corresponding atoms in the template structure with a force constant of 100 kcal mol^{-1} Å^2. Finally, the third starting model (denoted as A in the following) was a duplex structure (PNA·DNA or PNA·RNA) with the conformation of an A-type RNA duplex generated by using a procedure analogous to that reported above.

All of the starting models were immersed in a cubic box containing around 2000 TIP3P[58] water molecules. Sodium ions were added to give electroneutrality following the standard procedures.[51,53−55] The systems were then subjected to our standard seven-step equilibration process, which includes energy minimization and up to 130 ps of partially restrained molecular dynamics simulation.[51,53−55] The final structures after the equilibration process were used as the starting points for 2.5 ns unrestrained MD simulations, as well as for 0.5 ns restrained MD simulations using a harmonic restraint of 5 kcal mol^{-1}Å^{-2} on the nucleobases. Averaged structures were obtained for the last 2.0 ns of the unrestrained MD trajectories using the procedure noted elsewhere.[51,53−55] Molecular interaction potentials[51,53−55] were computed for the MD-averaged structures, which were determined by averaging of snapshots taken from either the three MD simulations performed for the PNA·DNA duplex or from the two MD simulations for the PNA·RNA duplex (see below). Hydration densities[51,53−55] were also determined from the last 2.0 ns of the unrestrained MD simulations. Averaged properties are then obtained using 4 (PNA·RNA) or 6 (PNA·DNA) ns of MD trajectories.

All MD simulations were performed in the isothermic−isobaric ensemble (NPT, $P = 1$ atm, $T = 298$ K), using periodic boundary conditions (PBC), and the particle mesh Ewald method (PME; ref 40,-59) to account for long-range effects. SHAKE[60] was used to constrain all the chemical bonds (conecting heavy or hydrogen atoms), which allowed us to use an integration time step of 2 fs. The AMBER-95 force-field[61] combined with previously developed parameters for the PNA backbone[53] was used.

All MD simulations were done using AMBER5.[61,62] Analyses were done using AMBER5, Curves,[63] as well as in-house programs. All calculations were carried out on the SGI-ORIGIN-2000 of the Centre de Supercomputació de Catalunya (C4-CESCA), as well as on workstations (SGI-ORIGIN-200) in our laboratories.

Results and Discussion

Trajectories Of The PNA·DNA Duplex. The three unrestrained MD simulations of the PNA·DNA duplex lead to stable trajectories in geometrical and energetic terms, and the sampled

(58) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(59) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden. T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(60) Ryckaert, J. P.; Ciccote, G.; Berendsen, J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(61) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K.; Ferguson, D. M. Spellmeyer, Fox, D. C.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(62) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 5*, University of California: San Francisco, 1997.
(63) Lavery, R.; Sklenar, H. *Curves 5.1*, Computer Program; Institut de Biologie Physico-Chimique: CNRS, 1996.

Table 1. RMS Deviations (in Å) of the Three MD-averaged Structures of the PNA·DNA Duplex; Comparisons with the NMR Structure in 1PDT Were Also Performed

	A-averaged	B-averaged	1PDT[a]
P-averaged	1.1	1.4	2.4
A-averaged		1.5	2.5
B-averaged			2.4

[a] The first structure deposited in 1PDT was used here for comparison.

configurational space corresponds to a double helix. Calculation of root-mean-square deviations (rmsds, see Figure 2) with respect to the MD-averaged structures confirms the stability of the trajectory. Furthermore, calculation of the rmsd for the three trajectories with respect to their respective starting structures shows the different behavior of trajectories A and B (starting from A- and B-models, as defined in ref 63), and trajectory P (starting from the NMR structure, ref 38). Thus, the P-simulation shows a very small increase in the rmsd with respect to the NMR structure, while A- and especially B-trajectories clearly diverge from their starting conformations.

The average rmsd (Figure 2) of the P-trajectory with respect to the NMR structure used as starting conformation is around 2.6 Å, and 2.4 Å if the MD-averaged structure of the P-trajectory is considered (see Table 1). These values are similar to those found by comparing the different NMR structures presented in PDB entry 1PDT (the rmsd of the NMR structure used here to start the P-trajectory from the remaining NMR-derived models deposited in 1PDT ranges from 0.3 to 2.0 Å). Interestingly, the P-trajectory stays closer to the A- form, (rmsd around 3.0 Å) than to the B-form (rmsd around 4.9 Å), as can be stated in Figure 2.

The rmsd (Figure 2) in the A- trajectory with respect to the canonical A-form shows a sharp increase in the first 200 ps to about 3 Å, and then remains stable for the rest of the simulation (average rmsd around 3.2 Å). There is also an increase of around 1 Å in the rmsd with respect to the B-form, which changes from 4.0 to 5.1 Å during the simulation. Finally, the rmsd from the NMR structure shows a slight decrease at the beginning of the simulation, and stabilizes around 2.8 Å during the second part for the A-trajectory. A similar value is obtained when the time-averaged structure of the A-trajectory is compared with the NMR structure (2.5 Å; see Table 1).

Finally, the B-trajectory shows a clear and rapid divergence from the starting B-form model (see Figure 2). The rmsd rises to 5 Å within a few hundred ps, and averages around 5 Å in the final stages of the simulation. The increase in rmsd with respect to the B-form is accompanied by a reduction of the rmsd with respect to both A-form and NMR-based models. Thus, the rmsd with respect to the A-form is reduced from 4.0 Å in the starting model to an averaged value of 2.9 Å during the second half of the trajectory. Similarly, the rmsd with respect to the NMR-based starting structure decreases from 3.5 Å in the beginning of the simulation to an average value of 2.6 Å during the last 2 ns the trajectory. Again, a similar rmsd value is obtained when the time-averaged structure of the B-trajectory is compared with the NMR structure (2.4 Å; see Table 1). It is worth noting that there is general agreement between our 2.5 ns B-trajectory and a previous CHARMM 0.6 ns trajectory.[46] In both cases the structure moves from canonical B-form to a structure with rmsd from standard B-form of 4−5 Å. This agreement gives strong confidence in the results.

In summary, the three trajectories, irrespective of the starting conformation, lead to configurations close in conformational space. This is noted in the fact that the MD-averaged structures obtained from the second half of P-, A-, and B-trajectories shows cross-rmsds of around 1 Å (see Table 1). Moreover, the results in Figure 2 and Table 1 also show that the structures sampled during the last part of the three trajectories are close to those determined from NMR data (rmsd around 2 Å). It is also clear that the general structural characteristics are closer to those of an A-type DNA duplex (rmsd around 3 Å) than to those of a B-type DNA duplex (rmsd around 5 Å). It is clear then that the structure of a PNA·DNA duplex is not similar to that of a physiological DNA duplex.

Trajectories Of The PNA·RNA Duplex. Trajectories started from the B-type conformation of the PNA·RNA duplex failed to reach equilibrium within a few hundred ps and began to unwind. Many attempts were made for equilibrating the structure, including extended dynamics with restraints, larger equilibration periods for solvent and counterions (see Methods), but all these attempts failed. This indicates that the B-form model is far from correct for a PNA·RNA hybrid.

The trajectories starting from the NMR[36] and A-forms have very similar rmsd profiles (see Figure 3). They both move slightly (rmsd around 2 Å) from their starting conformations in around 100 ps, and then remain stable for the rest of the simulation. Root-mean-square deviations with respect to the MD-averaged structures obtained from the last 2 ns of the two trajectories are around 1 Å, confirming the stability of the trajectories. Comparison of the structures sampled in A- and P-trajectories with the structures collected in PDB entry 176D show excellent agreement. Thus, the rmsds with respect to the first structure deposited in 176D were around 2 Å in both cases, a value not much larger than that found when the 10 experimental structures deposited in 176D were compared (average rmsd 1.62 Å, maximum 2.47 Å). The averaged structures obtained from the last 2 ns of the P- and A-trajectories are very similar (rmsd below 1.0 Å). Overall, there is clear convergence in the two trajectories. Indeed, they agree that the global structure of the PNA·RNA duplex in MD simulations is very similar to that reported from NMR data,[36] and can be defined as an A-type helix,[64] even though few structural and reactive characteristics of the hybrid are more similar to those of a B-type duplex (see below[64]).

In summary, 2.5 ns MD trajectories on the PNA·DNA and PNA·RNA duplexes show excellent convergence toward the experimental structures. Indeed, the main structural features resemble those of the canonical A-form, and are distant from those of the canonical B-form. This is illustrated, for instance, in the fact that the equivalent backbone atoms of the PNA·DNA (central hexamer) and PNA·RNA MD-averaged structures show an rmsd of 1.9 Å. For reference, a similar comparison of canonical A and B− form DNA·DNA hexamer duplexes gives an rmsd of 3.5 Å. In addition, the analysis of the high-resolution X-ray structure of the PNA·PNA duplex shows that the homoduplex of PNA also exhibits several characteristics of an A-type duplex.[64] We can conclude, therefore, that PNA is not a completely malleable molecule that can fit to any oligonucleotide counterpart, but shows marked conformational preferences. This point will be analyzed in detail below.

Structure Analysis. (a) The PNA·DNA Duplex. The convergence of the three trajectories allowed us to obtain reliable statistics of the structural dynamics of the PNA·DNA duplex using the second halves of the P-, B-, and A- trajectories. Analysis of the three MD-averaged structures shows that the structures sampled during the dynamics correspond to double

(64) Arnott, S.; Hukins, D. W. L. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1504−1510.

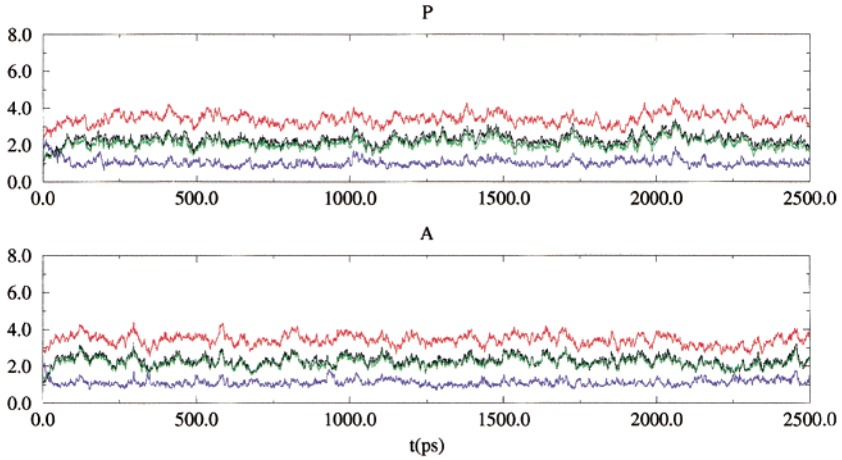

Figure 3. Root-mean-square deviation (rmsd in Å) in the A- and P-trajectories of the PNA·RNA duplex. Reference structures are: B-type duplex (red), A-type duplex (black), NMR structure (green), and MD-averaged structures for each trajectory (blue). Simulation times are in picoseconds. The B-trajectory leads to unfolded structures in around 500 ps and is not shown.

helices, with local and global twist angles in the range of 21−27°. These values are slightly smaller than those found in the NMR structures (26−28°) for this duplex, and very similar to those found in previous MD simulations.[46] A tendency for MD simulations of nucleic acids to lead to slightly underwound structures is well-documented.[39,65]

Interestingly, as shown in Figure 4, the twist angle changes dramatically along the sequence, being very high at the T−A step and very low at the C−A step (see Figure 4). This behavior is fully compatible with NMR data,[38] and recent statistical analysis of DNA structures.[66] The agreement between twist preferences for a given step between DNA·DNA and PNA·DNA duplexes strongly suggests that the sequence-dependent changes of twist found in this PNA·DNA duplex are due to intrinsic stacking preferences, and not to specific backbone-base interactions. It is worth noting that the same twist profile is obtained from the three trajectories, even though in the latter two the starting conformation was generated with the same twist angle for all of the steps. It seems then that MD simulations are powerful enough to accurately reproduce fine structural details, even if the starting structures are incorrect.

Amplitudes of sugar puckerings are typically 37−39°, as shown in most nucleic acid structures, and far from the very low values found in the NMR structures of this duplex, where the furanose rings are almost planar. We believe that the reason for this discrepancy might be related to the existence of a fast repuckering in the time scale of the NMR experiment, which might mask the real puckering amplitude of the sugars during the model refinement of the NMR structure. The phase angles move between the O1′-*endo* and the C1′-*exo* regions, sampling the *east* and *south-east* zones of the pseudorotational cycle, as found in the NMR structures. There is also good agreement with previous MD simulations, where the sugars were mostly found in the *south-east* region.[46] Thus, the MD simulations again support the NMR data and show that, despite the close overall similarity of the MD-derived structures to the canonical A-form, the sugar puckerings correspond to those of a B-type duplex. This illustrates the danger of classifying nucleic acid structures solely on the basis of their sugar puckerings, especially in the case of hybrids containing nucleic acids others than DNA or RNA.[37]

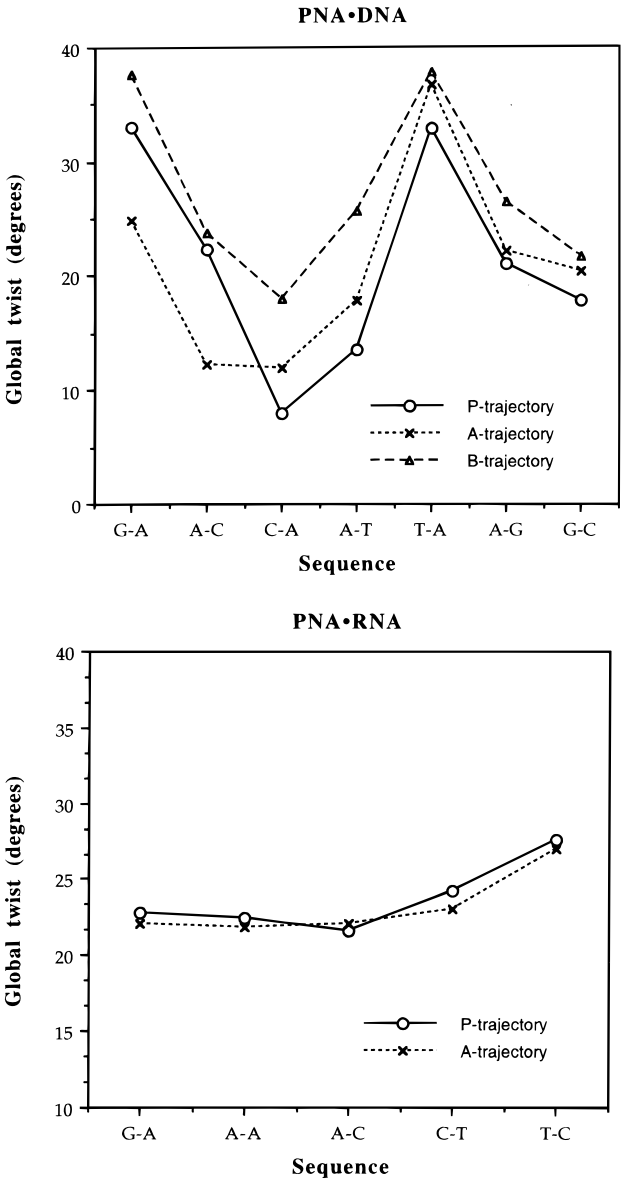

Figure 4. Global twist along the sequence for the MD-averaged structures of the PNA·DNA and PNA·RNA duplexes.

Finally, the MD trajectories can be used to analyze the formation of intra-strand H-bonds between peptide moieties of

(65) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dynam.* **1999**, *16*, 845−862.
(66) Gorin, A. A.; Zhurkin, V. B.; Olson, W. K. *J. Mol. Biol.* **1995**, *247*, 34−48.

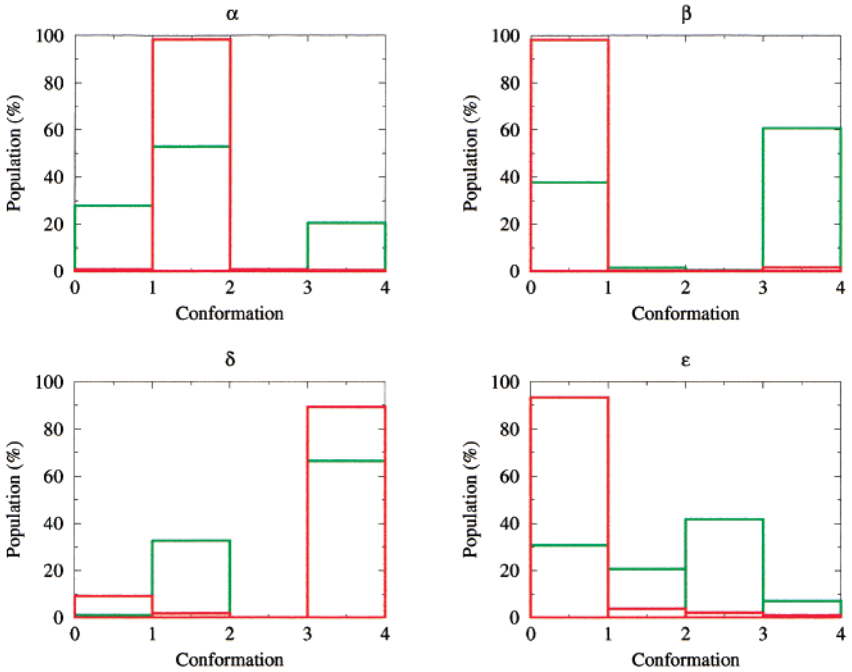

Figure 5. Histogram representation of the population of selected dihedral angles of the PNA (green) and DNA (red) backbones obtained in the three trajectories of the PNA·DNA duplex. Numbering of conformers is: *anti* (0−1), *gauche*$^{(-)}$ (1−2), *syn* (2−3), and *gauche*$^{+)}$ (3−4), see text for details.

the PNA backbone. Analysis of O7′-H1′ distances shows that in average they are at around 3.4 Å, with standard deviations around 0.5 Å. Analysis of the three independent trajectories shows that intra-strand O7′−H1′ H-bond (see Figure 1) exists only 0.8% of the time. This finding agrees with experimental evidence,[38] and other MD simulations,[46] and disagrees with previous modeling studies,[67] which suggested the existence of such intramolecular H-bonds based on gas-phase calculations.

(b) The PNA·RNA Duplex. The time-averaged structures of the PNA·RNA duplex have step-averaged local helical twist values of 23.7° (P-trajectory) and 23.2° (A-trajectory), whereas the average helical twist for the structures deposited in PDB entry 176D is 30.6°. The slightly greater unwinding of the MD-averaged duplex relative to the NMR structure may reflect the shorter length of the duplex and thus the greater importance of end-effects. In the NMR structures, the AA step shows the lowest twist (average 24.7°). Interestingly, while this is not evident in the P-trajectory, it is observed, albeit to a reduced extent, in the A-trajectory, which began from a canonical structure with identical twist for all of the steps. This supports again the ability of MD simulations to correct wrong structural features and to drive structures toward realistic regions of the conformational space.

All of the NMR structures show sugar puckerings tightly confined to the *north* region. This is also observed in the structures collected in P- and A-trajectories with the exception of G7, which in both trajectories is sampling equally the *south-east* and *north* regions. It is worth noting that the B-type trajectory, where all of the sugar puckerings were in the *south* region, leads to unfolding of the structure, demonstrating the impossibility to have *south* puckerings in PNA·RNA duplexes. Finally, as in the PNA·DNA simulation, we very rarely observe potential hydrogen-bonding interactions between PNA backbone O7′ and H1′ atoms. Distances less than 2.5 Å between these atoms happen only 4% of the time. This finding, which agrees with the available experimental evidence,[36,38] argues against the importance of intrastrand hydrogen-bonds for stabilizing PNA-hybrid structures in solution.[67]

Backbone Conformational Flexibility. (a) The PNA·DNA Duplex. The rmsd plots in Figure 2 show that the structure of the PNA·DNA duplex is quite flexible. To gain deeper insight into the flexibility of the hybrid duplex, we have examined the PNA and DNA dihedral angles. Figure 5 presents the histogram representation of selected PNA and DNA backbone angles (sequence-averaged) sampled in the A-, B-, and P-trajectories. These histograms were determined from population factors obtained after classification of each backbone angle of type Θ for configuration i and for step k into one of the following categories: (1) *anti* (Θ_{ik} in the range 180° ± 45°), (2) *gauche*$^{(-)}$ (Θ_{ik} in the range 270° ± 45°), (3) *syn* (Θ_{ik} in the range 0° ± 45°), and (4) *gauche*$^{(+)}$ (Θ_{ik} in the range 90° ± 45°). Figure 5 clearly reveals dramatic differences in terms of flexibility between the DNA and PNA backbones. Thus, all of the DNA backbone angles are grouped in a narrow region of the conformational space, while for PNA a larger portion of the conformational space is sampled. For instance, in the DNA strand the dihedral angles β and ϵ are most of the time *anti*, as is α in the *gauche*$^{(-)}$ region and δ in the *gauche*$^{(+)}$ region. In contrast in the PNA strand, α samples with similar intensity *gauche*$^{(-)}$, *gauche*$^{(+)}$, and *anti* regions, β samples *gauche*$^{(+)}$, and *anti* regions in similar proportions, δ show a significant population of *gauche*$^{(-)}$ and *gauche*$^{(+)}$, and ϵ samples almost all of conformational space.

The larger flexibility of sequence-averaged PNA backbone angles (Figure 5) is, in part, the consequence of the greater variability of each individual dihedral angle. This is noted in the standard deviations of the backbone angles and more clearly in the order factors (Ξ) determined as shown in eqs 1 and 2, where n stands for the maximum number of steps in the PNA·DNA duplex (excluding the 5′ and 3′ ends), Θ is the backbone angle, and the indexes A, B, and P refer to the trajectory used. As seen in eq 2, order factors determined for the three

(67) Almarsson, O.; Bruice, T. C.; Kerr, J.; Zuckermann, R. N. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7518−7522.

Table 2. Average Order Factors for the Backbone Torsional Angles Determined from MD-averaged Structures of PNA·DNA and PNA·RNA Duplexes (See Text for Details)

strand	backbone angle	order factor (DNA·PNA)	order factor (RNA·PNA)
PNA	α	0.71	0.31
	β	0.96	0.94
	δ	0.98	0.97
	ϵ	0.72	0.32
	γ	0.98	0.99
	χ_1	0.98	
	χ_2	0.97	
	χ_3	0.98	
	ζ	0.98	0.98
DNA or RNA	α	0.97	0.98
	β	0.98	0.99
	δ	0.97	0.96
	ϵ	0.95	0.98
	γ	0.98	0.99
	χ	0.95	0.99
	ζ	0.93	0.99

trajectories were averaged to obtain trajectory-independent values (see Table 2).

$$\Xi_{\Theta}^{\mathrm{sim}} = \frac{1}{n}\sum_{k=1}^{n}(\langle\cos(\Theta_k)\rangle^2 + \langle\sin(\Theta_k)\rangle^2)^{1/2} \quad (1)$$

$$\Xi_{\Theta} = \frac{1}{3}(\Xi_{\Theta}^{\mathrm{A}} + \Xi_{\Theta}^{\mathrm{B}} + \Xi_{\Theta}^{\mathrm{P}}) \quad (2)$$

where the brackets represent properties average over the MD ensembles.

Most of DNA torsions are restricted to move in relatively narrow range around the optimum values, yielding order factors (Table 2) very close to 1. On the contrary, the PNA backbone is flexible, as seen particularly in the α and ϵ dihedrals which have order factors of around 0.7. We cannot directly compare these findings with the NMR data, but the very large standard deviation in α and ϵ torsions reported in ref 38 and the correlation found experimentally between α and ϵ torsions[68] strongly support our conclusions. The largest flexibility of the α and ϵ torsions is also supported by previous MD simulations[46] using a different force-field; even MD results in ref 46 do not show a dramatic difference in terms of flexibility between the PNA and DNA backbones. In any case, it is worth noting that the high flexibility in α and ϵ is achieved without major displacements of the bases (for instance, no base pair opening is detected during the trajectories) but produces changes in the orientation of the amide group (see below).

A second reason for the wide sampling of PNA backbone torsions (Figure 5) is the sequence dependence of PNA backbone angles. Thus, inspection of MD-averaged B-, A-, and P-structures shows a large variation in PNA backbone angles along the sequence (see Figure 6). In sharp contrast, torsion angles in the DNA backbone remain largely constant along the sequence. In other words, not only do most dihedral angles in the PNA strand oscillate significantly about their equilibrium value, but also that equilibrium value is sequence-dependent.

The movements of the PNA backbone, particularly in α and ϵ bonds change the orientation of the secondary amide bonds. Two different arrangements of the carbonyl group are found. One corresponds to the carbonyl group pointing toward the C-terminus of the helix (forward, see Figure 7), and the other to the carbonyl group pointing toward the N-terminus (backward, see Figure 7). As noted above, no change in the positioning of the base pairs occurs due to these movements as clearly seen in Figure 7. Using the $C8_i'-N4_i'-C_i'-O1_i'$ dihedral angle (υ) as a measure of the placement of the carbonyl group, two regions can be defined ("forward" for υ in the range $-60 \pm 90°$, and "backward" for υ in the range $120 \pm 90°$). Analysis of the three trajectories yields to percentages of around 75% for the backward configuration and 25% for the forward configuration, yielding a free energy difference of 0.6 kcal/mol between the backward and forward configuration. We should note that these results were obtained by averaging the three trajectories, but very similar values were obtained by considering each individual trajectory, which demonstrates the excellent convergence of this free energy value. It is worth noting that the tertiary amide carbonyl groups always point toward the C-terminal of the helix, and thus should generate a significant dipole. In the backward conformation, the secondary amide carbonyls are oriented to oppose this dipole, whereas in the forward conformation they reinforce it. The consequences of this in regard to hydration are further discussed below. There are only eight structures in the PDB entry 1PDT, which precludes a statistical analysis as that performed here for MD trajectories. In any case, what is clear in NMR structures is the coexistence of backward and forward configurations, giving experimental support for the large flexibility of the PNA-backbone found in our MD simulations.

(b) The PNA·RNA Duplex. The PNA·RNA duplex is also quite flexible as noted in the rmsd fluctuations shown in Figure 3. Indeed, the PNA is much more flexible than the RNA strand as noted in the histograms in Figure 8. It is worth noting that the change in conformational region in RNA backbone angles is rare, and that ϵ and β are always in the *anti* region, α is almost all of the time in the *gauche*$^{(-)}$ region, and δ samples most of the time the *gauche*$^{(+)}$ region. On the contrary, the PNA backbone is quite flexible, especially in the α angle (which samples *anti*, *gauche*$^{(+)}$, and *gauche*$^{(-)}$ regions), and ϵ which as found for PNA·DNA duplexes samples most of the conformational space. It is worthwhile to compare histograms in Figure 8 with those for PNA·DNA duplexes in Figure 5. Despite the similar trends found in both plots, few differences emerge: (i) the RNA backbone is slightly more rigid than the DNA backbone, and (ii) even though there is general similarity between PNA·DNA and PNA·RNA duplexes (noted in the small rmsd), there are differences in the individual dihedral angles, which is expected to be a consequence of the different puckering of the sugars.

The variability in PNA backbone dihedral angles along the sequence is very large, even though the importance of end-effects in a short sequence precludes a quantitative analysis like that performed for the PNA·DNA duplex. Inspection of the order factors in Table 2 shows that all the conclusions drawn from the PNA·DNA simulations also hold for the PNA·RNA simulations. It is clear that PNA backbone is much more flexible than the RNA one, as suggested by histogram plots. Furthermore, the smaller order factors are those of the PNA backbone torsion angles, α and ϵ, suggesting again that these two bonds are the most flexible points of the PNA backbone. We should note that all these findings agree with experimental data, which indicates the large flexibility of the PNA backbone, especially in the α and ϵ bonds.[36,68]

In summary, the PNA backbone is very flexible irrespective of the nature of the other polinucleotidic strand. Most of the flexibility stems from the α and ϵ dihedral angles, but it does not produce major changes in the key helical parameters of the

(68) Topham, C. M.; Smith, J. C. *J. Mol. Biol.* **1999**, *292*, 1017−1038.

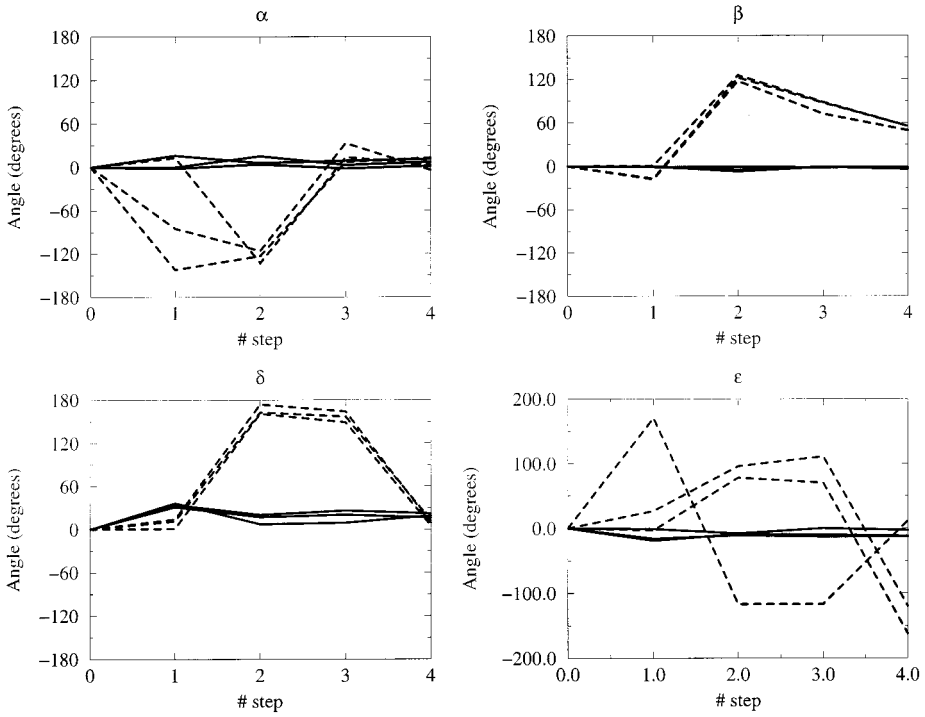

Figure 6. Variation of selected dihedral angles of MD-averaged structures along the sequence for the three trajectories of the PNA·DNA duplex. Solid lines correspond to the DNA strand, and dashed lines correspond to the PNA strand. All values are referred to the first dihedral angle in the sequence.

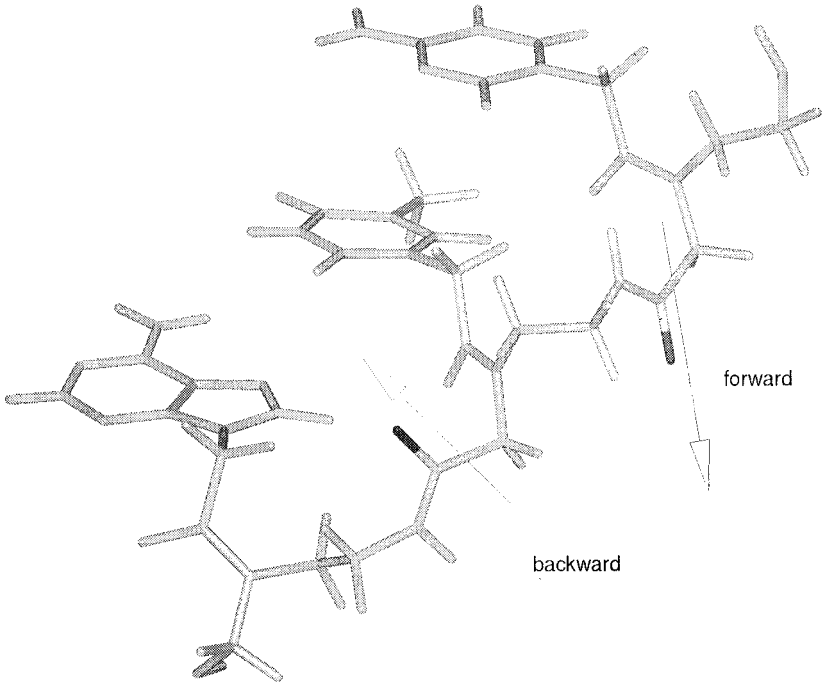

Figure 7. Detail of the "forward"−"backward" orientations of the carbonyl group (see text for details).

helix. It is worth noting that the large flexibility of the PNA backbone in the "bound" form suggests a very large flexibility in the "unbound" form, which might imply a poor "preorganization" of the PNA strand, and consequently an important energy penalty to interact with DNA or RNA strands. Our results suggest that derivatives of PNAs with restricted flexibility in the backbone might improve binding to natural oligonucleotides.

Using the dihedral angle $C8_i'-N4_i'-C_i'-O1_i'$ (v) to examine the forward and backward amide carbonyl conformations, we find the backward conformation to be slightly more populated (56%). This closely mirrors the results found for the PNA·DNA simulations, even though for PNA·RNA duplexes the free energy change associated with the transition between backward and forward configurations is less than 0.2 kcal/mol. Once again NMR structures deposited in PDB entry 176D support the existence of this conformational equilibrium.

Molecular Recognition. An initial insight into the molecular recognition properties of the PNA·DNA structure was obtained by inspection of hydration patterns (see Methods) determined by averaging water molecule distributions for the last 2.0 ns of each of the B-, P-, and A-trajectories (see Figure 9). The hybrid

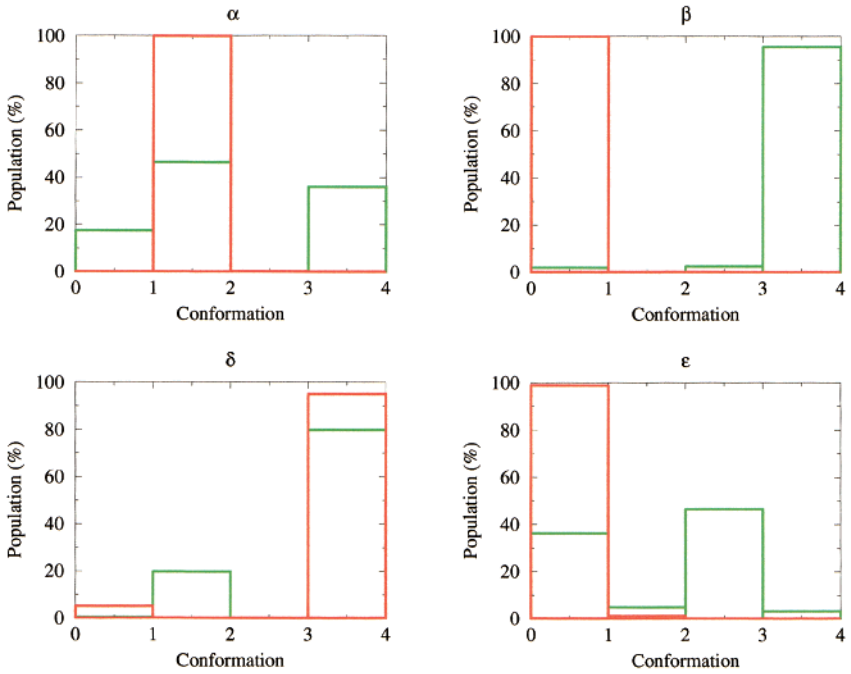

Figure 8. Histogram representation of the population of selected dihedral angles of the PNA (green) and RNA (red) backbones obtained in the two stable trajectories of the PNA·RNA duplex. Numbering of conformers is: *anti* (0−1), *gauche*$^{(-)}$ (1−2), *syn* (2−3), and *gauche*$^{(+)}$ (3−4), see text for details.

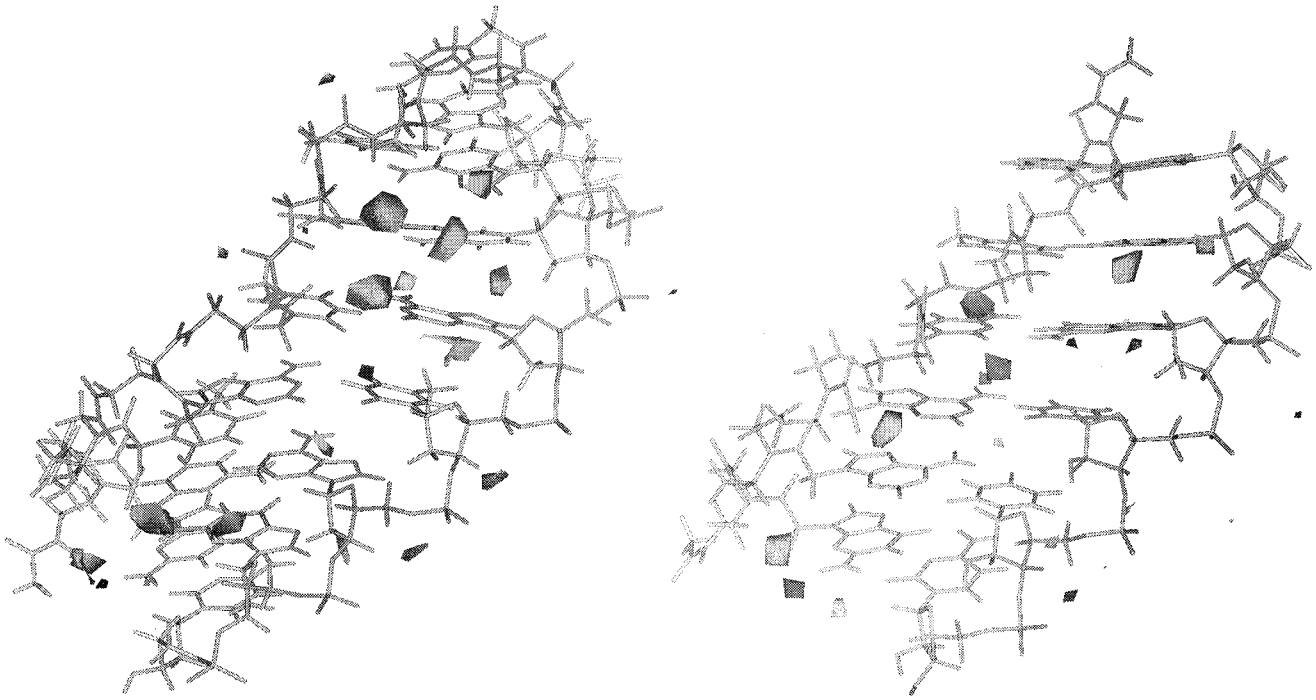

Figure 9. Hydration maps for the PNA·DNA (left) and PNA·RNA (right) duplexes. Contour levels correspond to 2.5 times the density of bulk water.

is quite well hydrated, just slightly worse than a DNA duplex, which is obviously due to the reduced charge density of the PNA·DNA duplex. In any case, the minor groove of the PNA·DNA duplex (especially in A·T regions) is the region with the highest apparent density of water, while there are no regions of high water density in the major groove (see Figure 9). This situation is very similar to that found in DNA duplex in the B-type conformation,[39,69,70] but different from that found for A-type oligonucleotides,[39] where the better hydration occurs in the major groove.[39] We conclude, therefore, that despite structural similarities to the A-form DNA, the grooves of the PNA·DNA duplex have recognition characteristics more similar to those of a B-type DNA duplex than to those of an A-type structure. These findings suggest that the sugar puckering is more important for the determination of recognition characteristics in nucleic acid structures than the general structure of the helix.

(69) Dickerson, R. E. *Methods Enzymol.* **1992**, *261*, 67−110.

(70) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, D. L. *Biochemistry* **1998**, *37*, 8341.

The PNA·RNA structure shows less regions of high water density than the PNA·DNA duplex. The greatest differences occur in the minor groove, where the "spine" of hydration found in the PNA·DNA duplex does not exist in the PNA·RNA, where small regions of large water density appear in the major groove. It is then clear that, despite the general structural similarity between the PNA·DNA and the PNA·RNA duplexes, the hydration pattern of these two molecules is completely different.

It was clear that the averaging procedure used above might mask distinct hydration features around the tertiary amide group that differed between forward and backward conformations. The P-form trajectory of the PNA·RNA simulation was filtered to leave only those snapshots in which the amide group of PNA base 3 (A) was in the forward orientation, and the hydration analysis repeated. The same was done for filtered snapshots in which this base had the backward orientation. This base was chosen because it showed an approximately equal population of both states. The hydration analysis showed that in the forward orientation the secondary amide carbonyl group is associated with a highly ordered water molecule, which appears to bridge the H1′ of the amide and N3 of the adenine base. In contrast, when the carbonyl group adopts the backward orientation there is no such hydration site (see Figure 10). We must note that even though only PNA·RNA results are shown here, very similar results were obtained for the PNA·DNA duplex.

The orientation of amides in a peptidic helix is a consequence of two opposite effects: (i) the alignment of amide dipoles antiparallel to the helix dipole generated by the other amide groups, and (ii) H-bonds and conformational constrains which prevent this. In the canonical structures of PNA·RNA and PNA·DNA helices, the alignment of dipoles might generate a helix dipole similar to that of pure peptides. We may therefore expect that this helix dipole would favor certain arrangement of the flexible secondary amide groups. To analyze this we have estimated the change in dipole−dipole interaction energy that is associated with an individual secondary amide group changing from the forward (parallel to the net helix dipole) to backward (antiparallel to the helix dipole) orientation. For this purpose, all partial charges in models of the PNA·DNA helix (very similar results are obtained for the PNA·RNA duplex) were set to zero, with the exception of the amide group atoms, which retained their standard AMBER values. In a model of the system in which all the amide groups were aligned, the dipole−dipole interaction energy in a *vacuum* between an individual secondary amide group and the rest of the helix was around +4 kcal/mol. When the orientation of the amide group was reversed, the calculated interaction energy was around −6 kcal/mol. Thus, in a field of otherwise aligned amide dipoles, the adoption of the forward orientation by an individual secondary amide is associated with an energy penalty of over 10 kcal/mol in the gas phase. We should note however that a large part of this energy difference should be compensated by the screening effect of water and DNA which reduces the strength of dipole−dipole interactions and by the better hydration of the structures with forward orientation of the secondary amide (see above).

Classical molecular interaction potentials (MIPs[39,48,51,53−55]) were used to determine the ability of the hybrids to recognize small cationic drugs. Analysis of MIP maps for the PNA·DNA duplex (see Figure 11) demonstrates that the phosphates are the regions yielding the most favorable interactions, followed by the minor groove in A·T regions. Comparison with duplex DNA of similar sequence allows us to identify two major differences in terms of MIPs. First the PNA·DNA duplex is clearly less attractive to a small cation than the DNA duplex

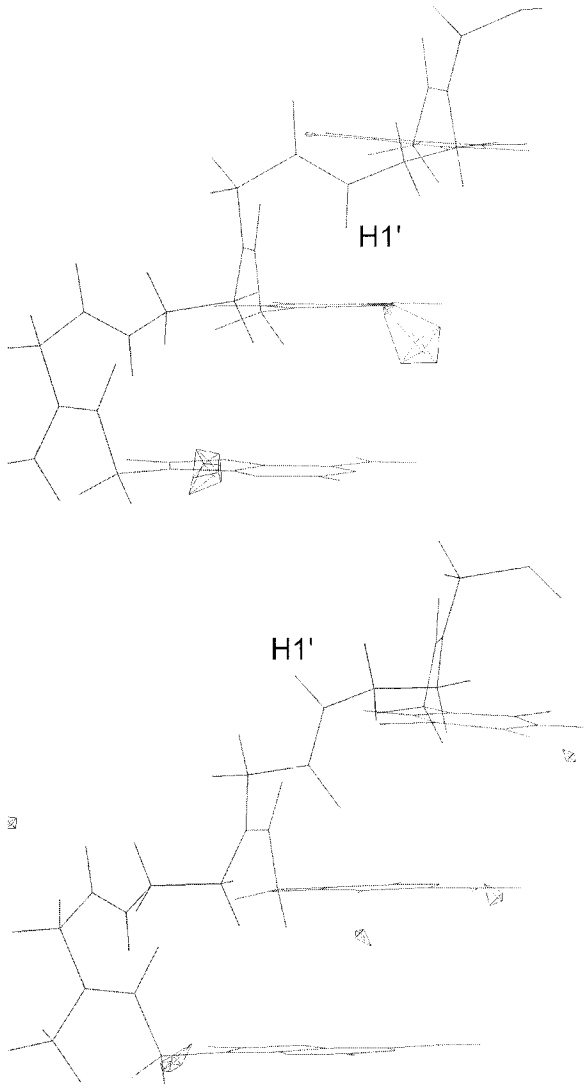

Figure 10. Diagram showing the hydration of PNA base 3 in the PNA·RNA simulation. When the amide group of this base adopts the "forward" orientation (top) a high-occupancy water site is evident, whereas no such site is evident when the amide group adopts the "backward" orientation (bottom).

(MIP values are around 2 kcal/mol smaller, in absolute terms, for PNA·DNA compared to duplex B-type DNA, see ref 39). Second, the minor groove is no longer the region of highest affinity, as is the case in the duplex B-type DNA. Overall, the MIP reflects the replacement of the phosphodeoxyribose backbone by a neutral peptide.

The analysis of the MIP for the PNA·RNA duplex shows also the phosphates as the regions yielding the most favorable interactions with a small cationic probe, followed by a G·C region of the major groove. The minor groove has a poor ability to interact with a cationic probe, in clear contrast with the behavior found for the PNA·DNA duplex. This shift in reactivity from minor to major groove is typical of B→A transitions,[39] and agree with the different pattern of hydration of both structures discussed above.

It is very remarkable to note the large discrepancy in recognition pattern found between PNA·DNA and PNA·RNA structures (see Figures 9 and 11), despite their global structural similarity (see above). It can be concluded that the different puckering of riboses is then the main responsible of the different recognition pattern displayed by the grooves of these two hybrid structures.

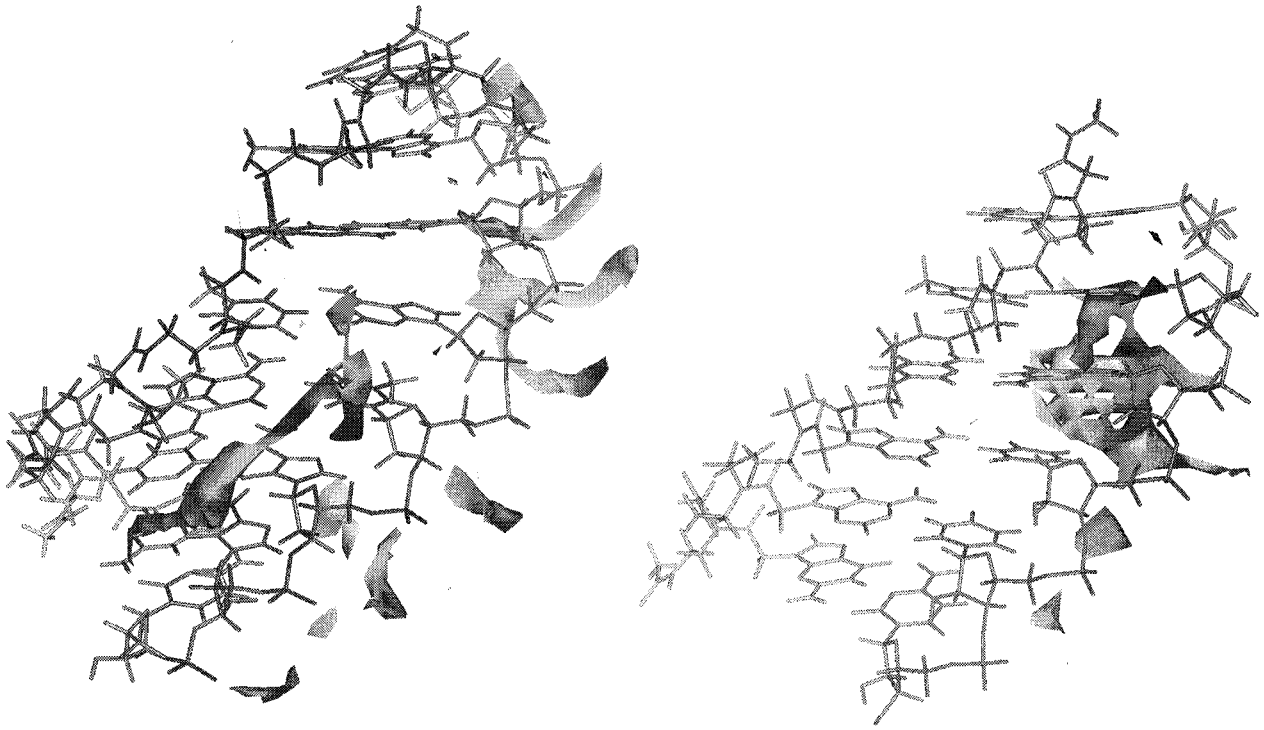

Figure 11. Molecular interaction potentials (MIPs) for the PNA·DNA (left) and PNA·RNA (right) duplexes. Contour levels correspond to −3.5 kcal/mol. The O^+ is used as probe molecule for MIP calculations.

Is The PNA Backbone Conformationally Neutral? MD simulations have demonstrated that the PNA backbone is much more flexible than the DNA or RNA backbones. Therefore, one can expect that the PNA backbone should be able to fit the structure of more rigid DNA or RNA backbones, which should have similar structures in the hybrids and in the homopolymers. The PNA·RNA duplex follows this prediction, since the backbone of the RNA strand in the hybrid is not much distorted from the conformation found in a RNA·RNA duplex, as expected from the similarity between the P- and A-forms of the hybrid. However, the prediction completely fails for the PNA·DNA duplex. In this case the PNA strand adopts a structure closer to its conformation in the homoduplex (PNA·PNA helix), than the DNA strand is to its conformation in canonical B-DNA (see above). The question is then: if the PNA is so flexible, what is the reason for the PNA-mediated change in the conformation of the DNA?

Analysis of the first picoseconds of the three trajectories for the PNA·DNA duplex reveals that the transition from B- to P-form is related to the low stability of the PNA backbone in the B-like structure (see Figure 12). On the contrary, the small transition from an A-type conformation to P-form is related to a subtle reduction of the backbone energy of the DNA strand (see Figure 12). No significant changes are obtained during the transitions for stacking interactions, in agreement with previous findings in the PNA·DNA·PNA triplex.[53] The same findings are also clearly shown by inspection of the energy profiles obtained in restrained MD simulations, where the nucleobases were restrained to their B-, A-, or P-like positions, while the backbones were completely free to move. Results collected in Table 3 show that when the PNA·DNA duplex is forced to be in a B-like conformation the PNA backbone is unstable, while if the structure is in the A-form, the DNA backbone is unstable. There are no statistically significant variations in the stacking energies of the three forms, and the DNA backbone energy of B- and P-forms and the PNA backbone energy of A- and P-forms, are very similar, which confirms the statistical value of the averages shown in Table 3. These results suggest that the internal energy of the PNA backbone is responsible for the distortion of the DNA backbone occurring in PNA·DNA duplexes.

MD simulations show that the PNA is flexible, but only if the conformational changes in the backbone are concerted in such a way that they do not lead to changes in the positioning of the nucleobases. Analysis of the experimental structures[8,36−38] of the PNA·PNA, DNA·DNA, and PNA·DNA duplexes shows that the twist of the DNA·PNA heteroduplex (around 28° from NMR results), lies between the twist values of the corresponding homoduplexes (36° for DNA·DNA and 20° for PNA·PNA duplexes). In other words, despite the greater flexibility of the PNA strand, the structural change induced by the DNA on the PNA strand is not greater than that induced by the PNA on the DNA strand.

Analysis of energy profiles collected during unrestrained and restrained A-, B-, and P-trajectories suggest that, despite its apparent plasticity, PNA has very clear conformational preferences for the positioning of the nucleobases and demands major changes in the conformation of the less flexible DNA backbone. Although less dramatic, changes to the conformation of an RNA strand in a PNA·RNA duplex are also imposed by the PNA. Our results suggest therefore that the formation of the PNA·DNA duplex involves a certain energetic cost related to the distortion of the DNA backbone, which might be reduced by other polymers with a relaxed structure closer to that of a B-DNA duplex.

Conclusions

Nanosecond molecular dynamics simulations using the AMBER-95 force field, periodic boundary conditions and the PME technique to introduce long-range effects are able to reproduce the structures of PNA·DNA and PNA·RNA duplexes as determined by NMR. The radius of convergence of the method is wide, although it can fail when the starting model is severely in error (the B-form model for the PNA·RNA duplex).

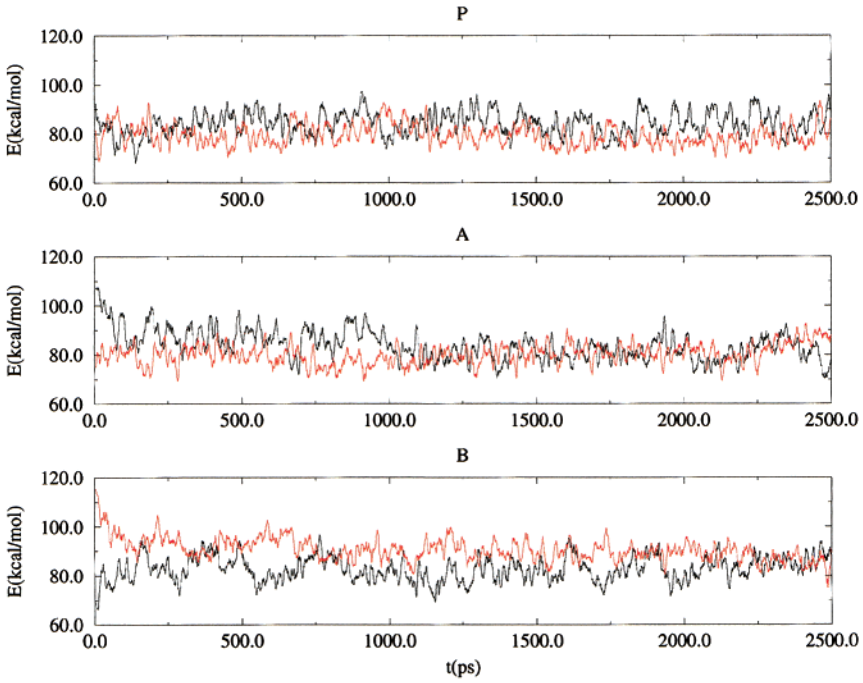

Figure 12. Internal energies (in kcal/mol) of the PNA (red) and DNA (black) backbones for the three simulations.

Table 3. Average Values of the DNA Backbone, PNA Backbone, and Nucleobase-Stacking Contributions to the Stability of PNA·DNA Duplexes in Restrained MD Simulations Forcing the Duplex To BE in P-, B-, or A-type Conformation[a]

conformation	DNA backbone	PNA backbone	stacking
P-	77.4 ± 7.6	83.9 ± 7.3	−412.2 ± 8.1
A-	94.4 ± 8.8	82.5 ± 7.6	−418.7 ± 7.8
B-	73.6 ± 6.8	100.4 ± 8.6	−415.9 ± 9.3

[a] All of the values are in kcal/mol. Standard deviations are shown.

It is found that the PNA·DNA duplex, despite the *south* type puckering of the sugars, adopts a general structure much more similar to an A-type duplex than to a B-type duplex. However, recognition characteristics of the PNA·DNA duplex are more similar to those of a B-type DNA than to those of an A-type duplex. It is clear then, that classification of PNA·DNA helices into one of the standard families of oligonucleotide helices is difficult and should be made with caution.

In contrast, the PNA·RNA duplex shows many of the features of a conventional RNA·RNA duplex, the major difference being a reduced helical twist. Interestingly, recognition patterns of the PNA·RNA duplex are qualitatively similar to those found in the RNA·RNA duplex (the A-form).

The PNA backbone is much more flexible than the DNA or RNA backbones. However, such flexibility does not lead to dramatic fluctuations in the helical parameters of the duplex, since rotations around PNA backbone bonds do not lead to changes in the positions of the PNA nucleobases. However, despite its flexibility the PNA backbone seems to have clear conformational preferences, which are not clear in terms of individual dihedral angles, but which are evident in term of helical parameters. These preferences are the driving force that leads to the important modifications of the structure of the less flexible DNA backbone during the formation of the PNA·DNA duplex.

Our results suggest that there is room for improvement in PNA structures by designing more rigid backbones, which would reduce the energetic cost of pre-organization, and by designing different PNAs adapted to bind specifically DNA or RNA strands in antigene or antisense experiments.

Acknowledgment. This work has been supported by the Fullbright Commission (E.C.S.), the Wellcome Trust (C.A.L., M.O.), the Catalan CIRIT (R.S.), and the Spanish DGICYT (PB96-1005 and PB98-1222). We thank Drs. C. González and J. L. Asensio for many useful discussions. We also thank the Centre de Supercomputació de Catalunya (Molecular Recognition Project) for computational support.

JA000259H