Molecular Dynamics of Duplex Systems Involving PNA: Structural and Dynamical Consequences of the Nucleic Acid Backbone

Srikanta Sen and Lennart Nilsson*

Contribution from the Centre for Structural Biochemistry, Karolinska institutet, Department of Bioscience, S-141 57 HUDDINGE, Sweden

Received July 7, 1997

Abstract: Molecular modeling and molecular dynamics simulation studies have been performed on homo- and heteroduplexes involving peptide nucleic acids (PNA) in aqueous solution under periodic boundary conditions. PNA is a DNA analogue that is homomorphous to DNA, but has an electrically neutral pseudopeptide backbone. In the present study we have investigated the structure and dynamics of duplex systems involving PNA in aqueous solution and how the overall structural and dynamical features of a double helix depend on the nature of the backbones of the constituent strands. Four different duplex systems have been studied: (i) PNA-PNA duplex (1.15 ns), (ii) PNA-DNA antiparallel duplex (0.64 ns), (iii) PNA-DNA parallel duplex (0.6 ns), and (iv) DNA-DNA duplex (0.64 ns). Comparison of the structural features obtained from this study on PNA-DNA antiparallel and PNA-PNA duplex systems with those obtained from NMR and X-ray crystallographic studies respectively has shown very good agreement. In all the cases the structures were stable over the entire period of simulations and the results indicate that the complementary bases and a backbone homomorphous to DNA are sufficient to maintain a stable double helix. The antiparallel PNA-DNA duplex and the PNA-PNA duplex have average structures between A- and B-helixes with certain A-like features while the parallel PNA-DNA double helix, as predicted in this study, is more close to the B-helix. No major difference in the geometries and dynamics of the base pairs in the different duplexes was found. However, the helicoidal parameters are found to be different for the different duplexes. These indicate that the actual structure is determined by the base pairing and the base stacking with the backbones causing some perturbations to this basic structure. The internal dynamics in the base linker region shows highly restricted motions even in the PNA strands where there is no ribose ring.

Introduction

Peptide nucleic acid (PNA) is a recently developed DNA analogue. A detailed investigation of the structural and dynamical properties of PNA is of considerable interest due to its great potential for use as a gene targeting drug in gene therapy.[1a,b] From a chemical point of view, a PNA molecule is a DNA strand in which the sugar−phosphate backbone has been replaced by a structurally homomorphous pseudopeptide chain consisting of N-(2-aminoethyl)glycine units (Figure 1a)[1a,b,2a,b] with successive monomeric units in the PNA strand connected by peptide bonds to form a polymeric PNA. As a result, PNA and DNA have striking similarities as well as dissimilarities. Both PNA and DNA contain the same types of ordinary nucleobases and a PNA backbone and a DNA backbone that are homomorphous to each other, but while the DNA backbone contains a sugar ring and negatively charged phosphates, the PNA is electrically neutral and connects the base to the main backbone through a short chain linker. These similarities and differences make PNA and DNA a unique system for addressing the interesting and fundamentally important question of how the structural and dynamical properties of the resulting duplex are affected by these features.

Properties of PNA have been probed by different physical techniques such as thermal melting, CD-spectroscopy, calorimetry, NMR spectroscopy, and very recently X-ray crystallographic methods.[2a,b,3−13] It has been clearly demonstrated that PNA strands with complementary base sequences can form a stable homoduplex and adopt a double helical structure with

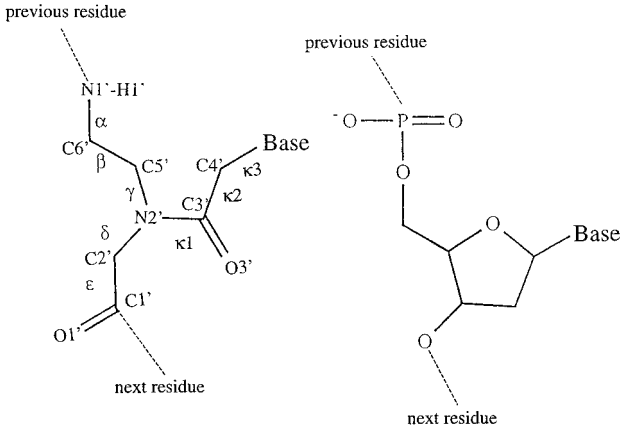

Figure 1. Schematic representation of (a) the chemical structures of the monomeric units of PNA and DNA and (b) the base sequences of the different duplex systems studied by MD simulation.

* To whom correspondence should be addressed.

(1) (a) Nielsen, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 167−183. (b) Eriksson, M. E.; Nielsen, P. E. *Q. Rev. Biophys.* **1996**, *29*, 369−394.

S0002-7863(97)02234-8 CCC: $15.00
Published on Web 01/20/1998

Watson−Crick type base pairing between them.[7,13] PNAs being intrinsically achiral are capable of forming both right-handed and left-handed double helixes.[7,13]

It is also found experimentally that a PNA strand can form a heteroduplex with a nucleic acid fragment (DNA/RNA) with complementary base sequence and adopts a double helical structure with Watson−Crick type base pairing between the strands.[1a,b,2a,b,3] A PNA-DNA(/RNA) heteroduplex is thermally more stable than the corresponding DNA-DNA or RNA-RNA homoduplexes.[1a,b,2a,b,7] A PNA strand can bind to its complementary DNA strand both in an antiparallel, where the N-terminus of the PNA strand is oriented toward the 3′ end of the complementary DNA strand, and in a parallel fashion.[1a,b,2a,b] Interestingly, the antiparallel complex has been found to be more stable than the parallel one.[1b,2a] The thermal stabilities of the different homo- and heteroduplexes involving PNA and DNA is the following order:

PNA-PNA > PNA-DNA (anti) > PNA-DNA (para) > DNA-DNA

Thus, all duplexes involving PNA are thermally more stable than the DNA-DNA duplex.

The ability of PNA to form stable duplex structures with DNA strands with higher affinity compared to the DNA-DNA duplex allows the PNA to bind to the DNA by strand invasion, making PNA a potential gene targeting drug.[1a,b] Hence, the characterization of the structure and dynamics of complexes involving PNA in solution is also very important for better understanding and improving the possible role of PNA as a gene targeting drug. The purpose of the present study is thus 2-fold: (i) to characterize the structural and dynamical features of duplexes involving PNA in aqueous solution and (ii) to get some insight in understanding the effects of backbone on the structure and dynamics of the resulting duplexes.

Molecular dynamics (MD) simulation is a powerful method for obtaining dynamical and structural information of macromolecular systems in atomic detail.[14a−c] In the present work, molecular modeling and MD simulation techniques have been used to probe the above questions. We have carried out a systematic study of unrestrained molecular dynamics simulations of duplexes with the base sequences shown in Figure 1b in (i) PNA-PNA, (ii) PNA-DNA (antiparallel), (iii) PNA-DNA (parallel), and (iv) DNA-DNA systems in aqueous solution, starting from B-like conformations and allowing the systems to relax toward stable structures. To our knowledge, the present report is the first detailed MD simulation done on duplex systems involving PNA in aqueous solution. In a few early works,[15a−c] some molecular mechanics studies on PNA-RNA and PNA-DNA heteroduplex (antiparallel binding) were made but only in vacuo energy minimization of the starting model was done, and no MD simulation was attempted at all. Only recently has an NMR derived solution structure of an octamer PNA-DNA antiparallel duplex been reported,[11] and the crystal structure of a hexamer PNA-PNA duplex has been presented very recently.[13] The corresponding average structures from our simulations, which were completed before the X-ray crystallographic data on a PNA-PNA duplex appeared early this year, are very similar to the experimental structures of the PNA-DNA antiparallel duplex and of the PNA-PNA duplex. The present molecular modeling and MD results also predict the structural features of a PNA-DNA parallel duplex, as well as providing data about the internal dynamics of the duplexes in all three cases.

Methods

(i) Topology, Parameter, and Partial Atomic Charge Setup. PNA molecules are not common biomolecules and hence the topology and parameters for PNA residues are not directly available in standard molecular mechanical and molecular dynamics packages such as CHARMM.[16] We have therefore prepared these relevant parts for PNA (all atoms including nonpolar hydrogens) for use in CHARMM. In doing so, we have assigned the atom types of the atoms involved in PNA residues following CHARMM[16] atom type definitions. This makes most of the bond, angle, and dihedral and all of the nonbonded parameters directly available from the combined CHARMM all-hydrogen parameter set for nucleic acids and proteins.[17] Those which are not available were obtained by comparison with similar groups in the CHARMM parameter set (a complete list is given in the Appendix, see Supporting Information). In the PNA backbone, there are several local planar regions (see Figure 1) mainly arising from the sp^2 hybridization of the carbon atom involved. The planarity of these regions is ensured by improper dihedral angle specifications (see Table A1 in the Supporting Information).

The partial atomic charges for the atoms in PNA residues were estimated by the ESP (ElectroStatic Potential) method in the MOPAC 6.0 package[18] by fitting the partial atomic charges to reproduce the electrostatic potentials at 1448 space points around the molecule. The necessary semiempirical quantum mechanical calculations were done with the AM1 parameter set. Thus, we obtained the partial atomic charges for the entire residue including the backbone and the base. However, we found some differences in the charges of the backbone atoms which are common to all four PNA residues for different conformations and different bases attached to it. In each case the backbone and the base separately contained small net charges, although the residue as a whole was electrically neutral. Moreover, the nucleobases are the same in DNA and PNA and there is only one bond connection between the base and the backbone. It may therefore be assumed that the charge distribution in the bases should not significantly depend on the nature of the backbone. On the basis of these

(2) (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497−1500. (b) 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) Leijon, M.; Gräslund, A.; Nielsen, P. E.; Buchardt, O.; Norden, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry* **1994**, *33*, 9820−9825.

(5) Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, F.; Bonham, M. A.; Au, K. G.; Carter, S. T.; Bruckenstien, D. A.; Boyd, A. L.; Nobel, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481−1485.

(6) Knudsen, H.; Nielsen, P. E. *Nucleic Acids Res.* **1996**, *24*, 494−500.

(7) Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Norden, B. *Nature* **1994**, *368*, 561−563.

(8) Wittung, P.; Nielsen, P. E.; Lyng, R.; Eriksson, M.; Norden, B. *J. Am. Chem. Soc.* **1995**, *117*, 10167−10173.

(9) Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994**, *265*, 777−780.

(10) Tomac, S.; Sarkar, M.; Ratilinen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Gräslund, A. *J. Am. Chem. Soc.* **1996**, *118*, 5544−5552.

(11) Eriksson, M.; Nielsen, P. E. *Nature Struct. Biol.* **1996**, *3*, 410−413

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

(13) Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nature Struct. Biol.* **1997**, *4*, 98−101.

(14) (a) McCammon, J.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: London, 1987. (b) Brooks, C.; Karplus, M.; Pettit, B. M. *Proteins: A theoretical Perspective on Structure, Dynamics and Thermodynamics*; Wiley: New York, 1988. (c) van Gunsteren, W. F.; Weiner, P.; Wilkinson, A. *Computer Simulations of Biomolecular Systems*; ESCOM: Leiden, 1993.

(15) (a) Almersson, Ö.; Bruice, T. C.; Kerr, J.; Zuckermann, R. N. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7518−7522. (b) Almersson, Ö.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9542−9546. (c) Torres, R. A.; Bruice, T. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 649−653.

(16) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus. M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(17) MacKerell, A. D., Jr.; Wirkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.

(18) Coolidge, M. B.; Stewart, J. J. P. *MOPAC 6.0*, 1990.

considerations, we have taken the partial atomic charges for the backbone atoms of the PNA residue with a guanine base and have used the partial atomic charges of the atoms in the same bases as for DNA bases in CHARMM.[17] In addition, the CHARMM charges keep the bases electrically neutral. So, we have slightly adjusted the computed charges (ESP) of the terminal atoms (C1′, O1′ and N1′, H1′) of the backbone of the PNA residue to make the backbone separately neutral. Thus the entire PNA residue as well as the backbone and the base separately remain electrically neutral and more consistent with the CHARMM parameters. We have compared the charges for the atoms in the base as obtained from MOPAC with those in CHARMM to see how compatible are the MOPAC charges with CHARMM charges. We found that although some charges are different, most of the charges, on the average, are not very different in these two independent charge sets, indicating that these two charge sets are reasonably compatible with each other. The charges on the peptide group obtained in this way are also similar (N1: −0.37; H1: 0.24; C1: 0.33; and O1: −0.25) to the CHARMM charges for peptide groups in proteins. The charge set for the PNA backbone atoms used here is given in Table A2 in the Supporting Information.

(ii) Choice and Preparation of Starting Models. It is quite clear that a double helix is stabilized mainly by (i) the horizontal interactions arising from the base pairing between the complementary bases and (ii) the vertical interactions arising from the stacking of the base pairs on top of each other, constrained by the covalent connectivity of the backbone. On the other hand, the electrostatic repulsion between charged backbones tends to destabilize the structure in the case of DNA. So, one can say that in a double helix, it is the interactions between the base pairs which actually determine the basic structure and the different backbones can only introduce some characteristic perturbations to this basic structure. PNA and DNA contains the same nucleobases and are structurally homomorphous to each other. Hence, it is quite logical to assume that a PNA-PNA or a PNA-DNA duplex may adopt structures close to the standard B-form of DNA with some structural perturbation induced by the PNA backbone. Thus, for such systems involving PNA, a structure similar to the B-form DNA seems to be a reasonable choice as a starting geometry.

The PNA backbone is superimposable on the DNA backbone, giving a one-to-one mapping of the PNA backbone atoms (non-H) onto the DNA backbone atoms. For an antiparallel PNA-DNA duplex, the mapping scheme is (DNA) O5′-C5′-C4′-C3′-C2′-C1′-O3′-+P-+O1P → N1′-C6′-C5′-N2′-C3′-C4′-C2′-C1′-O1′ (PNA). Thus, in generating a B-DNA-like helix involving PNA, we started with a DNA-DNA duplex in the B-form having the base sequence given in Figure 1b. Then we replaced the non-H atoms of the backbone of the respective DNA strand by the PNA backbone atoms following the above mapping scheme. However, this mapping cannot include the carbonyl oxygen atoms O3′ (in the base linker part). The coordinates of these atoms were generated by geometric calculation on their local topology and coordinates of their nearest connected atoms. The orientations of the C3′−O3′ bond were oriented toward the C-terminal of the respective PNA strand as obtained in NMR studies.[11] The coordinates of the H-atoms were generated by the HBUILD[19] facility in CHARMM.

The resulting coordinates were then energy minimized in vacuum for 300 steepest descent steps keeping the bases fixed in position. Then the restraints were removed and further energy minimization for 200 steepest decent steps was performed. This structure was then used for subsequent MD simulation in a water box.

The PNA-DNA parallel duplex was prepared in a similar fashion only with a different mapping scheme: (DNA) O5′-C5′-C4′-O4′-C1′-C3′-O3′-+P → C1′-C2′-N2′-C3′-C4′-C5′-C6′-N1′ (PNA). The PNA-PNA duplex was generated by replacing both the DNA strands of the DNA-DNA duplex, according to the antiparallel mapping scheme. The initial energy minimization protocol was the same for all systems.

(iii) Setup of Solvated Systems. In the systems involving the DNA strand, one Na^+ counterion per phosphate group was included to neutralize the system[20] (i.e., 9 Na^+ ions for a PNA-DNA heteroduplex and 18 Na^+ ions for a DNA-DNA homoduplex). Each Na^+ ion was placed at a distance 3.5 Å from the phosphorus atom of the respective phosphate group, on a line bisecting the line joining the two oxygen atoms of the phosphate group of the DNA strand. For PNA molecules no counterions were needed. The molecular system was then inserted in a rectangular water box of size 31.5 Å × 31.5 Å × 41.5 Å, containing 1288 preequilibrated TIP3P water molecules.[21] Water molecules which were closer than 2.8 Å from any atom of the solute molecule or counterion were deleted. This solvated system was then energy minimized for 500 steepest descent steps keeping the solute fixed in position to allow only the counterions and water molecules to reorient around the solute, and then the constraint on the solute was removed and energy minimization for another 500 Powell steps was done. The resulting systems were then used to perform the subsequent MD simulations.

Dynamic Simulation Methodology. All MD simulations were done by employing the program package CHARMM, version 25, with its standard empirical potential energy function.[16] In the simulations, Newton's equation of motion for each atom was integrated by using the leapfrog-Verlet algorithm[22a,b] with a time step of 2.0 fs. The SHAKE algorithm[23a,b] was applied to constrain the bond lengths involving hydrogen atoms to their equilibrium positions. Periodic boundary conditions with minimum image conventions were applied to calculate the nonbonded interactions. The nonbonded pair list was updated every 10 steps and the nonbonded interactions were smoothly shifted to zero at 11.0 Å. For electrostatic calculations a relative dielectric constant of 1.0 was used. It should be emphasized here that in all these simulations no direct experimental constraint was used.

In each simulation, the water box including the solvated solute was heated to 298 K during the first 2 ps and then equilibrated for 2 ps by assigning velocities to the atoms from a Gaussian distribution at 298 K. Then the simulation was continued with the temperature being checked every 100 steps and adjusted by scaling velocities only if the average temperature of the system was outside the 298 ± 10 K window. Thus, the average temperature was maintained around 298 K. The trajectory was saved every 200 steps for future analysis. Dynamic simulation was continued until the RMSD of the structure with reference to the starting structure reached a steady average value over a few hundred picoseconds. Thus, MD simulations on the PNA-PNA duplex were conducted for 1150 ps. Each of the DNA-DNA, PNA-DNA antiparallel, and PNA-DNA parallel duplex systems were simulated for about 600 ps or more. Structural and dynamical analysis for the PNA-PNA duplex system was made over the last 700 ps of the trajectory, and for the other duplexes the last 300 ps were used.

The correlation coefficients between pairs of dynamical quantities were computed to investigate the correlated motions present in the dynamics of the duplex systems. The linear cross correlation coefficient between two dynamical quantities "x" and "y" is given by the expression[24]

$$C_{xy} = \langle \Delta x_i \cdot \Delta y_i \rangle / [\langle \Delta x_i^2 \rangle \cdot \langle \Delta y_i^2 \rangle]^{1/2}$$

where $\langle ... \rangle$ denotes an average over the trajectory and $\Delta x_i = x_i - \langle x \rangle$ and $\Delta y_i = y_i - \langle y \rangle$.

From molecular dynamics trajectories the correlation function $C(t)$ for the orientational dynamics of a bond vector can be obtained from the relation

$$C(t) = \langle P_2(\cos\theta(t,\tau)) \rangle$$

where P_2 is the second-order Legendre polynomial and θ is the angle between the bond positions at initial time and after a period τ.

(19) Brunger, A.; Karplus, M. *Struct. Funct. Genet.* **1988**, *4*, 148−156.
(20) Jayaram, B.; Beveridge, D. L. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 367−394.
(21) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926−935.
(22) (a) Hockney, R. W. *Methods Phys.* **1970**, *9*, 136−211. (b) Potter, D. *Computational Physics*; Wiley: New York, 1972; Chapter 5.
(23) (a) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311−1327. (b) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(24) Press: W. H.; Teukolski, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipies in Fortran*, 2nd ed.; Cambridge University Press: Cambridge, p 630.

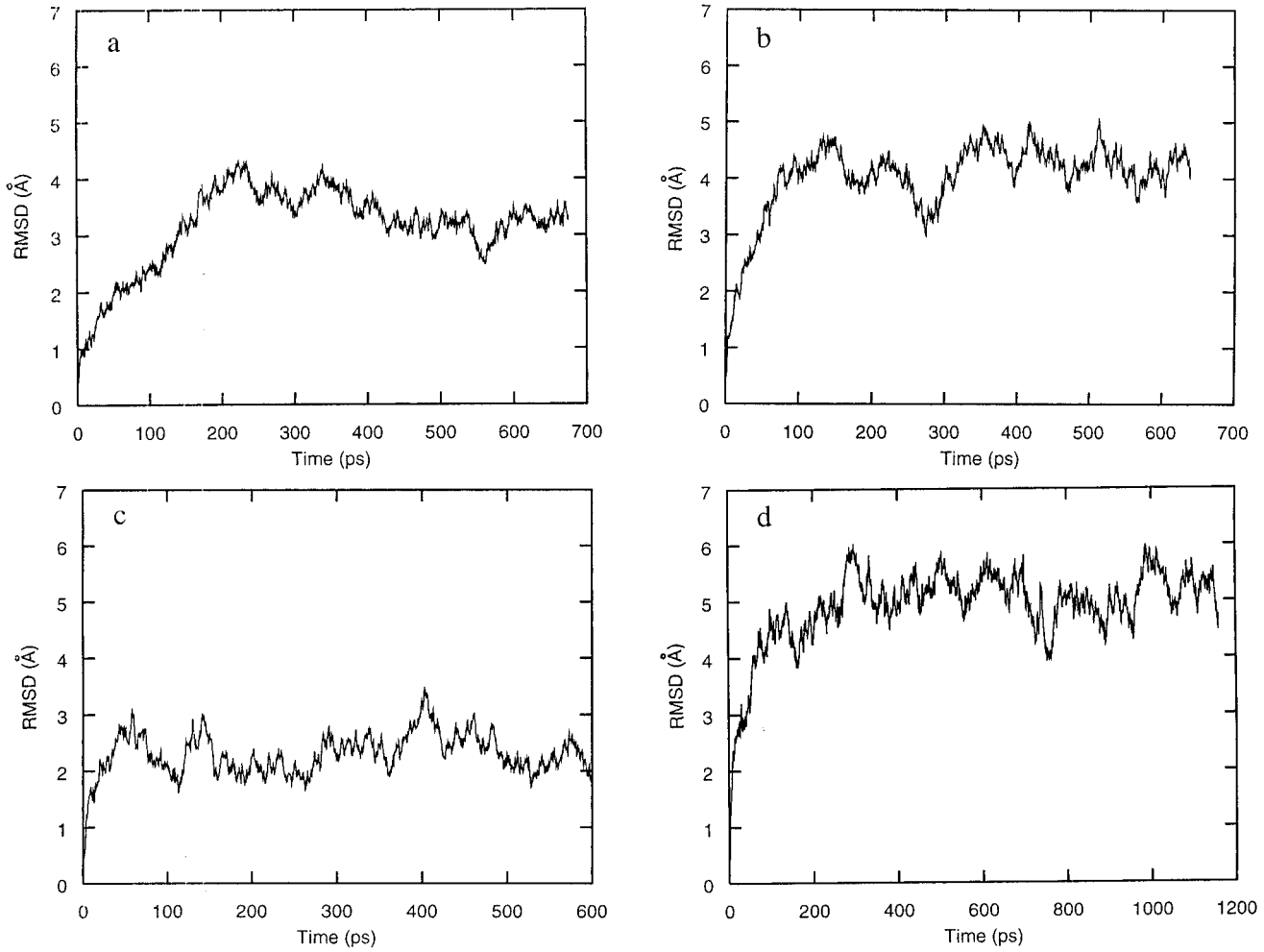

Figure 2. Time evolution of the RMSD values for the different duplex systems with respect to their starting structures: (a) the DNA-DNA duplex system, (b) the PNA-DNA antiparallel duplex, (c) the PNA-DNA parallel duplex, and (d) the PNA-PNA antiparallel duplex.

According to the model free approach of Lipari and Szabo,[25a,b] the correlation function for the internal dynamics of a bond can also be represented in terms of a generalized order parameter S^2 and an effective reorientational correlation time τ_e of the bond vector, through a relation

$$C(t) = S^2 + (1 - S^2)e^{-t/\tau_e}$$

The generalized order parameter S^2 describes the spatial restriction of the reorientational motion of the bond vector.

Results

All the dynamic simulations were found to be quite stable with the average temperature well maintained throughout the simulation periods. We have described the results for the individual systems as well as compared the results for all four duplex systems.

RMSD. First we have compared the overall structural features of the four different duplex systems by looking at the time evolution of their RMSD values with respect to their starting B-like double helical structures (Figure 2a−d). It is clearly seen that in each case the RMSD has reached a steady average value over a substantial period of the later part of the simulation, indicating that in each case a stable structure has been reached. However, the steady average RMSD values for the different duplex systems are different from each other, indicating that the final structures are different, although the initial model structures were the same (B-like) for all the systems. The plots also show that the RMSD value is least (∼2 Å) for the PNA-DNA parallel duplex, indicating that it is closer to the canonical B-form compared to the other duplexes. It is further interesting to note that the average RMSD value for the PNA-DNA parallel duplex is considerably different from the average RMSD of the antiparallel duplex, implying that their relaxed average structures are different from each other. It may be mentioned here that structural differences have also been inferred experimentally on the basis of CD spectra of these two duplexes.[3] The RMSD value is found to be maximum (∼5 Å) for the PNA-PNA duplex. This is also not very surprising because in a PNA-PNA duplex both the strands contain the unusual pseudopeptide backbone, which naturally introduces more structural perturbations into the starting canonical B-helix-like structure and thus makes the resulting relaxed average structure considerably different from the starting one.

Figure 3a,b shows the comparison of the RMSD values between the two PNA and DNA strands in the antiparallel and parallel PNA-DNA heteroduplexes, respectively. The comparison indicates that despite a substantial difference in the chemical compositions of the two strands (PNA and DNA), the time developments of their RMSD do not show significant differences in either case. Figure 3c represents the comparison between the time dependence of the RMSD values of the atoms in the backbones (dashed line) and the atoms in the bases (solid line) for the PNA-PNA duplex system. It is clearly seen that larger rearrangements occurred for the atoms in the backbone

(25) (a) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559. (b) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.

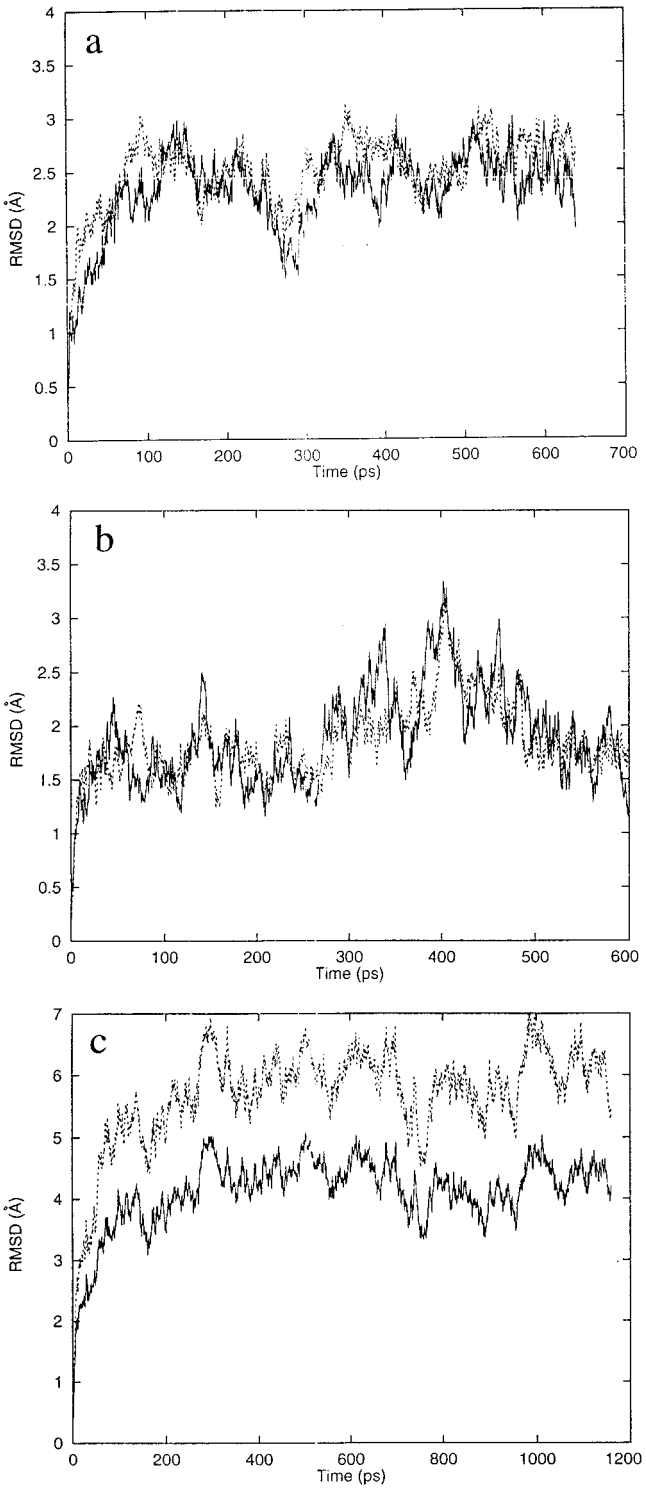

Figure 3. Comparison between the time evolution of the RMSD of the two PNA (solid) and DNA (dashed) partner strands in the (a) antiparallel and (b) parallel duplex systems. (c) Comparison of the time evolution of the RMSD values between the backbone atoms (dashed) and the atoms in the bases (solid) of the PNA-PNA duplex.

compared to that in the bases. This is consistent with our primary assumption that the basic nature of the overall duplex structure is mainly determined by the bases.

Geometry and Dynamics of Base Pair H-Bonds. Dynamic averages and the root mean-square (rms) fluctuations of the base pair H-bond geometries are given in Table A3 (in the Supporting Information), clearly indicating that all the base pairs are strongly H-bonded. The rms fluctuations of the dynamical averages of each bond length and bond angle over all the base pairs indicate the regular nature of these quantities over the entire duplex system. The averages of these properties among the four different duplex systems are strikingly similar, with slight differences found for the PNA-DNA parallel duplex. Thus it is found that the geometries and dynamics of the H-bonds in base pairs are not significantly dependent on the detail nature of the backbones.

The lengths of the two H-bonds in all the AT base pairs are correlated by a cross correlation coefficient value in the range 0.2 to 0.3, implying that the AT base pair exhibits a linear stretching−compressing motion, while the lengths of the two outer H-bonds in GC base pairs are anticorrelated with a correlation coefficient in the range −0.2 to −0.3, which is characteristic of inplane bending motion. Similar observation was also reported earlier.[26] Here we have found that these types of motion are quite general and do not depend on the nature of the associated backbone.

Backbone Torsional Angles and Their Dynamics. Table 1A represents the data on the PNA backbone torsional angles for the duplexes involving PNA. Here also the range of the dynamic averages and the rms fluctuations for the torsional angles indicate that the backbones are quite rigid. Similar behavior is also obtained for the torsional angles of the DNA backbones in the duplex systems containing DNA strands, as shown in Table 1B. However, the PNA strand in the PNA-DNA parallel duplex shows some differences in the backbone torsional values compared to the other duplexes, indicating that the relative orientation of the two partner strands in the heteroduplex has some effect on the overall base pair geometry and backbone torsional angles of the duplex.

In Figure 4 we have plotted the distribution of each torsional angle of the PNA backbone over different values during dynamics of the PNA-DNA antiparallel duplex. It is found that for all the torsional angles, there is only one preferred value about which the torsional angle fluctuates and the fluctuation occurs over a narrow range of angles (Figure 4a-h) excepting the case of torsional angle ϵ which fluctuates over a wider range of values (Figure 4e). Similar features were observed for the backbone torsional angles in the PNA strands involved in the other duplexes.

Most of the different torsional angles of the PNA backbone are anticorrelated to varying degrees, with correlation coefficients in the range 0 to −0.3. It may be mentioned that the torsional angles in a DNA backbone are also in general mostly anticorrelated.[27] However, there are a few torsional angle pairs, $\beta \leftrightarrow \delta$, $\beta \leftrightarrow \epsilon$, and $\gamma \leftrightarrow \epsilon$, which have a positive correlation coefficient in the range 0.2 to 0.3. It is also interesting to note that the dynamics of the torsional angle pair $\epsilon \leftrightarrow \zeta$ in the DNA strand of both the antiparallel and the parallel PNA-DNA duplex systems are positively correlated (correlation coefficient in the range 0.2 to 0.3) and $\alpha \leftrightarrow \gamma$ are anticorrelated (correlation coefficient in the range −0.1 to −0.2) while in the case of the standard DNA-DNA duplex both of these angle pairs are anticorrelated.

Sugar Pucker in the DNA Strand. It is found that although the general preference of sugar pucker is toward the C2′-endo, there is a substantial population of the C3′-endo conformation which also implies that the sugars in the PNA-DNA antiparallel duplex frequently visit the C3′-endo conformation and as a result the average structure possesses some A-helix-like features. The

(26) Harvey, S. C.; Prabhakaran, M.; Mao, B.; McCammon, J. A. *Science* **1984**, *223*, 1189−1191.
(27) Saenger, W. *Principles of nucleic acid structure*; Springer: New York, 1984.

Table 1. Comparison of the Range of Dynamical Averages (rms fluctuations) and the Averages (rms fluctuations) Over the Entire Duplex for the Backbone Torsional Angles of the PNA and DNA Backbones of the Three Different Duplexes Involving PNA and DNA Strands

A. PNA Backbone

torsion angle	range of dynamic av (antiparallel)	av over all residues (antiparallel)	range of dynamic av (parallel)	av over all residues (parallel)	range of dynamic av (PNA-PNA)	av over all residues (PNA-PNA)
α	169−176	174	142−178	158	171−180	173
	(21−30)	(3)	(16−39)	(11)	(24−35)	(3)
β	61−74	67	65−73	71	61−70	65
	(10−11)	(4)	(10−14)	(3)	(10−12)	(3)
γ	74−81	77	78−91	85	72−80	75
	(8−9)	(3)	(9−13)	(4)	(8−10)	(2)
δ	75−81	78	77−87	83	74−80	77
	(10−12)	(2)	(11−15)	(3)	(11−16)	(2)
ϵ	61−73	66	66−95	78	60−75	68
	(18−29)	(5)	(11−36)	(9)	(23−29)	(5)
*κ*1	−6 to −16	−11	−6 to −16	−13	−3 to −12	−8
	(12−15)	(3)	(13−15)	(4)	(13−15)	(3)
*κ*2	−169 to −177	−172	−107 to −138	−122	−167 to −177	−172
	(9−10)	(3)	(16−31)	(9)	(9−14)	(3)
*κ*3	92−104	98	−96 to −131	−115	94−103	98
	(11−13)	(4)	(13−26)	(12)	(11−13)	(3)

B. DNA Backbone

torsion angle	range of dynamic av (antiparallel)	av over all residues (antiparallel)	range of dynamic av (parallel)	av over all residues (parallel)	range of dynamic av (DNA-DNA)	av over all residues (DNA-DNA)
α	−67 to −78	−71	−66 to −81	−71	−64 to −78	−69
	(1−14)	(3.3)	(11−13)	(4.9)	(10−12)	(5.4)
β	149−165	160	133−165	152	161−170	169
	(1−19)	(7.4)	(9−21)	(9.8)	(10−16)	(5.9)
γ	52−60	57	51−59	55	54−61	59
	(9−10)	(2.5)	(9−10)	(2.1)	(8−111)	(3.3)
δ	102−123	115	109−136	126	81−107	86
	(15−24)	(7.3)	(7−23)	(10.6)	(5−16)	(2.9)
ϵ	−142 to −160	−150	−104 to −159	−139	−140 to −159	−149
	(11−23)	(6.9)	(10−29)	(16.2)	(8−14)	(7.2)
ζ	−84 to −149	−106	−80 to −165	−113	−74 to −95	−82
	(14−42)	(20.6)	(10−47)	(27.7)	(9−14)	(8.6)
κ	−122 to −142	−134	−109 to −133	−123	−129 to −175	−149
	(13−24)	(9.8)	(11−22)	(15.6)	(7−18)	(15.2)

distributions of the sugar pucker phase angle for each sugar moiety in the DNA strand of the PNA-DNA antiparallel duplex system are shown in Figure A1 (Supporting Information). On the other hand, in the case of the PNA-DNA parallel duplex, the sugar puckers are found mainly in the C2′-endo conformation. This is also consistent with the more B-like character in the parallel duplex.

Helicoidal Parameters, Minor Groove Width, Solvent Accessible Area and Bending Flexibility. Table 2A compares the structural features of the different duplexes in terms of the helicoidal parameters "twist", "roll", and "tilt" values of their average structures (of the central eight base pairs) calculated by the program Curves 5.1.[28] The PNA-PNA duplex has an average twist value of 23.3° (±1.9) while the DNA-DNA duplex has a twist of 32.3° (±1.1), indicating that the PNA-PNA duplex is considerably underwound with reference to the DNA-DNA duplex. The values in parentheses indicate the fluctuation of twist over the central eight base pairs. The PNA-DNA antiparallel duplex has an intermediate value of 26.3° (±3.9) as a result of the partial influences of both the PNA and DNA partner strands. However, the PNA-DNA parallel duplex has a twist value of 33.7° (±8.9) which is close to that of the DNA-DNA duplex but is associated with larger fluctuation over the entire duplex structure, indicating a strong sequence dependence or irregularities. Comparison of "roll" and "tilt" as shown in Table 2A, does not show significant differences between the different duplexes. The average values of the "rise" and "inclination" values for the different duplexes are given in Table 2B, which also summarizes some differences among the different duplexes.

Table 2B also compares the width of the minor groove in the four types of duplexes and shows that both PNA-PNA and PNA-DNA antiparallel duplexes have narrower minor grooves compared to the other two duplexes. A narrowing of the minor groove in the duplex involving PNA is not very unlikely because of the absence of the strong electrostatic repulsion between the backbones as it occurs in a DNA-DNA duplex system.

Since composition wise a PNA backbone seems to be more hydrophobic than a DNA backbone, we have compared the solvent accessible surface areas[29] of the average structures in each case. Table 2B indicates that the PNA-PNA duplex has a slightly smaller solvent accessible area compared to the other duplexes which is consistent with the above fact.

To get some qualitative idea about the relative bending flexibilities of the different duplex systems, we have looked into the time evolution of the end-to-end distance of the first strand of each duplex. Figure 5a−d represents the corresponding data over the last 400 ps of the respective dynamic trajectories in each case. Comparison of the patterns of the time evolution of the end-to-end distance in the different cases clearly indicates that it is much more steady in the case of the DNA-DNA duplex (Figure 5d) compared to that in the case of the PNA-PNA duplex (Figure 5a), implying larger bending flex-

(28) Ravishankar, G.; Swaminathan, S.; Beveridge, D. L.; Levery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 669−699.

(29) Lee, B.; Richard, F. M. *J. Mol. Biol.* **1971**, *55*, 379−400.

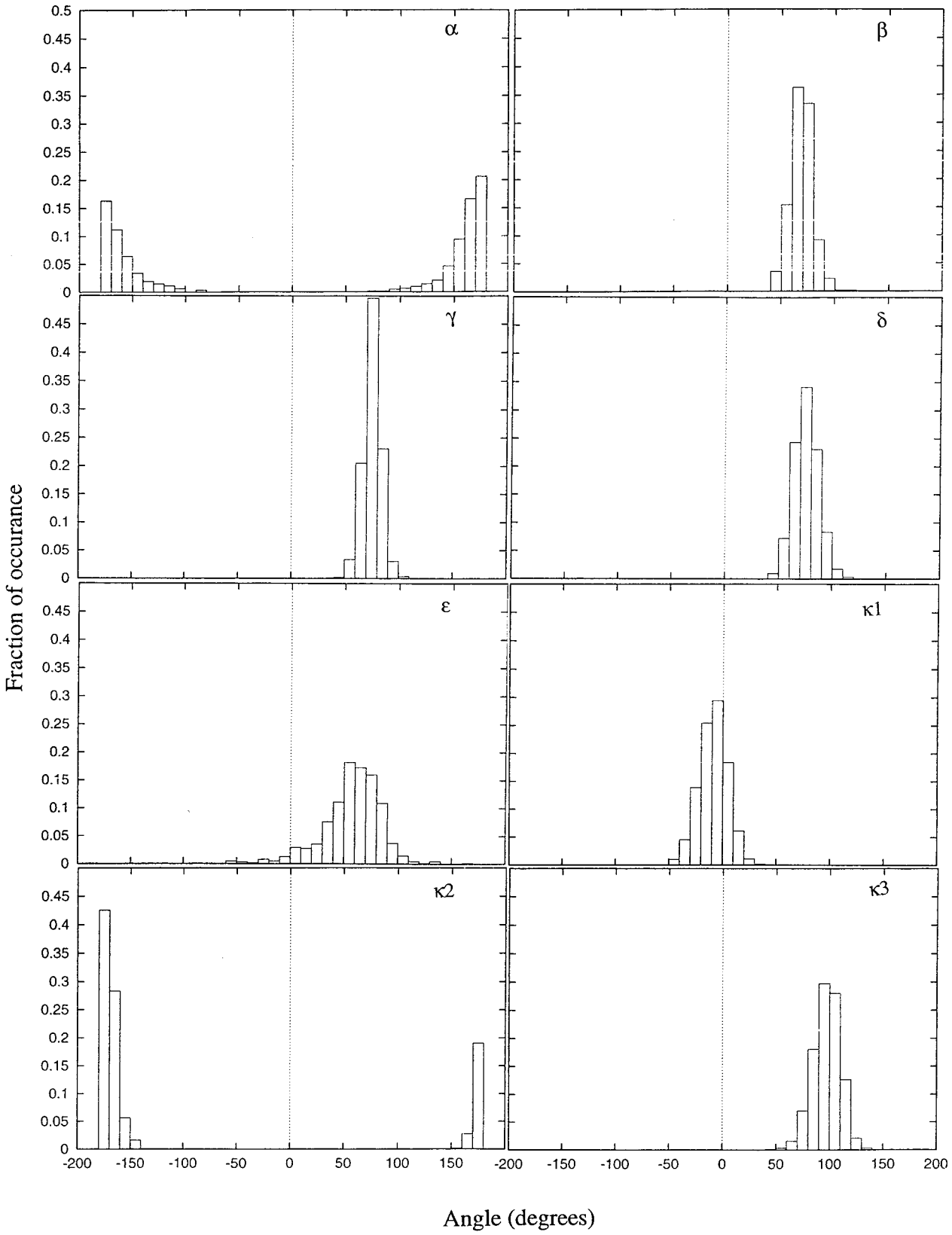

Figure 4. Probability distributions of the different backbone torsional angles of the PNA backbone in the antiparallel PNA-DNA duplex.

ibility of the PNA-PNA duplex compared to the DNA-DNA duplex. Figure 5b and 5c also indicate trends of higher bending flexibility, but lower than the PNA-PNA case. Thus, in general, duplexes involving PNA strands show higher bending flexibility.

Internal Dynamics of C−O and N−H Bond Vectors. As a fundamental difference between a PNA and a DNA molecule is the absence of any ring structure in the base linker region of the PNA molecules, it is of interest to investigate the internal dynamics of this region. For this, we have calculated the correlation function for the carbonyl bonds C3′−O3′ in the base linker region and the C1′−O1′ in the main backbone for comparing the effect of the local atomic arrangements. Figure 6a shows three curves representing typical correlation function plots for the C3′−O3′ bond in the base linker region (solid line)

Table 2. Comparison of the Helicoidal Parameters, Rise, Inclination Parameters, Minor Groove Widths, and Solvent Accessible Surface Area for the Average Structures of the Different Duplex Systems

A. Helicoidal Parameters

	twist				roll				tilt			
base pair	ppd	anti	para	dna	ppd	anti	para	dna	ppd	anti	para	dna
G2−C19	23.7	16.6	44.8	33.7	−0.1	10.4	−22.6	2.1	−0.6	2.5	−17.9	−1.3
T3−A18	22.1	28.8	22.7	32.7	−0.6	5.0	17.0	19.9	−0.7	−0.9	5.2	−1.9
G4−C17	26.4	27.6	33.9	32.1	0.8	5.9	10.7	9.0	0.3	−4.7	−3.0	7.3
A5−T16	23.9	26.4	26.6	32.5	0.7	2.1	11.4	−7.2	−0.8	−0.5	0.1	−5.5
T6−A15	22.4	29.3	45.4	32.3	−2.8	−3.6	−15.5	12.7	0.7	1.4	−3.3	1.9
C7−G14	21.7	27.5	29.2	32.6	−4.5	7.2	−2.5	−4.0	−0.8	−2.9	4.2	−1.8
T8−A13	26.2	25.4	23.8	29.8	−2.2	3.8	23.4	13.7	0.2	3.1	−1.5	−4.7
A9−T12	20.5	28.7	43.2	32.1	13.4	2.5	−24.6	4.5	0.5	1.4	−14.1	−0.4

B. Rise, Inclination Parameters, Minor Groove Widths, and Solvent Accessible Surface Area

duplex system	ris (Å)	inc (deg)	minor groove width (Å)[a]	solvent accessible surface area (Å^2)
PNA-PNA	2.9 ± 0.1	14.9 ± 0.6	10.1	3949
PNA-DNA antiparallel	3.1 ± 0.3	4.6 ± 1.4	10.2	4054
PNA-DNA parallel	3.4 ± 0.3	13.9 ± 1.8	12.4	4056
DNA-DNA	3.2 ± 0.3	−0.6 ± 1.4	13.2	4020

[a] The minor groove width of the PNA-PNA duplex was measured as the average distance between the C5′ atom of the *i*th residue of one strand and the N2′ atom of the (13 − *i*)th residue of the other strand. For both the PNA-DNA antiparallel and parallel duplexes the atom pairs C5′ and C5′ were used and for DNA-DNA duplex the standard P and P atoms were used for this purpose.

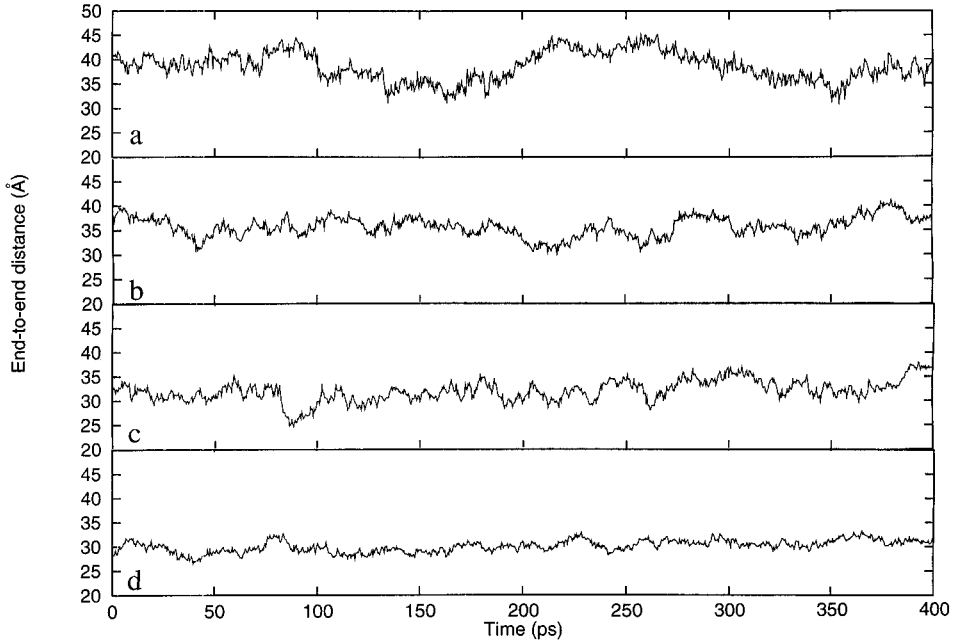

Figure 5. Time evolution of the end-to-end distance of the first strand of the different duplexes over the last 400 ps dynamics trajectory: (a) PNA-PNA duplex, (b) antiparallel PNA-DNA duplex, (c) parallel PNA-DNA duplex, and (d) DNA-DNA duplex.

and the C1′−O1′ bond (dashed line) and the N1′−H1′ bond (doted line) in the main backbone chain of the PNA strands in a PNA-PNA duplex.

Figure 6b represents the plot of the generalized order parameter (S^2) values of the C3′−O3′ bonds (◇, +) in the two PNA strands and for the C1′−O1′ bonds (□, ×) against residue number. The S^2 values for the bond C3′−O3′ are highly regular over the residues at around 0.85, indicating a highly restricted reorientational motion. On the other hand, the order parameter values for the C1′−O1′ bonds are less regular over the entire double helix with an average value close to 0.6, implying that here the motion is less restricted. The highly restricted motion of the C3′−O3′ bond may result from the steric hindrance of the local atoms in the solute and some weak interaction with the backbone amide group. On the other hand, the C1′−O1′ bonds in the main backbone chain are completely exposed to the surrounding solvent and get more reorientational freedom. However, the important fact is that even in the absence of any ring structure in the base linker region, the local internal dynamics is highly restricted. The small rms fluctuation of the local torsional angles $\kappa 1$, $\kappa 2$, and $\kappa 3$ (Table 1A) also support this fact. The N1′−H1′ bond vector also shows order parameter value similar to that of the C1′−O1′ bond vector as they are parts of the same peptide bonded rigid region.

Figure 6c represents the plot of the effective reorientational correlation times τ_e for the above three bond vectors. Comparison of the values of the different bonds indicates that although the restriction on the reorientational dynamics is different for the different bonds, the corresponding values of τ_e for them are not very different. The τ_e values indicate very fast decay of motion. However, the estimation of τ_e is sensitive to the accuracy of the value of S^2, which is estimated as the average value of $C(t)$ over the long time plateau region. Thus the S^2 value depends on the statistical behavior of $C(t)$ over the

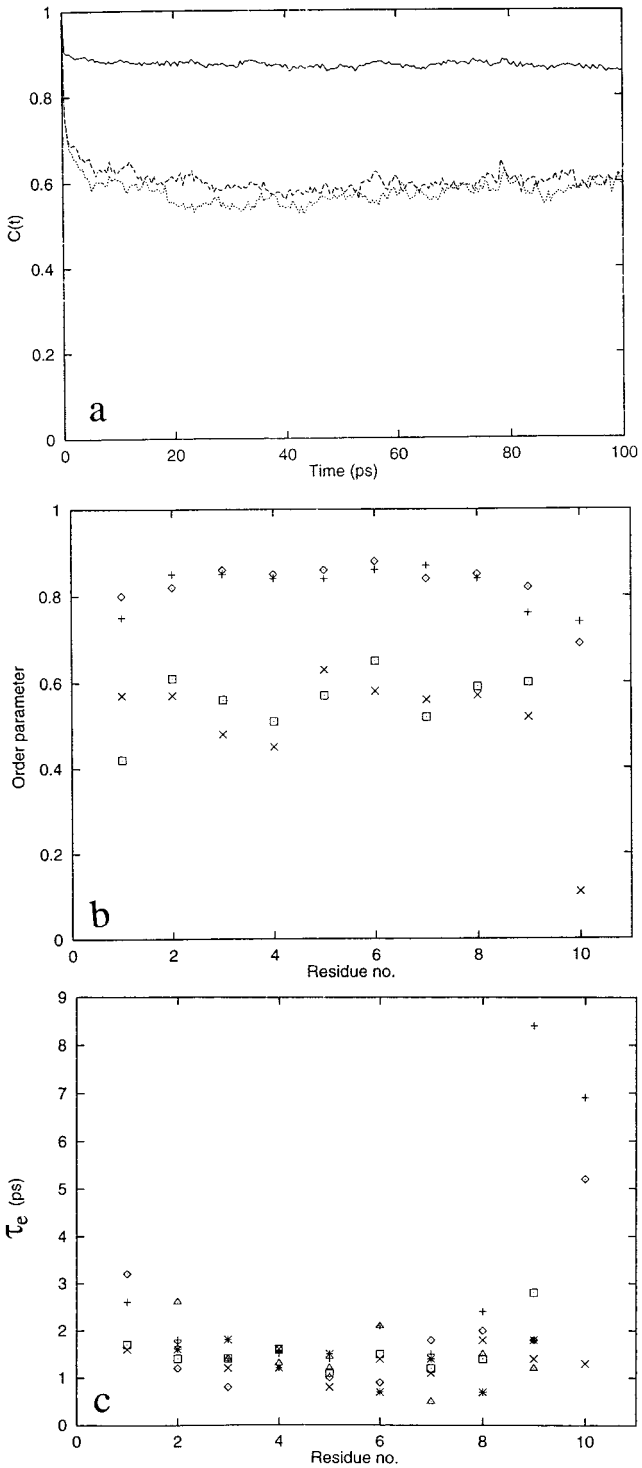

Figure 6. (a) Typical correlation function plots for the reorientational dynamics of the C3′−O3′ (—), C1′−O1′ (— —), and N1′−H1′ (- - -), bond vectors of the PNA strands in the PNA-PNA duplex system. (b) Comparison of the generalized order parameter (S^2) values against the residue numbers, for the reorientation dynamics of the bond vectors for the C3′−O3′ (◇, +), and for the C1′−O1′ (□, ×). The first symbol in each case corresponds to the first strand and the other for the second strand. (c) Comparison of the effective correlation time (τ_e) values against the residue numbers, for the reorientation dynamics of the bond vectors for the C3′−O3′ (◇, +), C1′−O1′ (□, ×), and N1′−H1′ (△, *).

plateau region and hence there are some statistical uncertainties associated with the estimation of S^2 and τ_e. For some bond vectors (particularly for the end residues) reasonably good plateau regions were not found in the corresponding correlation function plots, and hence in such cases, estimation of S^2 and τ_e was not possible.

Energetics and Stabilities of Different Duplexes. Considering the fact that the different types of duplexes of the same base sequence differ considerably in their thermal stabilities,[2−3,10] it is of interest to make a comparative study of the different interactions involved. Table 3 gives a summary of some of the important energetics. In each case, the average and the rms fluctuation have been calculated over 126 frames from the dynamic trajectory spanning the entire time range with stable average rmsd value. The self-energy of a molecular system represents the sum of all the interaction energies (both bonded and nonbonded) present in the entire molecular system. Comparison of the self-energies of the antiparallel and parallel PNA-DNA duplexes clearly indicates that the antiparallel duplex is energetically more favorable than the parallel one. As the reference state, where the component single stranded PNA and single stranded DNA are separated in solution, is the same for both the antiparallel and parallel complexes, the difference in their self-energies is relevant for their stability difference. This result is also consistent with the experimental data which indicate that the antiparallel duplex is indeed more stable than the parallel one.[2,3,10] On the other hand, although the self-energy of a PNA-PNA duplex is found to be more favorable than that of the corresponding DNA-DNA duplex, it cannot be directly correlated with their stability difference because their reference states, where the constituent strands are separated in solution, are not the same. It is also interesting to see from Table 3 that interaction of PNA with water is much less favorable than the interaction of DNA with water. This is consistent with the fact that the PNA backbone with several CH_2 groups is more hydrophobic than a DNA backbone with its charged phosphate groups.

It may further be pointed out that despite the difference in the chemical compositions of PNA and DNA strands and also the difference in their net electric charge, both the van der Waals and the electrostatic energies separately are similar for the interactions between the two complementary strands and also in the self-energies of the different duplexes, a result which indicates that the different types of duplexes adopted structures to minimize the unfavorable interactions with the charges in the DNA backbones, while maintaining similar base-base interactions.

Overall Structures. Figure 7 shows the average structures of the different duplex systems. Clearly in all the cases very regular structures with well-defined base pairing and base stacking are observed. The PNA-DNA antiparallel duplex and the PNA-PNA duplex systems appear to be very similar. Both are inbetween B- and A-helical structures, although in all the cases the starting structures were B-helix like. The PNA-DNA parallel duplex also shows well-defined structure and the base pair planes are more B-like. In the PNA strands in all the duplexes involving PNA, the C1′−O1′ bonds are exposed to the solvent and the C3′−O3′ bond is always oriented toward the C-terminal of the respective strand and is partially exposed to the solvent through the major groove. The top views indicate A-helix-like central cavities in the cases of antiparallel and PNA-PNA duplexes, indicating the existence of some A-like characters in their overall structure while no such cavity in the parallel complex indicates its B-like feature. In Table 4 we have given the rmsd values of the average structures in each case in reference to the canonical B- and A-helix structures for comparison. In the cases of both the antiparallel duplex and the PNA-PNA duplex, the 10 residues did not make a complete

Table 3. Comparison of the Energetics of the Different Types of Duplexes[a]

	($S^1 + S^2$) with water			S^1 with S^2			self-energy[b] of ($S^1 + S^2$)			counterion
	total	vdw	elec	total	vdw	elec	total	vdw	elec	total
anti-PNA-DNA	−1430	−202	−1227	−201	−47	−154	−763	−71	−1423	−58
	±65	±13	±65	±7	±4	±8	±21	±11	±13	±19
para-PNA-DNA	−1510	−199	−1311	−200	−44	−157	−709	−69	−1385	−96
	±44	±12	±44	±12	±5	±13	±21	±11	±14	±28
PNA-PNA	−618	−232	−387	−198	−50	−149	−841	−61	−1429	−
	±25	±14	±23	±7	±6	±4	±18	±10	±11	
DNA-DNA	−2277	−144	−2133	−200	−46	−154	−716	−88	−1445	−260
	±70	±15	±56	±6	±5	±7	±23	±12	±14	±23

[a] The averages and rms fluctuations are calculated over 126 different frames from the dynamic trajectories. The energies are expressed in kcal/mol. S^1 and S^2 represent the first and the second strands of a duplex, respectively. [b] The total self-energy also contains the contributions from the bonds, bond angles, and dihedral energy terms.

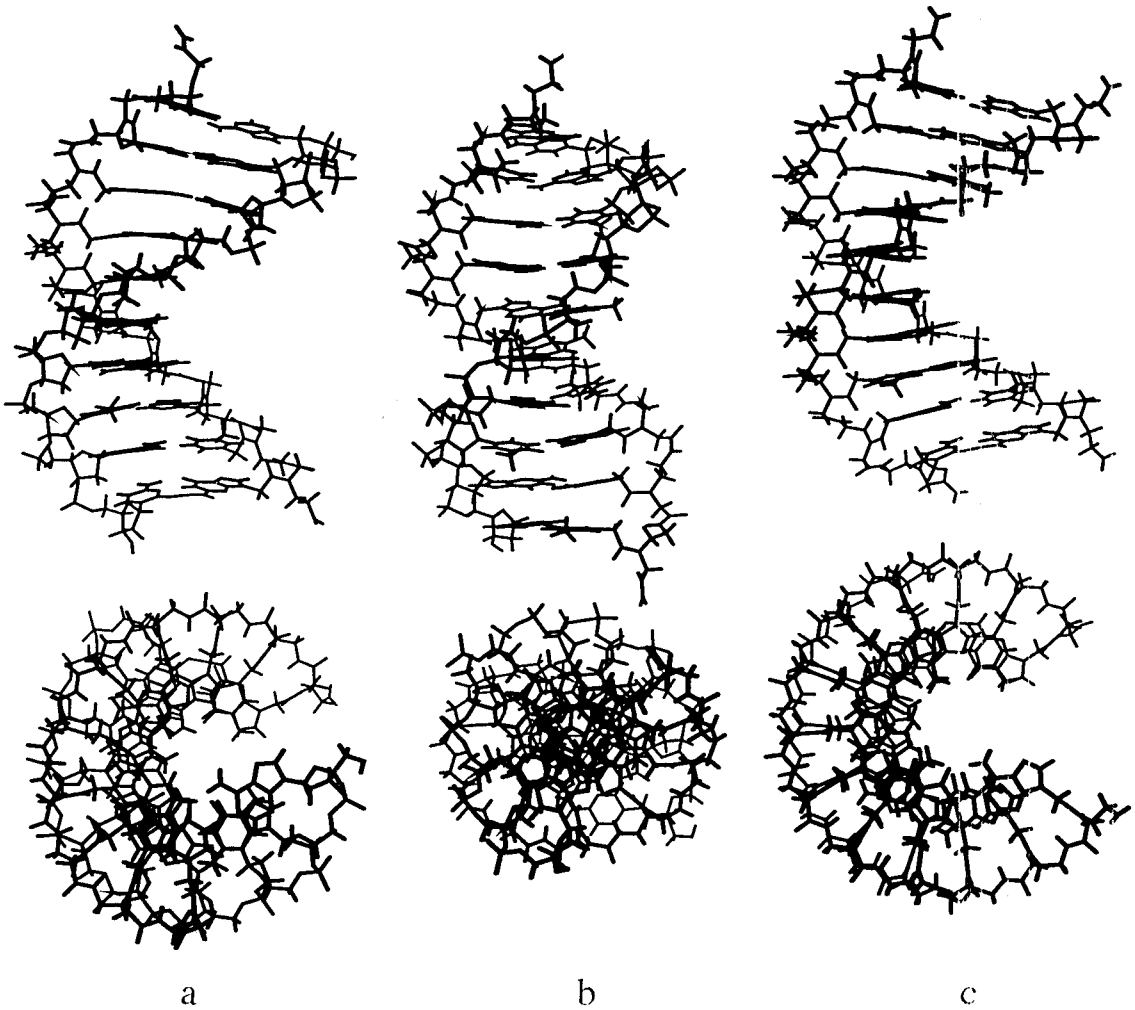

Figure 7. The average structures of the different duplex systems: (a) PNA-DNA antiparallel duplex, (b) PNA-DNA parallel duplex, and (c) PNA-PNA duplex. The top view of each structure is given at the bottom of each structure.

Table 4. Comparison of the rmsd (Å) of the Different Duplex Systems with Reference to the Canonical B-Form and Canonical A-Form

	antiparallel (PNA-DNA)	parallel (PNA-DNA)	(PNA-PNA)	(DNA-DNA)	canonical B-form
canonical B-form	4.2	2.1	4.9	3.1	
canonical A-form	3.9	5.4	4.3	3.3	5.8

turn as evidenced from the top views. This appears to be due to the presence of the rigid peptide part in the PNA backbone, which effectively acts as a linear stretch in the PNA backbone as appeared in Figure 8. This straight part reduces the effective bending of the strand and as a result the duplex becomes wider and underwound.

Inter- or Intraresidue H-Bonding. From early molecular mechanics studies on energy minimization of a PNA-DNA structure it was suggested that the helical structure of a PNA strand could be stabilized by an intra- or an interresidue H-bonding between the atom O3′ and the peptide N1′−H1′ group.[15a−c] To probe these possibilities, we have looked into the time evolutions of the acceptor−donor distances and angles for the intra- or interresidue hydrogen bonds. In both cases, the combination of the values of the average distance (4.6 Å for intra- and 3.8 Å for interresidue H-bonding) and the average angle (126° for intra- and 122° for interresidue H-bonding) does not support the existence of any strong H-bond between them.

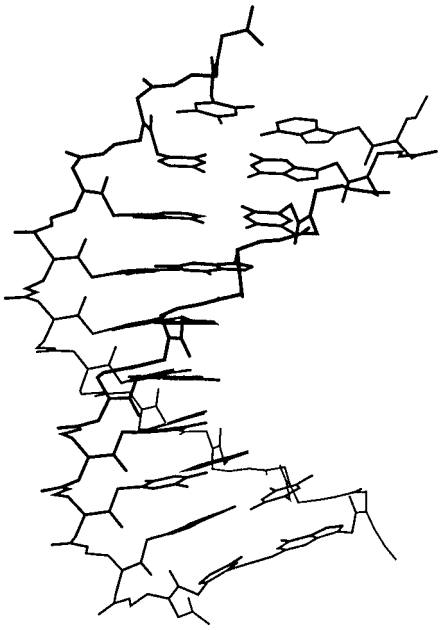

Figure 8. The PNA-PNA duplex structure (without H atoms) showing the rigid and effectively linear nature of the peptide bonds.

The typical time evolutions of the distance and angle for the interresidue H-bonding between O3′ and N1′−H1′ are represented in Figures A2a and A2b (Appendix, see Supporting Information), respectively, and Figures A2c and A2d (Appendix, see Supporting Information) represent the same for the intraresidue H-bonding between the *i*th O3′ and the $(i + 1)$th N1′−H1′. Calculation of the electrostatic energy of the interaction between these atoms also does not indicate any significant interaction between them.

Comparison with Experimental Data. In Table 5 we have compared the average torsional angles of the DNA and PNA backbones for the PNA-DNA antiparallel duplex system obtained by NMR studies[11] and the same for the PNA-PNA duplex as obtained from X-ray crystallographic studies[13] to the corresponding values obtained in our modeling and MD simulation studies. Comparison indicates excellent quantitative agreements between the MD derived values and the X-ray crystallographic data for the backbone torsional angle values of the PNA strands in the PNA-PNA duplex. The overall agreement of the backbone torsional angles for both the DNA and PNA strands in a PNA-DNA antiparallel duplex, with the NMR data, is reasonably good particularly for the DNA strand. The differences in some values for the PNA strand may be due to the difference in the refinement process of NMR derived data. In the NMR study refinement was done in vacuum while we have performed all our simulations in aqueous solution.

As the base sequence of our PNA-DNA antiparallel duplex is different from that of the NMR study, we could not compare the NOE distances. However, we have calculated the distances between a base hydrogen atom and a hydrogen atom on the PNA backbone where in a similar situation there are NOE data in the NMR study. We have found that in our MD simulation also the dynamic averages of such distances are within a distance of <5 Å, which is generally the limit for observing a NOE peak.

The rise value, which indicates the average separation of the successive two base pairs along the axis of a double helix, was found to be 3.2 Å in the crystallographic study[13] while in MD it is 3.0 Å. Moreover, the average twist value in the MD calculation is 23.1°, which corresponds to 15.5 base pairs per turn of the PNA-PNA double helix while crystallographic data give 18 base pairs per turn.[13] In the MD derived structure, the orientation of the C1′−O1′ bond was found to point to the solvent and the C3′−O3′ bond was found to point toward the C-terminal of the respective PNA strand in both the PNA-DNA duplex and the PNA-PNA duplex in complete agreement with NMR and crystallographic findings. The minor groove was narrower in both the PNA-DNA antiparallel duplex and the PNA-PNA duplex than in a B-helix as indicated in the NMR and crystal structure. In MD study we did not find any strong inter- or intraresidue H-bonds in agreement with the feature obtained in NMR and crystallographic results. Overall, the difference in structure between the parallel duplex and the antiparallel duplex is consistent with the CD results[3] which showed different CD spectra for parallel and antiparallel duplexes.

Distribution and Dynamics of Counterions. One of the most fundamental differences between a PNA-DNA and a DNA-DNA duplex system is that in the PNA-DNA duplex only one strand has a net negative charge and the other strand is neutral while in the DNA-DNA duplex both strands have net negative charges. We have analyzed and compared the movements and distribution of the counterions around the similar DNA strand in a PNA-DNA and a DNA-DNA duplex system. Although each of the Na^+ ions was initially placed at a distance of 3.5 Å from the phosphate atom of the respective phosphate group, during dynamics they were moved, and most of the time most of them kept an average distance of $\sim$5 Å from the closest phosphate group. Figure A3 (Supporting Information) compares the probability distribution of finding a counterion at a distance from the phosphorus atom of any phosphate group. There are clear differences between these distributions. In the case of the DNA-DNA duplex the structure in the distribution is more pronounced compared to the PNA-DNA case; probably the nearby phosphate groups of the complementary strand in a DNA-DNA duplex interact favorably with the counteions to keep them more ordered and close to the DNA.

Interactions with Solvent. As the chemical composition of the PNA backbone is quite different from that of a DNA, it is also interesting to have an idea about the interaction of PNA with solvent. We have looked into the radial distribution function around different donor and acceptor atoms on the PNA backbone. The results are plotted in Figure A4 (Appendix, see Supporting Information). The shoulder at the position at $r = 3$ Å indicates the possible existence of some preferred ordering around the C1′ atom of the PNA strand, but the effect is not very pronounced. The plot of $g(r)$ around N1′ atoms of PNA does not show any convincing evidence of local solvent ordering.

The H-bond analysis of the trajectories shows that H-bonds formed between water and donor or acceptor atoms of the backbone (N1′, H1′, O1′) as well as the bases (N3, N7, O2, O4, O6, H21, H22, H61, H6) of the PNA strand. It is found that most of the time during dynamic simulation, the donors/acceptors of the bases (both in the major and the minor grooves) are engaged in H-bonding with water molecules but the particular water molecule to which it is H-bonded is renewed frequently. The average residence time of such H-bonds varies in the range of a few picoseconds to 10 ps. In particular cases, a residence time as high as 30 ps has been found. On the other hand, the acceptor/donor atoms in the PNA backbone are engaged in H-bonding to water molecules with a much reduced frequency and with residence times of the order of a few picoseconds or less. The O3′ atom in the base linker region is

Table 5. Comparison of the Values of the Backbone Torsional Angles Obtained from Our Modeling and MD Data with Those Derived from NMR Data for the Solution Structure of the PNA-DNA Antiparallel Duplex and the X-ray Data on the PNA-PNA Duplex

system	method	α	β	γ	δ	ϵ	ζ	κ_1	κ_2	κ_3
DNA[b]	NMR	-71 ± 10	-165 ± 12	49 ± 5	110 ± 6	-171 ± 8	-75 ± 5			-134 ± 5
	MD	−72	161	56	110	−152	−95			−133
PNA[b]	NMR	105 ± 55	141 ± 13	78 ± 16	139 ± 13	35 ± 42		-3 ± 3	151 ± 9	-103 ± 19
	MD	170	65	78	78	61		−9	−172	99
PNA[c]	X-ray	−117	63	74	83	−15[a]		6	−173	89
	MD	173	65	75	77	68		−8	−172	98

[a] In the X-ray crystallography study this torsional angle was defined in a different way (N2′−C2′−C1′−O1), so this value has been converted according to our definition (N2′−C2′−C1′−+N1′), where +N1′ is in the next residue. [b] Strands in the antiparallel duplex. [c] PNA in the PNA-PNA duplex.

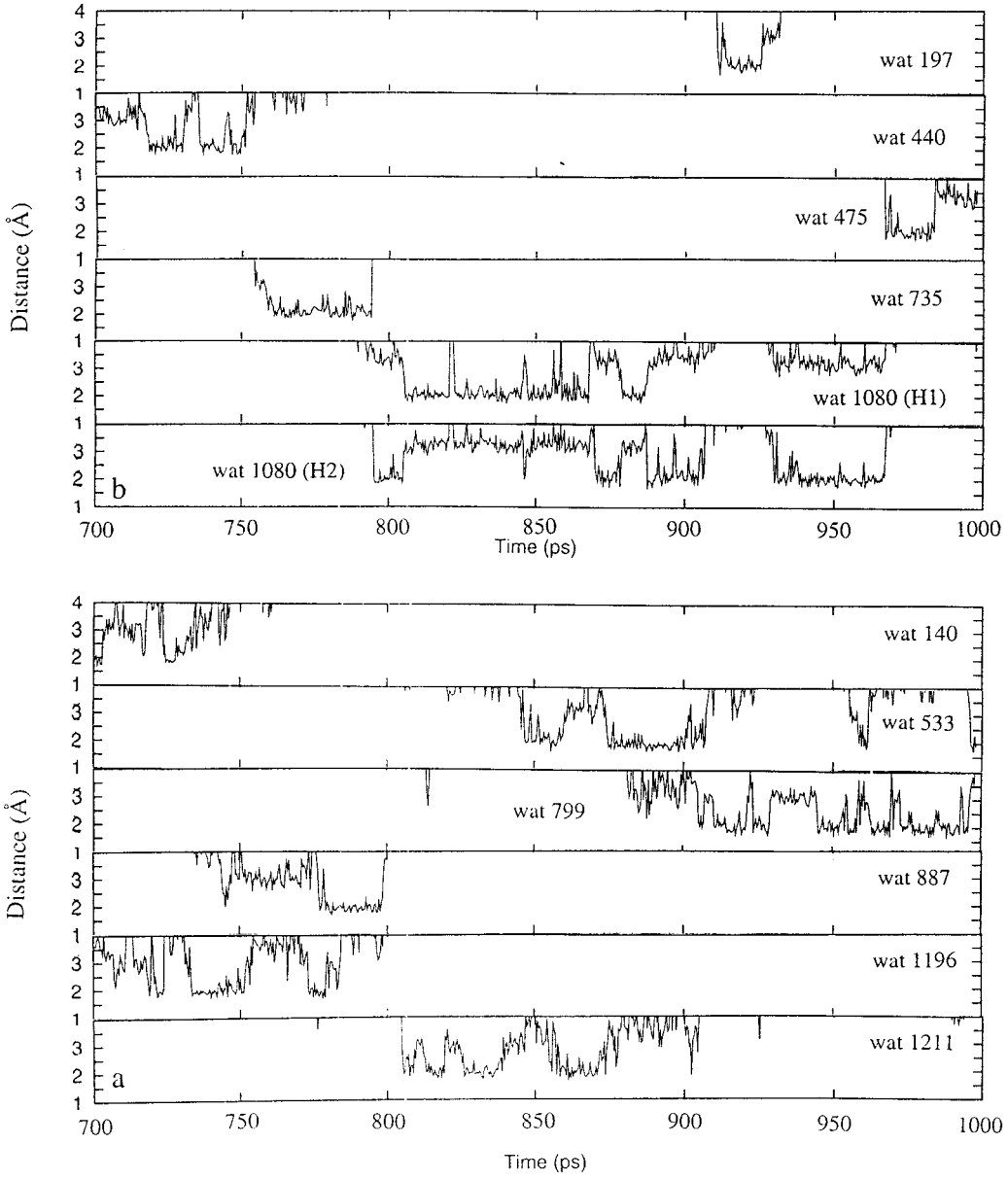

Figure 9. Time evolutions of the distance between the N7 atom (a) of the fourth residue and the N3 atom (b) of the second residue of the first strand of the PNA-PNA duplex and a hydrogen of some of the nearby water molecules. The number of that water molecule is also given in the figure.

also found to frequently make H-bonds with water molecules and the average residence time is a few picoseconds. Such short average residence times are consistent with experimental data, which indicates that, in general, the residence time of such H-bonding water molecules on the molecular surface is only a few picoseconds.[30] However, the phosphate oxygens of the DNA strand in a PNA-DNA antiparallel duplex are found to form H-bonds with water more frequently and the average residence time is longer. Figure 9a,b shows some typical examples of the time evolutions of the distances between the acceptor and the hydrogen atoms in such H-bonding between bases and water.

We have also looked for water molecule between the O3′ and N1′−H1′ groups bridging them to see if such a water bridge

(30) Billeter M.; Guntert, P.; Luginbuhl, P.; Wuthrich, K. *Cell* **1996**, *85*, 1057−1065.

can account for the regularity of their dynamics. We found a few water molecules staying close to both of them, but the average residence time was very short, less than a picosecond, and it occurred very infrequently.

Discussion and Conclusions

The present article reports the possible structures and internal dynamics as obtained from unrestrained MD simulation studies on decamer duplexes containing PNA-PNA, PNA-DNA antiparallel, and PNA-DNA parallel strands in aqueous solution and also compares the structural features of the decamer PNA-DNA antiparallel duplex and the PNA-PNA duplex with those of a NMR derived structure of an octamer PNA-DNA antiparallel duplex and the crystallographic data of a hexamer PNA-PNA duplex, respectively. The comparison indicates good agreement of our MD results on the PNA-DNA antiparallel duplex with the NMR derived results and excellent quantitative agreement between MD and crystallographic data on the PNA-PNA duplex system. The differences on the quantitative level with NMR data may be due to the fact that in the NMR work the structure was refined by MD simulations in vacuo while we have done our MD simulations in aqueous solution. Considering the facts that the starting model structures were B-like in all cases and no direct experimental constraints were used, the agreement with experimental data is very satisfactory and also illustrates the improved utility of the present days molecular modeling and molecular dynamics approaches.

All the duplexes involving PNA strand(s) maintained stable and well-defined structures over the entire period of simulation. Even though in all cases the initial structures were B-helix-like, the final average structures varied with the nature and the orientation of the strands. The antiparallel PNA-DNA duplex has a structure between that of A- and B-helixes while the parallel complex has a more B-like conformation. The PNA-PNA duplex is similar to the antiparallel PNA-DNA duplex and is between that of A- and B-helixes.

The fact that the differences in the structural and dynamical properties of the different duplexes in aqueous solution are small, as judged by the rmsd from A- and B-helixes (Table 4) which is smaller than the rmsd between canonical A- and B-DNA of the same size, can be rationalized as follows. The presence of the same nucleobases constrained by the homomorphous backbones tends to keep similar types of base pairing and base pair stacking geometries. On the other hand, the chemically and electrostatically different backbones introduce some perturbations to the structures, and as a result, structurally slightly different duplexes appear. For example, a RNA molecule has a slightly different backbone compared to a DNA of the same base sequence, due to the presence of the 2′OH groups which induce structural differences in the resulting RNA double helix. The degree of perturbation in different duplex systems depends on the diversity of the backbone atoms and hence, quite expectedly, the maximum structural difference in the present work was found in the PNA-PNA duplex structure, where both the inter- and intrastrand backbone interactions are most different from the DNA-DNA case. The observed differences in the structures between the antiparallel and parallel complexes between the same PNA and DNA strands may arise due to the difference in the arrangements of the covalent bonds along the PNA strand with respect to the DNA strand which also cause differences in the immediate atomic neighborhood.

In all the duplexes involving PNA, the fluctuations of PNA backbone torsional angles are mostly anticorrelated. The main motions of the AT base pairs are in-phase linear correlated motions while the GC base pairs exhibit in-plane bending motions, independent of the nature of the backbone.

Despite the absence of any ring structure in the base linker region of the PNA strand, the local internal dynamics still remains highly restricted. The physical reason behind the highly restricted motion of the base linker region in the PNA strand, even in the absence of any ring structure, could be some weak interaction between the C3′−O3′ and N1′−H1′ groups and steric hindrance by the surrounding atoms in that locality.

Energetically the difference in self-energies between the antiparallel and parallel PNA-DNA duplexes of the same base sequences is consistent with the experimentally observed stability difference between these two duplexes. The PNA molecule also appears to be more hydrophobic compared to the DNA molecules.

Even in the absence of any chiral center the PNA-PNA duplex maintained a stable and regular, well-defined double helical structure over the entire simulation (1.15 ns). This indicates that only base pairing and the base pair stacking along with a proper covalent distance between successive bases are quite sufficient to stabilize a double helical structure. A sugar−phosphate backbone is not a prerequisite for such a double helical structure, but it should be noted that within these double helical frameworks, these different complexes have different conformations.

No significant intra- or interresidue H-bonding between O3′ and N1′−H1′ was found in the PNA strand in any duplex.

The excellent agreement of the structural data from the present molecular modeling and MD simulation study with available NMR and crystallographic data indicates the reliability of the results obtained by MD simulations in the present work, and also for the parallel PNA-DNA duplex for which an experimentally determined structure is not yet available.

Acknowledgment. This work was supported by the Swedish Natural Science Research Council. S.S. would like to thank Prof. Astrid Gräslund for drawing his attention to PNA.

Supporting Information Available: An appendix containing three tables (Tables A1−A3) describing additional force field parameters for the PNA backbone and base pair H-bond statistics and four figures (Figures A1−A4) describing sugar puckering, solvent distribution, and time evolution of intra- and interresidue H-bonding in PNA backbone (8 pages). See any current masthead page for ordering and Internet access instructions.

JA972234X