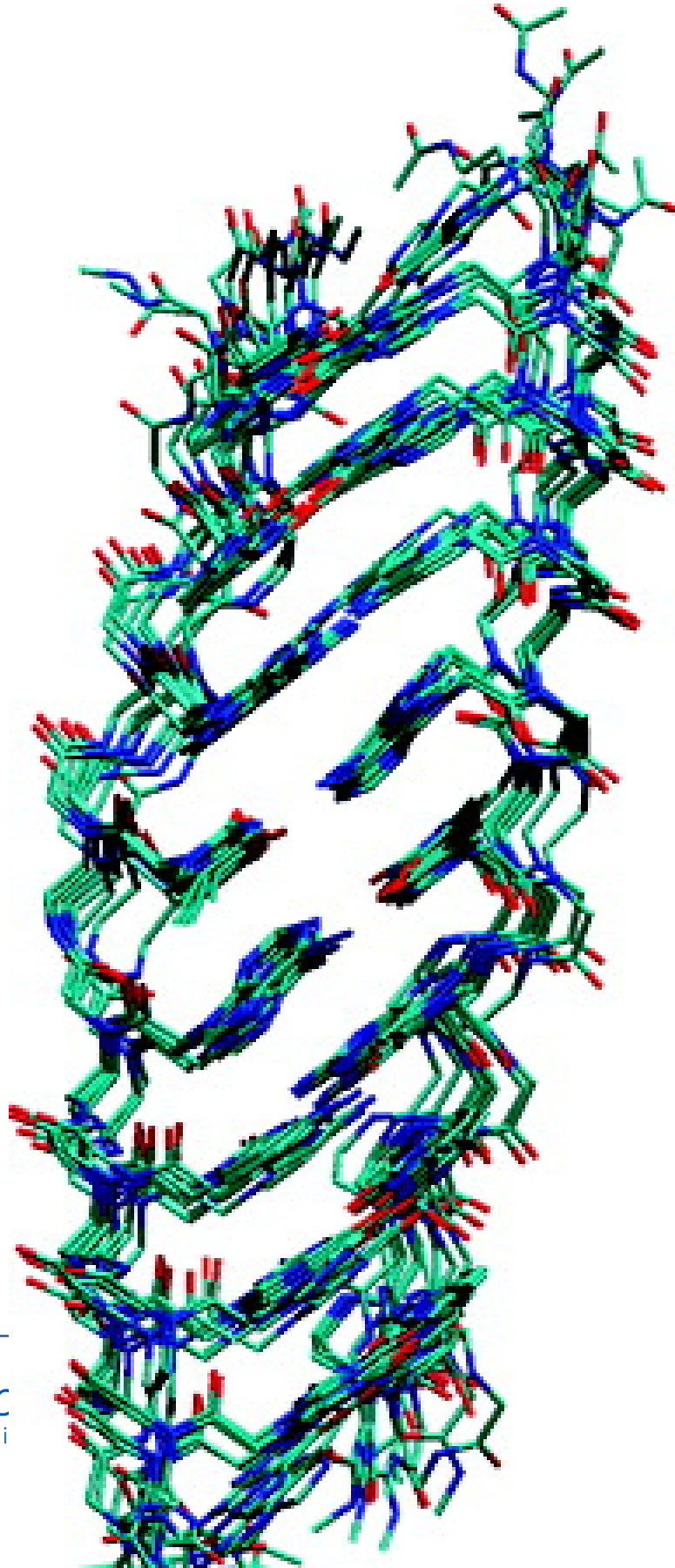
Article

Solution Structure of a Peptide Nucleic Acid Duplex from NMR Data: Features and Limitations

Wei He, Elizabeth Hatcher, Alexander Balaeff, David N. Beratan, Roberto R. Gil, Marcela Madrid, and Catalina Achim

J. Am. Chem. Soc., **2008**, 130 (40), 13264-13273 • DOI: 10.1021/ja800652h • Publication Date (Web): 10 September 2008

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY



ublished by the American Chemical
on, DC 20036

More About This Article

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

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 09/10/2008

Solution Structure of a Peptide Nucleic Acid Duplex from NMR Data: Features and Limitations

Wei He,[†] Elizabeth Hatcher,[‡] Alexander Balaeff,[‡] David N. Beratan,[‡] Roberto R. Gil,*[,†] Marcela Madrid,*[,§] and Catalina Achim*[,†]

Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, Department of Chemistry, Duke University, Durham, North Carolina 27708, and Pittsburgh Supercomputing Center, Pittsburgh, Pennsylvania 15213

Received January 26, 2008; E-mail: rgil@andrew.cmu.edu; mmadrid@psc.edu; achim@cmu.edu

Abstract: This paper describes the results of a 1D and 2D NMR spectroscopy study of a palindromic 8-base pair PNA duplex GGCATGCC in H_2O and H_2O-D_2O solutions. The ^1H NMR peaks have been assigned for most of the protons of the six central base pairs, as well as for several amide protons of the backbone. The resulting 36 interbase and base-backbone distance restraints were used together with Watson−Crick restraints to generate the PNA duplex structure in the course of 10 independent simulated annealing runs followed by restrained molecular dynamics (MD) simulations in explicit water. The resulting PNA structures correspond to a P-type helix with helical parameters close to those observed in the crystal structures of PNA. Based on the current limited number of restraints obtained from NMR spectra, alternative structures obtained by MD from starting PNA models based on DNA cannot be ruled out and are also discussed.

Introduction

Peptide nucleic acid (PNA) is a synthetic analogue of DNA that has a pseudopeptide backbone instead of a sugar-diphosphate backbone (Figure 1). The first PNA was reported in 1991 and had a backbone based on aminoethyl-glycine.[1] PNA contains the same nucleobases as DNA, and it forms homoduplexes based on Watson−Crick base pairing (Figure 1). PNA does not contain chiral centers, and consequently the homoduplexes can have a right or left handedness, which can be influenced by the inclusion of chiral amino acid residues in the PNA oligomers. This molecule holds promise for the development of gene therapeutic agents based on antisense or antigene strategies, because it binds with high sequence selectivity to DNA and RNA to form PNA·DNA and PNA·RNA heteroduplexes that are more stable than DNA·DNA and RNA·RNA homoduplexes.[2]

The structure of a double stranded (ds) PNA·PNA duplex has been determined by X-ray crystallography to 2.7 Å resolution.[3] That structure was determined for a palindromic, six base pair duplex based on a PNA oligomer with the sequence $H-C_1G_2T_3A_4C_5G_6-NH_2$. The helical structure of the PNA was distinct from that of DNA·DNA duplexes and was named P-form. The P-form PNA·PNA helix has a pitch of 18 base pairs per helical turn compared to 10−12 base pairs for DNA, and a diameter of 28 Å, compared to ∼24 Å for A- and B-DNA and 20 Å for Z-DNA duplexes. The PNA duplex has a wide and deep major groove and a shallow and narrow minor groove, which leads to base pairs being displaced toward the minor groove. The close-to-zero base pair inclination observed for a PNA duplex whose structure was determined by X-ray crystallography[3] was much smaller than that in A- and Z-DNA duplexes and comparable to that in B-DNA where the bases are perpendicular to the helical axis.[4] Interestingly, both right-handed and left-handed duplexes have been identified in the crystals of the PNA, both in the case in which the PNA oligomers did not contain a chiral center and in the case in which it contained an L-lysine at the C-terminus.[3,5] The terminal L-lysines were involved in cation−π interactions with the adjacent guanines and in hydrogen bonding with the linker carbonyl group and N7 of the adjacent guanine residues in PNA. Both interactions were mediated by water molecules. Nevertheless the influence of these L-lysines on the structure of the PNA duplex was not strong enough to lead to crystals containing exclusively left handed PNA duplexes.[5] A similar P-type structure was determined by X-ray crystallography to 2.2 Å resolution for a backbone-modified, N-methylated PNA·PNA duplex.[6] The N-methylation had an effect only on the conformation of the backbone amide bond, which was *cis*,[6] in contrast to the *trans* conformation observed in nonmethylated PNA duplexes.[3,5] The structure of a six base pair PNA duplex based on a PNA oligomer that contained a modified base, namely

[†] Carnegie Mellon University.
[‡] Duke University.
[§] Pittsburgh Supercomputing Center.

(1) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497–500.
(2) Nielsen, P. E. *Lett. Pept. Sci.* **2004**, *10*, 135–147.
(3) Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat. Struct. Biol.* **1997**, *4*, 98–101.
(4) Blackburn, G. M., Gait, M. J., Eds. *Nucleic Acids in Chemistry and Biology*, 2nd ed.; Oxford University Press: Oxford, New York, 1996.
(5) Rasmussen, H.; Liljefors, T.; Petersson, B.; Nielsen, P. E.; Kastrup, J. S. *J. Biomol. Struct. Dyn.* **2004**, *21*, 495–502.
(6) Haaima, G.; Rasmussen, H.; Schmidt, G.; Jensen, D. K.; Kastrup, J. S.; Wittung Stafshede, P.; Norden, B.; Buchardt, O.; Nielsen, P. E. *New J. Chem.* **1999**, *23*, 833–840.

 10.1021/ja800652h CCC: $40.75

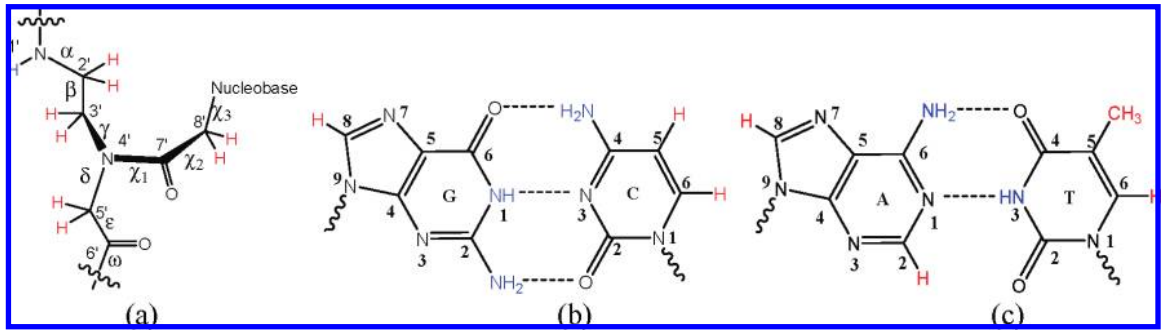


Figure 1. Chemical structure of backbone (a), GC (b), and AT (c) Watson−Crick base pairs of PNA. Exchangeable protons (imino, amino protons in the bases and amide proton in the backbone) and nonexchangeable protons (aromatic and T-methyl protons in the bases and methylene protons in the backbone) are labeled in blue and red, respectively. Torsion angles are marked in (a), and their definitions are α = C6′−N1′−C2′−C3′; β = N1′−C2′−C3′−N4′; γ = C2′−C3′−N4′−C5′; δ = C3′−N4′−C5′−C6′; ϵ = N4′−C5′−C6′−N1′; ω = C5′−C6′−N1′−C2′; χ_1 = C3′−N4′−C7′−C8′; χ_2 = N4′−C7′−C8′−N1(py)/N9(pu); χ_3 = C7′−C8′−N1(py)/N9(pu)−C2(py)/C4(pu).

H-G***bT***ATAC-L-Lys-NH_2 (***bT***: 1,8-naphtyridin-2(1*H*)-one), has also been determined by X-ray crystallography to 1.8 Å resolution.[7] That duplex adopted a P-type structure with a base pair displacement from the helical axis smaller and a base pair rise slightly larger than those observed for the nonmodified PNA duplex. More recently, the X-ray structure at 2.60 Å resolution of a partly self-complementary PNA decamer with the sequence H-$G_1T_2A_3\underline{G_4A_5T_6C_7}A_8C_9T_{10}$-L-Lys-$NH_2$ has been reported.[8] The central, self-complementary four base pair part of the decamer formed a P-type duplex by Watson−Crick base pairing. The structure differed from that of the six base pair duplex reported earlier only in the base pair inclination.[3]

Several structures of hybrid nucleic acid duplexes or triplexes that contained at least one single stranded (ss) PNA have been obtained by X-ray crystallography or NMR spectroscopy.[9−13] Several heteroduplexes adopted a structure that bears resemblance to that of homoduplexes of the nucleic acid to which PNA is hybridized, which led to the suggestion that PNA adapts to its nucleic acid partner.[9−11,13−16] The PNA strands in those PNA·DNA, PNA·RNA, and $(PNA)_2$·DNA structures share several properties with the PNA strands of PNA·PNA duplexes. The backbone amide adopts a *trans* conformation similar to that observed in PNA·PNA duplexes with a nonmodified backbone. With few exceptions, the C═O groups of the linkers that connect the nucleobases to the PNA backbone are oriented toward the C terminus of the PNA strand.[3,9,11−13] In several structures, the NH group of the backbone amides was involved in hydrogen bonding to the purine N3 or pyrimidine O2.[3,6,7,13] The majority of the C═O groups of the backbone amide was oriented toward the solution in nonmodified PNA[3] and inward in the N-methylated PNA duplexes.[6] The methylene groups showed broad peaks in the NMR spectra[9,11] and, thus, have been difficult to assign, but the X-ray and NMR structures showed that the ethylene diamine adopts a staggered conformation with a NCCN torsion angle β of approximately $+75(\pm15)°$ and $-75(\pm15)°$ in the right- and left-handed structures, respectively.[12,13]

Molecular dynamics (MD) simulations have been a valuable tool for studying the structure and dynamics of PNA. MD simulations have been performed on ss PNA,[17] ds PNA,[18] hybrid PNA·DNA and PNA·RNA duplexes,[18,19] and PNA_2·DNA triplexes.[20,21] All the duplexes involving PNA strands maintained stable and well-defined structures over the duration of the simulation.[18,19] MD simulations of a ds PNA led to a structure in which the torsion angles, the width of the minor groove, and the helical twist were similar to those determined by X-ray crystallography.[18] The width of the minor groove and solvent-accessible surface of homoduplexes are smaller for PNA than for DNA, which is consistent with the nature of the neutral, hydrophobic PNA backbone.[18] The small twist of the helix, i.e. ~23.3° for the MD model of ds PNA, is attributed to the presence of rigid peptide bonds in the PNA backbone, which effectively act as linear stretches and unwind the helix.[18] MD simulations led to the suggestion that the significantly larger x-displacement of the PNA·DNA duplexes when compared to that of DNA·DNA duplexes causes fluctuations in hydrogen bonding, water rearrangement, and interruptions in base stacking around AC mismatches and consequently leads to higher mismatch discrimination in PNA·DNA and possibly in PNA·PNA duplexes.[22] An important role for base stacking in the stabilization of PNA duplexes was supported by the fact that the root-mean-square deviations (RMSDs) of the bases were consistently smaller than those of the PNA backbone.[18] Furthermore, MD studies of ss nucleic acids concluded that ss PNA maintains the initial base-stacked helical conformation throughout MD simulations, while ss DNA or ss RNA shows large fluctuations, thus suggesting that the coupling between the

(7) Eldrup, A. B.; Nielsen, B. B.; Haaima, G.; Rasmussen, H.; Kastrup, J. S.; Christensen, C.; Nielsen, P. E. *Eur. J. Org. Chem.* **2001**, *178*, 1–1790.
(8) Petersson, B.; Nielsen, B. B.; Rasmussen, H.; Larsen, I. K.; Gajhede, M.; Nielsen, P. E.; Kastrup, J. S. *J. Am. Chem. Soc.* **2005**, *127*, 1424–1430.
(9) Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994**, *265*, 777–780.
(10) Eriksson, M. *Nucleosides Nucleotides* **1997**, *16*, 617–621.
(11) Eriksson, M.; Nielsen, P. E. *Nat. Struct. Biol.* **1996**, *3*, 410–413.
(12) Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995**, *270*, 1838–41.
(13) Menchise, V.; De Simone, G.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, R.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12021–6.
(14) Kastrup, J. S.; Pilgaard, M.; Jorgensen, F. S.; Nielsen, P. E.; Rasmussen, H. *FEBS Lett.* **1995**, *363*, 115–17.
(15) Leijon, M.; Graeslund, A.; Nielsen, P. E.; Buchardt, O.; Norden, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry* **1994**, *33*, 9820–5.
(16) Menchise, V.; De Simone, G.; Corradini, R.; Sforza, S.; Sorrentino, N.; Romanelli, A.; Saviano, M.; Pedone, C. *Acta Crystallogr., Sect. D* **2002**, *58*, 553–555.
(17) Sen, S.; Nilsson, L. *J. Am. Chem. Soc.* **2001**, *123*, 7414–7422.
(18) Sen, S.; Nilsson, L. *J. Am. Chem. Soc.* **1998**, *120*, 619–631.
(19) Soliva, R.; Sherer, E.; Luque, F. J.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **2000**, *122*, 5997–6008.
(20) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1997**, *119*, 7463–7469.
(21) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 5895–5904.
(22) Rathinavelan, T.; Yathindra, N. *FEBS J.* **2005**, *272*, 4055–4070.

Scheme 1. Sequence of the Palindromic Eight-Base Pair PNA Duplex

$$H\text{-}G_1\ G_2\ C_3\ A_4\ T_5\ G_6\ C_7\ C_8\text{-}Lys\text{-}NH_2$$

$$H_2N\text{-}Lys\text{-}C_8\ C_7\ G_6\ T_5\ A_4\ C_3\ G_2\ G_1\text{-}H$$

nucleobases and the backbone is weaker in ss PNA than in ss DNA or ss RNA.[17] Nevertheless, in ds PNA the dynamics of the linker that connects the base to the PNA backbone is highly restricted, possibly due to steric interactions with the surrounding atoms or to weak interactions between C7′−O7′ and N1′−H1′.[18] The orientation of the C═O group of the linker is toward the C′ terminus of the PNA strand, as observed in the crystal structures.[3,9,11,13]

All experimental PNA·PNA structures reported to date have been obtained by X-ray crystallography, and no solution structure is available for a PNA·PNA duplex. Application of NMR spectroscopy techniques to PNA is challenging because of the chemical characteristics of the PNA backbone. Proton assignment methods usually applied to DNA are not applicable to PNA because of the lack of sugar rings. Moreover, all methylene protons in the PNA backbone have similar chemical shifts. This paper reports the first NMR information about a PNA·PNA duplex. We have conducted 1D and 2D ^{1}H correlated NMR spectroscopy (COSY) and nuclear Overhauser enhanced NMR spectroscopy (NOESY) studies of an eight-base pair PNA duplex of the self-complementary sequence shown in Scheme 1. The duplex is palindromic, and consequently, the number of ^{1}H NMR signals is one-half that of a nonpalindromic duplex with the same number of base pairs. To ensure that the duplex is stable at room temperature despite its relatively short length, we situated GC base pairs at the ends of the duplex. The structures were then calculated by means of simulated annealing followed by MD simulations in explicit water, with distance restraints derived from the NOESY spectra and Watson−Crick restraints. The simulations were repeated for several different starting structures. The properties of PNA duplexes in the resulting structural ensemble are discussed.

Results

Assignment of NMR Resonances and Determination of NOE Restraints. The imino (G-H1 and T-H3) and amino (G-H2N, C-H4N, and A-H6N) protons of the nucleobases and the amide (H1′) protons of the backbone are exchangeable (blue in Figure 1), while the aromatic protons (G-H8, C-H5/6, A-H2/H8, and T-H6) and methyl protons (T-H5) of the nucleobases and the methylene protons (H2′/H2″, H3′/H3″, H5′/H5″, H8′/H8″) of the backbone are nonexchangeable (red in Figure 1). Consequently, 1D and 2D NMR spectra of PNA solutions in D_2O solution show peaks for the nonexchangeable protons only (Figure S1), while the corresponding spectra of PNA dissolved in H_2O/D_2O (9:1) contain peaks for both the exchangeable and nonexchangeable protons (Figure 2). Nonexchangeable aromatic protons have been assigned by first examining the NMR spectra of PNA in D_2O. Protons connected by chemical bonds experience scalar coupling, which is revealed by COSY cross peaks. Protons situated in close proximity of each other, whether chemically bonded or not, experience through-space dipolar coupling and produce cross peaks in the NOESY spectra. In ^{1}H NMR studies of DNA, the sequential assignment of the protons is based on NOE interactions between the sugar and the nucleobase protons.[23,24] In contrast to DNA, PNA contains six methylene protons in each backbone unit and these protons have similar chemical shifts because they are relatively far from the bases. Therefore, the peaks of the methylene protons significantly overlap in the narrow 2.5−4 ppm chemical shift range and have a broad line shape. This property prevented us from assigning the methylene protons' peaks and from using these peaks to assign the nucleobase protons.

Strong COSY cross-correlation peaks have been observed between the C-H5 and C-H6 protons of the C3, C7, and C8 cytosines at δ (6.67, 5.41), (6.86, 5.43), and (7.17, 5.71) (Figure 3a). The existence of a NOE cross peak at δ (7.17, 6.67) indicates that the protons to which these peaks pertain are in adjacent bases and thus they must be the cytosine protons C7-H6 and C8-H6 (Figure 3b). Consequently, the peak at δ (6.86, 5.43) observed in the COSY spectrum was assigned to NOE between C3-H5 and H6, and the cross peak at δ (6.86, 7.87) has been assigned to a NOE between C3-H6 and A4-H8. The cross peak δ (6.67, 7.41) must be due to a NOE between either C7-H6 or C8-H6 and a proton in an adjacent base other than a cytosine and, therefore, was assigned to a NOE between C7-H6 and G6-H8 protons. Based on this analysis, the peaks at 6.67 and 7.17 ppm were assigned to the C7-H6 and C8-H6 protons, respectively, and the peaks at 5.43, 5.41, and 5.71 ppm, to the H5 protons of the C3, C7, and C8 cytosines, respectively (Figure 4).

The T-CH_3 and T-H6 protons and the C-H5 and C-H6 protons exhibit cross peaks in the COSY spectra. The 8-mer PNA contains only one thymine, T5, whose T5-CH_3 and T5-H6 protons have shown a COSY cross peak at δ (1.35, 6.33) (Figure 5a). In the D_2O NOESY spectrum, the T5-CH_3 protons (1.35 ppm) show also strong cross peaks with the aromatic protons of the bases that make the A4·T5′ and G6·C3′ base pairs adjacent to the T5·A4′ one, i.e. A4-H2/H8 and G6-H8 (see peaks at δ (1.35, 7.34), (1.35, 7.87), and (1.35, 7.41), respectively, in Figure 5b). As the G6-H8 and A4-H8 protons have been shown to have peaks at 7.41 and 7.87 ppm, respectively, the 7.34 ppm peak is assigned to the A4-H2 proton. The only two peaks yet to be assigned in the aromatic region of the NMR spectrum of the PNA in a D_2O solution (Figure 4) are at 7.2−7.3 ppm and must belong to the G1-H8 and G2-H8 protons.

Four imino protons, G1-H1, G2-H1, T5-H3, and G6-H1, are present in the palindromic PNA duplex. Only three sharp signals were observed in the imino region of the ^{1}H NMR spectrum of the duplex in H_2O/D_2O (9:1) (Figures 2 and 6a), at 13.43, 12.99, and 12.37 ppm. The fourth peak is not observed because of the fast exchange between the terminal G1-H1 and water due to fraying of the PNA duplex. The peak at 13.43 ppm was assigned to T5-H3 because the imino protons of thymines have the lowest field signals in NMR spectra of nucleic acid duplexes.[11,23,24] The cross peak at δ (13.43, 12.37) in the NOESY spectrum must be an imino−imino NOE between the T5-H3 and the adjacent G6-H1 imino proton. A weak NOE peak observed at δ (12.99, 12.37) has been assigned to an interstrand interaction between G2-H1 and G6-H1 (Figure 6b). The small splitting observed for the C8-H5/H6 and G2-H1 signals in the 1D NMR spectrum can be due to either duplex asymmetry or the existence of the (terminal base pairs of the) duplex in two slightly different conformations (Figures 4 and 6a).

(23) Hare, D. R.; Wemmer, D. E.; Chou, S. H.; Drobny, G.; Reid, B. R. *J. Mol. Biol.* **1983**, *171*, 319–36.
(24) Scheek, R. M.; Boelens, R.; Russo, N.; Van Boom, J. H.; Kaptein, R. *Biochemistry* **1984**, *23*, 1371–6.

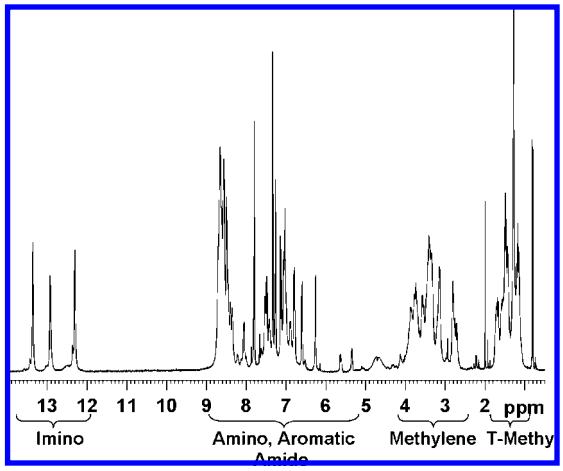


Figure 2. ^{1}H NMR spectrum of the PNA duplex in H_2O/D_2O (9:1), pH 7.0 10 mM sodium phosphate buffer ([ds PNA] = 0.4 mM, 27 °C).

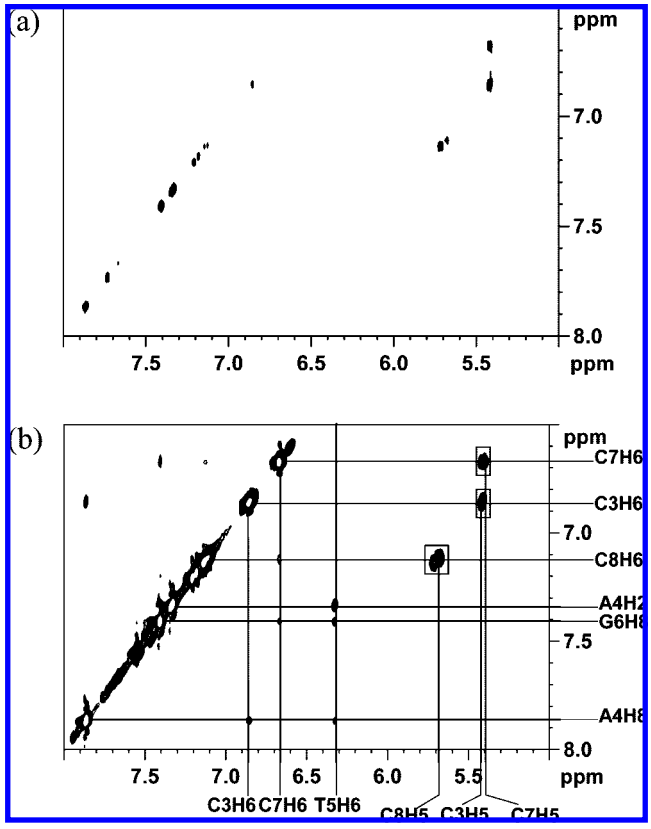


Figure 3. Aromatic region of COSY (a) and NOESY (b) spectra of the PNA duplex in D_2O ([ds PNA] = 0.4 mM, 27 °C). NOESY peaks corresponding to the H5−H6 correlation in each of the three cytosines are circled.

The imino protons of each guanine show NOEs with the amino protons of their complementary cytosines (G2-H1 to C7-H4N at δ (12.99, 8.78) and (12.99, 7.08) and G6-H1 to C3-H4N at δ (12.37, 8.76) and (12.37, 6.96)) (Figure 6c). Using these NOEs and the known peaks for the imino protons of G2-H1 and G6-H1 (see above), the amino protons of C7 and C3 have been assigned. The assignments of cytosine amino protons above were corroborated by the observation of strong and broad NOE peaks at δ (8.78, 7.08) and (8.76, 6.96) (Figure 7). Similarly, the imino proton of T5 showed an NOE with the

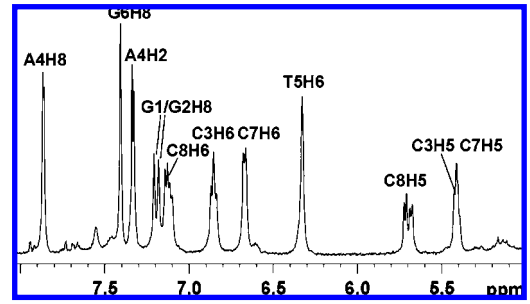


Figure 4. Aromatic proton region of the ^{1}H NMR spectrum of the PNA duplex in D_2O ([ds PNA] = 0.4 mM, 27 °C).

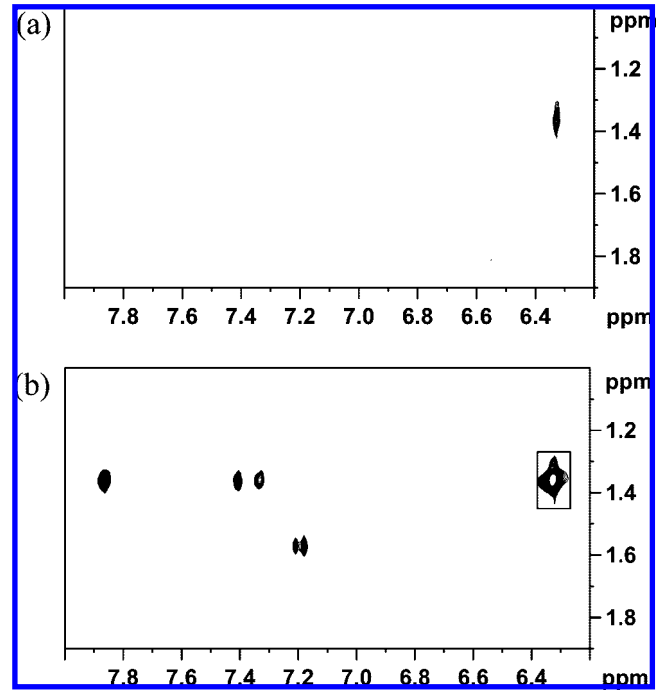


Figure 5. Methyl-aromatic region of COSY (a) and NOESY (b) spectra of the 8-bp PNA duplex in D_2O, respectively ([ds PNA] = 0.4 mM, 27 °C). The NOESY cross peak (circled) corresponds to the T5 CH_3−H6 correlation.

amino proton of the complementary adenine A4-H2 at δ (13.43, 7.33) (Figure 6c).

The amide protons in the backbone (H1′) have chemical shifts that are comparable to those of the amino protons (8.5−8.8 ppm). Therefore, cross peaks corresponding to aromatic-amide and aromatic-amino NOEs appear in the same region of the H_2O/D_2O NOESY spectrum (Figure 7). Strong and narrow NOE cross peaks in this region are observed and assigned as the amide-aromatic interaction between the amide proton and the base. Each amide proton gives two NOE peaks, one to the H6/H8 proton of its own nucleobase and the other to the H6/H8 proton of the nucleobase in the C′-flanking PNA residue.

The analysis of the 1D ^{1}H NMR, COSY, and NOESY spectra of the eight base pair PNA duplex in H_2O/D_2O and D_2O led to assignments of most of the nucleobase protons and of several protons of the backbone (Tables 1 and 2, and Figure 8). An estimate of distances between protons that are in close proximity of each other in the PNA duplex has been made using the dependence of the intensity of the NOE peaks on $(r_{ij})^{-6}$, where r_{ij} is the distance between protons i and j. The linear range of NOE buildup with respect to mixing time was determined

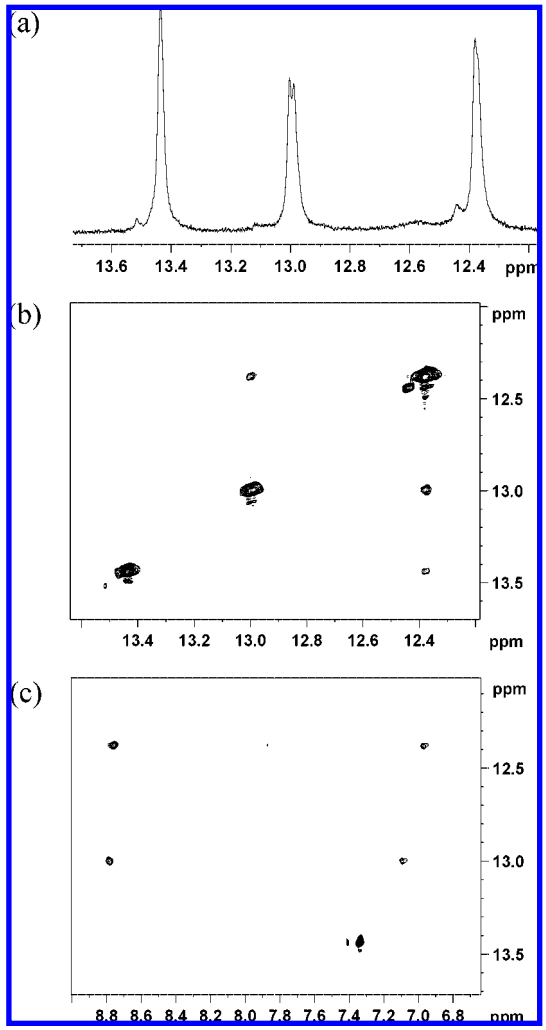


Figure 6. Imino region of the 1D ^{1}H NMR spectrum (a), and imino−imino (b) and imino−amino (c) regions of the NOESY spectrum of the PNA 8-mer in H_2O/D_2O (9:1), pH 7.0 10 mM sodium phosphate buffer ([ds PNA] = 0.4 mM, 27 °C).

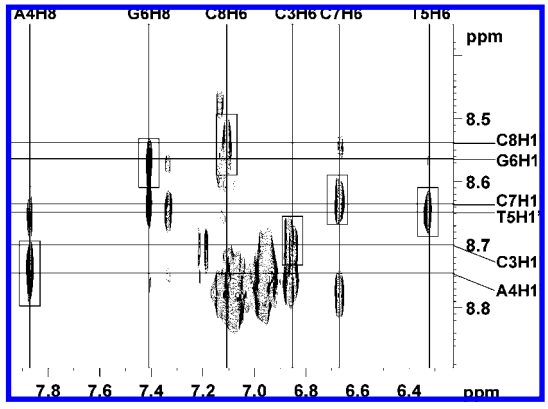


Figure 7. Amide-aromatic region of NOESY spectrum of the 8-bp PNA duplex in H_2O/D_2O (9:1), pH 7.0 10 mM sodium phosphate buffer ([ds PNA] = 0.4 mM, 27 °C). Peaks marked in boxes are the NOE peaks between the backbone amide proton (H1′) and aromatic (H6/H8) proton in the base.

(Figure S2), and the proton distances have been calibrated using a distance of 3.0 Å between the T-CH_3 and the T-H6 protons. This analysis led to 36 interbase and base-backbone distances and 8 intrabase H5−H6 and NH2−H5 distances in cytosine.

Table 1. Proton Chemical Shifts of the Central Bases of the 8-mer PNA

residue	G2	C3	A4	T5	G6	C7
H2	-	-	7.34	-	-	-
H5	-	5.43	-	-	-	5.41
H6/H8	n.a.	6.86	7.87	6.33	7.41	6.67
imino	12.99	-	-	13.43	12.37	-
amino	-	6.96/8.76	-	-	-	7.08/8.78
CH_3	-	-	-	1.35	-	-
amide	n.a.	8.70	8.74	8.65	8.56	8.64

The 36 interbase and base-backbone distances together with 36 Watson−Crick hydrogen bond distance restraints were used in subsequent molecular dynamics simulations. We have also restrained the torsion angle β to values of $-60° \pm 10°$, which are characteristic for a left-handed PNA duplex, because the CD spectrum of the eight-base pair PNA duplex is indicative of a left-handed structure. The ideal values for β in left-handed PNA duplexes is $-60°$.[25] Furthermore, β had values between $+60°$ and $+90°$ in all crystal structures of right-handed homo- or heteroduplexes that contained at least one PNA strand.[3,12,13]

Structures Simulated by Molecular Dynamics. To obtain the structure of the PNA duplex, the NOE restraints were employed in a series of MD simulations, conducted as described in the Experimental Section. The main round of MD simulations consisted of 10 independent simulated annealing runs followed by 5 ns restrained MD simulations in explicit water. These simulations were started from the structure of the eight base pair PNA duplex created from the crystal structure of ds PNA.[3] Two more runs employed starting structures corresponding to the canonical A- and B-DNA helices.[26] The structures obtained from each of these three starting points are called below X-PNA, A-PNA, and B-PNA, respectively. Time-average structures were calculated over the last 400 ps for each of the MD trajectories.

The time evolution of the RMSDs of the backbone and nucleobases calculated with respect to the starting structure for each MD trajectory showed that all X-PNA trajectories were

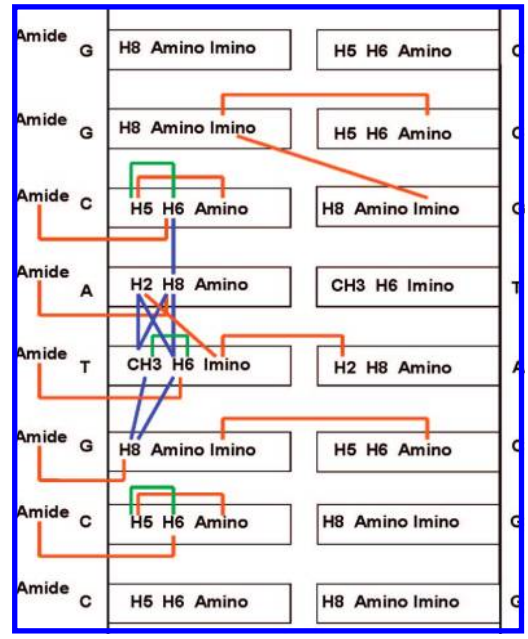
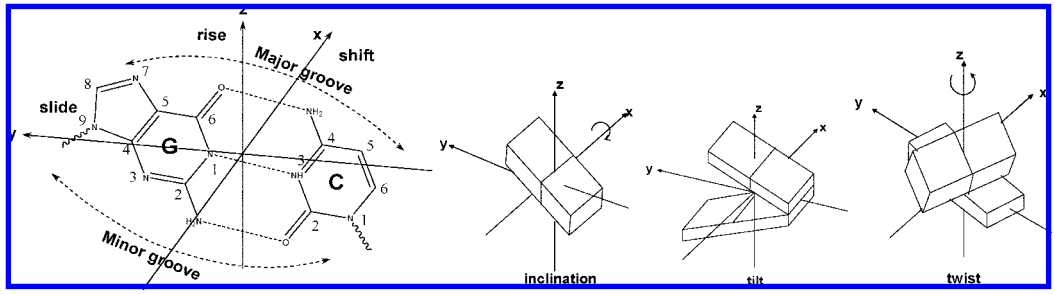


Figure 8. Schematic NOE correlation map throughout the PNA 8-mer duplex, showing COSY correlations (green), nonexchangeable NOEs (blue), and exchangeable NOEs (orange). For simplicity, only NOE interactions of one strand of the duplex are shown.

Table 2. Helical Parameters for PNA, DNA, and RNA[a]

PNA sequence	handedness	disp. (Å)	rise (Å)	incl. (deg)	tilt (deg)	twist (deg)	bp/turn		ref.
H-GGCATGCC-L-Lys-NH_2 (X-PNA)	L	7.9 (0.7)	3.7 (0.1)	−5.0 (2.8)	−0.2 (1.0)	−17.3 (0.6)	21	NMR	this paper[b]
H−CGTACG-NH_2	R	8.3	3.2	0.3	1.0	19.8	18	X-ray	3
H−CGTACG-L-Lys-NH_2	L	6.6	3.4	9.4	0	−19	18	X-ray	5
	R	7.7	3.2		0	20.5	18		
H−$C_{Me}GT_{Me}AC_{Me}$G-L-Lys-NH_2	L1/L2	7.2/7.2	3.5/3.6	−1.4	0/0.3	−20/−20	18/18	X-ray	6
	R1/R2	4.8/4.9	3.8/3.8		0.2/0.1	18.9/19	19/19		
H-GTAGATCACT-L-Lys-NH_2	L	7.9	3.3	−8.9	0	−18	20	X-ray	8
	R	3.5	3.4	−26.7	0	19.1	19		
H-G**bT**A TAC-L-Lys-NH_2	R	3.3	3.7	0.5	0.1	18	18	X-ray	7
	L	2.7	3.8		0.7	−18	18		
H-AGTGATCTAC-H H-TCACTAGATG-H	R		2.9	14.9		23.3	15	MD	18
A-DNA	R	−5.3	2.56	12	20	33	11		4
A-RNA	R	−5.3	2.8	15.8	0	32.7	11		4
B-DNA	R	0.0	3.4	2.4	0	36	10		4
Z-DNA	L	3.3	3.8	7.1	−1.7	−60/2	12		4



[a] Definitions of parameters are illustrated below the table. [b] Each parameter is reported as the mean value of the global averages (with standard deviation) calculated for each of the 10 X-PNA structures using Curves.[30,31]

equilibrated after approximately 1.5 ns and that the backbone deviated from the starting structure more than the nucleobases did (Figures 9a and S3, navy and magenta, Table S2). The RMSDs with respect to the time-average structures for each of the 10 MD trajectories showed also that the backbone atoms of the duplex deviated more than the nucleobases from the average structure, as expected due to the higher flexibility of the backbone (Figures 9a and S3, yellow and cyan, Table S3). The A-PNA and B-PNA simulations resulted in larger deviations from the corresponding initial structures than the X-PNA simulations (Figures 9b and S4a,b, Table S2).

The helical parameters of the X-PNA structures were calculated for the six central base pairs of the duplex of each of the 10 average X-PNA structures (Tables 2 and S4) and are similar to those previously reported for P-form PNA, which is relatively unwound and exhibits a larger base pair displacement from the helical axis than B-DNA (Tables 2 and S4). Superposition of the base pairs of the 10 time-averaged X-PNA structures showed that the structural similarity between these structures is higher in the base pair region than in the backbone region, and it is also higher in the central part of the duplex than at the end of the duplex (Figure 10a and Table S3). The A- and B-PNA simulations yielded helices in which the nucleobases are relatively close to those of X-PNA (Figure 10b), but whose backbone differs significantly from that of X-PNA. These two duplexes are also unwound, but the displacement is

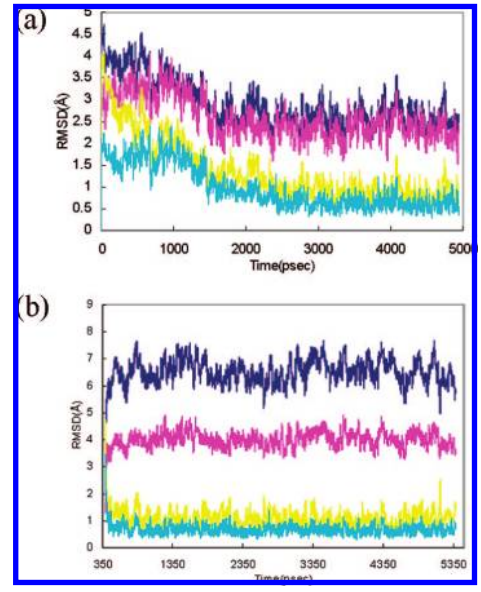


Figure 9. Time evolution of RMSDs of the backbone and nucleobase atoms with respect to the starting structure (backbone in navy and base pair in magenta) and to the time-averaged structure (backbone in yellow and base pair in cyan) for trajectory 9 of the X-PNA simulations (a) and for the A-PNA simulation (b).

significantly smaller than that of PNA and much closer to that of B- or Z-DNA (Tables 2 and S4). The difference in the calculated helical parameters of A-, B-, and X-PNA structures, obtained using the same set of NMR restraints, is likely due to the fact that the number of the restraints is limited. The restraints involving the backbone are especially scarce, including only those between the amide hydrogens and their nucleobases. We note that, in contrast to the global helical parameters, the local helical parameters, e.g., the displacement, shift, and slide, for A-, B-, and X-PNA are close

(25) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 10258–10267.
(26) Case, D. A. et al. *AMBER*, 9; University of California, San Francisco, 2006.
(27) Topham, C. M.; Smith, J. C. *J. Mol. Biol.* **1999**, *292*, 1017–1038.
(28) Topham, C. M.; Smith, J. C. *Biophys. J.* **2007**, *92*, 769–786.
(29) Srinivasan, A. R.; Olson, W. K. *J. Am. Chem. Soc.* **1998**, *120*, 492–508.
(30) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63–91.
(31) Lavery, R.; Sklenar, H. *Curves 5.1, Computer Program*; Institut de Biologie Physico-Chimique: CNRS, 1996.

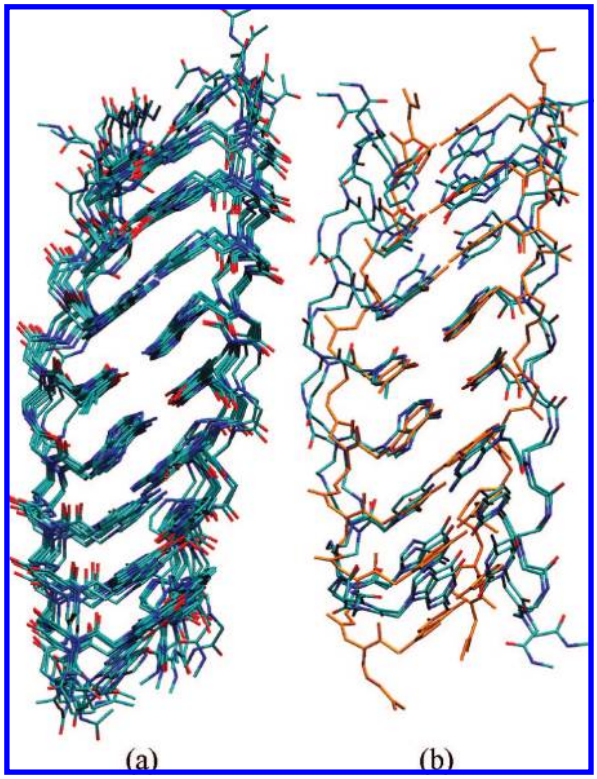


Figure 10. (a) Superposition of the 10 time-average X-PNA structures obtained by restrained MD simulations. (b) Superposition of A- and B-PNA structures with the average of the 10 time-average X-PNA structures (orange).

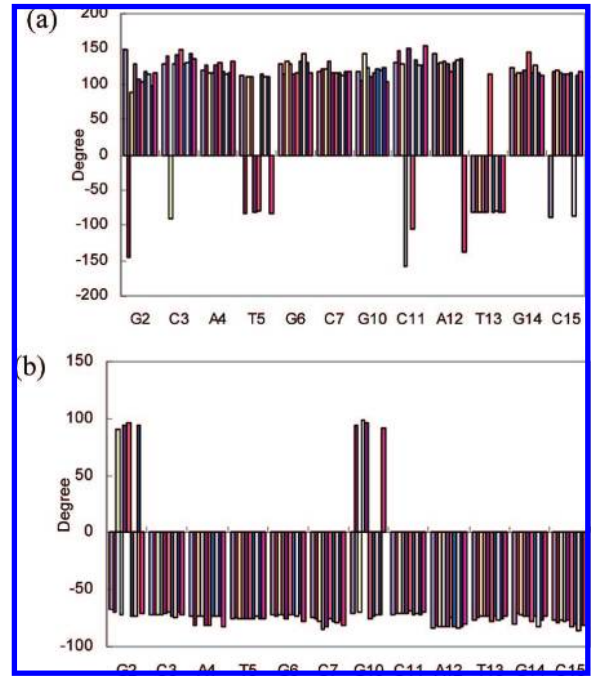


Figure 11. Torsion angles α (a) and γ (b) for the central nucleobases of the 10 time-average X-PNA structures.

to each other (Figure S11) and within the range of values for the corresponding parameters measured in the PNA duplexes whose structure was determined by X-ray crystallography (Figure S12).[5–7]

The torsion angles for the 10 average X-PNA structures fall within relatively narrow ranges (Figures 11 and S5), and most variations occur near the end of the two strands. In all 10 X-PNA structures and in the A- and B-PNA, the angles $\chi_1-\chi_3$ of all PNA monomers are each narrowly distributed, indicating the same structure of the linker that connects the nucleobase to the backbone. χ_2 is close to 180°, which means that the amide carbonyl group of the linker and the C−N bond that connects the nucleobase to the linker are coplanar and parallel to each other. The torsion angle γ adopted values of ∼ −75° in all X-PNAs and in A- and B-PNA with only a few exceptions at the terminals (Figures 11b, S5 and S7, Table 3). In all monomers of the 10 X-PNA structures, the torsion angles ω and χ_1, which describe the backbone and linker peptide bonds, respectively, have values of ∼175° and ∼2°, as expected due to the planarity of the peptide bond. The carbonyl group of the linker is consistently oriented toward the C′ terminus of the PNA strand because χ_1 is ∼0. In contrast, the angle ω in the A- and B-PNA has a value close to 0°, indicative of a *cis* form, which is unfavorable and consequently was observed previously only for prolines in peptides and never in PNA. Therefore, we consider that it is unlikely that such a structure would be adopted by PNA in solution given that a structure in which the peptide bond has an angle ω of ∼180° as in X-PNA is accessible.

In all the structures, the angles α and ϵ were correlated (Figures 13a, S5 and S8a). Two sets of pairs of values for (α_{i+1},ϵ_i) were found for the X-PNA structures: (−93°(24°),−155°(17°)) and (123°(16°), −3°(13°)). These angles were also correlated with the pseudo torsion angle ν, which is defined by the atoms C8′−N4′−C6′−O6′ (Figures 12, 13b, and S8b). Specifically, the pseudo torsion angle ν_i of each residue i and the torsion angle α_{i+1} of the subsequent residue always had opposite signs (Figure 13b). These correlations have been reported in the previous studies of PNA-containing duplexes.[22,27,28] The correlation between ν_i and α_{i+1} determines opposite orientations of the backbone carbonyl, namely toward the N′ terminus if $\nu_i > 0$ and $\alpha_{i+1} < 0$, and toward the C′ terminus if $\nu_i < 0$ and $\alpha_{i+1} > 0$ (Figure 12). The distance between the two oxygen atoms of the backbone and linker carbonyl groups within the same residue in the structures from the first category was 3.2−4.1 Å (Figure 12a), which suggests that a water molecule could hydrogen bond the two atoms, while in the structures from the second category the corresponding distance was 4.1−4.5 Å (Figure 12b). In the A- and B-PNA structures the (α_{i+1},ϵ_i) pair of angles adopts values of (−100°, 150°) (Figures S9 and S10).

Discussion and Conclusion

The structures of PNA calculated in this study based on both PNA and DNA initial structures are unwound when compared to A- and B-DNA, with A- and B-PNA being more unwound than X-PNA. The helical parameters depend on the starting structure, although all the structures satisfy the NMR restraints to the same extent (Table S1). In general, the 10 X-PNA structures that were obtained from a model based on the PNA X-ray structure are more similar to each other than to those that evolved from either A-DNA or B-DNA. Both A- and B-PNA ended up having a *cis* peptide bond ($\omega = 0$) in the backbone, which suggests that they are metastable states from which the 5 ns MD simulation could not escape. Simulated annealing could conceivably change those structures and likely bring them closer to the X-PNA.

The 7.9 Å displacement of the base pairs from the global axis of the double stranded X-PNA helix is comparable to that measured for ds PNA in the crystal structure, which is 8.3 Å. The 3.7 Å rise per base pair observed for the X-PNA is larger than that of PNA in the crystal structure (3.2 Å). It is rather

Table 3. Average Torsion Angle Values (with Standard Deviation) (deg) of PNA Strands Involved in Watson−Crick Base Pairing in Several Duplexes and Triplexes That Contain at Least One PNA Strand

helix type	method		α	β	γ	δ	ϵ	ζ/ω	χ1	χ2	χ3	ref.
PNA·PNA[a]	NMR	L	−93(24) 123(16)	−66(5)	−75(4)	−88(10)	−155(17) −3(12)	175(10)	2(4)	169(10)	−82(8)	this work
PNA·PNA	X-ray	R	−120 (6) 83(11)	66(5)	74(9)	91(11)	172(11) −19(3)	−168(12)	6(6)	−175(4)	83(3)	3
modified PNA·PNA	X-ray	R	−116(9) 69(7)	73(12)	72(6)	92(10)	150(0) −13(20)	−177(5)	3(2)	−173(4)	85(7)	7
		L	−79(32) 119(11)	−75(10)	−72(4)	−92(8)	−143(2) 3(11)	176(5)	−3(2)	177(6)	−87(6)	
modified PNA·PNA	X-ray	R	−106(8) 74(2)	79(11)	68(8)	91(7)	160(6) −13(12)	177(9)	0(3)	185(6)	79(5)	6
		L	−72(4) 119(19)	−81(20)	−71(8)	−84(7)	−169(5) −1(18)	181(6)	1(3)	171(3)	−77(5)	
PNA·PNA	MD	R	173(3)	65(3)	75(2)	77(2)	68(5)	180	−8(3)	188(3)	98(3)	18
PNA·DNA	X-ray	R	−114(15) 72(7)	73(9)	67(4)	90(5)	179(10) −21	−178(2)	4(6)	−178(3)	−91(6)	13
PNA·DNA	MD	R	−100 ± 10[b]	75 ± 15	75 ± 15	75 ± 15		195 ± 15	∼0	180−260	30−90	22
PNA·RNA	NMR	R	165(10)	67(7)	81(3)	77(6)	83(17)	180	6(3)	−158(14)	70(11)	9
PNA_2·DNA	MD	R	+120 ± 20	60 ± 40	60 ± 40	60 ± 40	60 ± 40	180 ± 40	60 ± 40	180 ± 40	60 ± 40	29
PNA_2·DNA	X-ray	R	−103	73	70	93	165	180	1	190	89	12

[a] Each torsion angle is reported as the average value of all 10 X-PNA structures. [b] The ± values define the whole range of values adopted by the respective parameter.

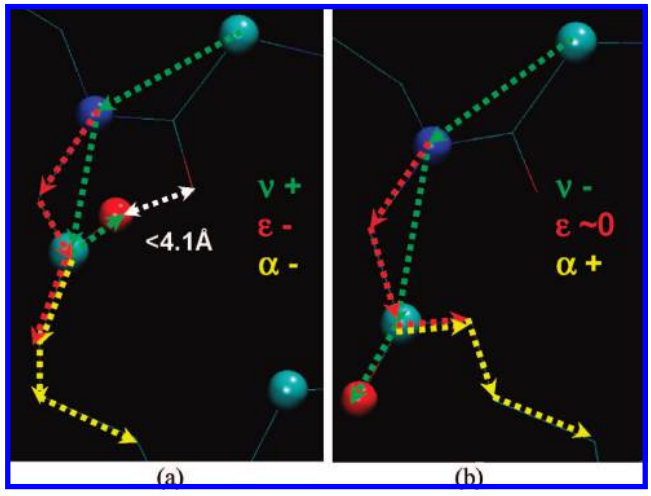


Figure 12. Angles ν, ϵ, and α in residues observed in the X-PNA structures (a and b). The examples shown in the figure are for the base pairs A_4·$T_{5'}$ (a), and T_5·$G_{6'}$ (b) of the X-PNA 1.

comparable to the rise observed in the crystal structures of the N-methyl backbone-modified PNA[6] and ***bT*** base-modified PNA[7] (Table 2). The rotational parameters of the PNA duplex in solution, i.e., inclination, tilt, and twist, are small and comparable to those reported for nonmodified and modified PNA duplexes in several crystal structures.[3,5−7] Those parameters are indicative of (1) the base pairs being perpendicular to the PNA helical axis, and similar to B- and Z-DNA duplexes, and (2) a significantly unwound PNA duplex, with a number of base pairs per helical turn (∼20 bp/turn) larger than that in either DNA or RNA (10−12 bp/turn). In contrast, a relatively large inclination was observed for the four base pairs involved in the duplex portion of the crystal structure of a single-strand PNA recently reported by Petersson et al.,[8] which suggests that PNA has a large structural flexibility and is capable of adopting a broad range of structural motifs when forming duplexes. The width of the minor groove of the 8-bp X-PNA duplex is 9.8(±0.9) Å (Figure S13, Table S5), which is close to that of the ds PNA in crystals (8.8(±0.3) Å) and is smaller than that of DNA duplexes (∼11 Å). This property is consistent with previous observations that, by virtue of being neutral, PNA has a higher hydrophobicity

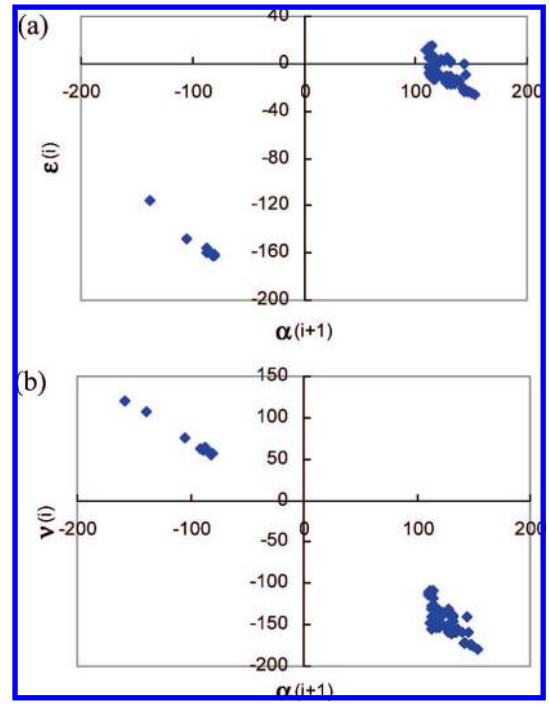


Figure 13. Correlation plot for α_{i+1} and ϵ_i (a) or α_{i+1} and ν_i (b) in the residues of the 10 time-average structures for X-PNA.

than DNA and forms helices with a narrower minor groove and a smaller solvent accessible surface than those of ds DNA.[18] The values of the torsion angles for the X-PNA duplexes are consistent with those for PNA in the reported crystallographic structures of the 6-bp PNA·PNA double helix,[3] a D-Lys PNA·DNA,[13] and a PNA_2DNA triplex,[12] and the structures obtained by MD simulations for a PNA·DNA duplex and a PNA_2·DNA triplex.[22,29]

Since the discovery of PNA in 1991, several X-ray diffraction studies have revealed that PNA shows a propensity to form P-type homoduplexes, which are more unwound and have a larger off-axis displacement than ds DNA (Table 2). During the same period of time, a solution structure determined by NMR spectroscopy and MD simulations has been reported for a

heteroduplex of PNA with RNA, but not for PNA homoduplexes, which is due to the difficulty posed by the assignment of the methylene protons of the backbone of PNA. In the current study, we were able to assign for the first time the protons of the PNA nucleobases and of the backbone amide group of a PNA homoduplex. Use of these distance restraints in MD simulations that started with a model based on the crystal structure of PNA and included a simulated annealing step led to a family of X-PNA structures that have helical parameters similar to those previously reported for PNA in crystals and torsion angles similar to those of PNA in homo- and heteroduplexes. Based on the relatively small number of restraints derived by NMR spectroscopy, we cannot rule out alternative structures obtained by MD simulations using DNA as the initial model for PNA, which are similar to X-PNA in the nucleobase part but differ from X-PNA in the backbone part of the duplex. A better determination of the structure must await the assignment of the protons of the PNA backbone. Isotope labeling or small chemical modifications of the PNA backbone are strategies which can help to accomplish that goal.

Experimental Section

PNA Monomer and Oligomer Synthesis. PNA monomers A, T, C, and G were purchased from Applied Biosystems and used without further purification. PNA nonmodified 8-mer H-GGCAT-GCC-Lys-NH_2 has been synthesized using the Boc-protection strategy.[32] After cleavage, PNA has been precipitated using ethyl ether and was purified by reversed-phase HPLC using a C18 silica column on a Waters 600 Controller and Pump. Absorbance has been measured at 260 nm with a Waters 2996 Photodiode Array Detector. Characterization of the oligomers has been performed by MALDI-TOF mass spectrometry on an Applied Biosystems Voyager Biospectrometry Workstation with Delayed Extraction and an R-cyano-4-hydroxycinnamic acid matrix (10 mg/mL in 1:1 water/acetonitrile, 0.1% TFA). MALDI-TOF calculated/found for (8-mer + H)$^+$ *m/z*: 2315.24/2314.90. PNA stock solution of the 8-mer has been prepared by dissolving PNA solid in sodium phosphate buffer (10 mM, pH 7.0). The stock solution concentration has been checked by UV−vis spectroscopy using a value for ε_{260nm} of 77 200 $M^{-1}\cdot cm^{-1}$ calculated using common values for the absorption coefficients for the nucleobases.[32]

^{1}H NMR Studies. The 1D ^{1}H NMR WATERGATE water suppression spectra of the PNA duplex have been collected both in H_2O/D_2O (9:1) and in D_2O. The PNA oligomer concentration was 800 μM, and 512 scans have been collected for each ^{1}H NMR spectrum. COSY and WATERGATE-NOESY spectra of PNA 8-mer have been collected both in H_2O/D_2O (9:1) and in D_2O. The PNA oligomer concentration was 800 μM. A total of 64 scans and 640 t_1 increments were collected in the NOESY WATERGATE water suppression experiment. A set of five NOESY experiments with different mixing times of 50, 100, 150, 200, and 300 ms, respectively, has been obtained and used to determine the linear range of NOE buildup. Repetition delay (d1) was set to 1.5 s. The 50 and 100 ms experiments fall within the linear range of the NOE buildup (Figure S2). All NMR spectra were collected at 300 K on a Bruker Avance DMX-500 operating at 500.13 MHz.

Molecular Modeling. All molecular modeling was performed using AMBER.[26] The force field ff99SB[33] was complemented with the electrostatic charges of the PNA atoms. These charges were computed separately for each of the four nucleic bases bound to a section of PNA backbone, following the Cornell et al. protocol.[34] The sections of PNA backbone were capped with an acetyl group ($-COCH_3$) on the N′-terminus and an N-methylamide group on the C′-terminus. The geometries were optimized using the HF/6-31G* basis set.[35] The ESP (electrostatic potential) charges on each atom were obtained by running a single-point calculation using the MP2/6-31G* basis set[35] and then refined by the restrained electrostatic potential (RESP) method implemented in the Antechamber module of AMBER.[26] Using these initial charges, RESP was run twice more: first constraining the charges of the acetyl and N-methylamide groups to the values defined in the Cornell et al. force field,[34] and then restraining the charges of chemically identical hydrogens. The obtained charges were self-consistent as well as consistent with the force field of Cornell et al.[34]

The simulations were conducted using three different starting structures of the PNA duplex. Ten main runs were based on the structure obtained by modifying the crystal structure of dsPNA[3] (PDB code 1PUP). Two more runs started from the structures derived from the canonical A and B DNA forms.[26]

The crystal structure of a PNA·PNA hexamer[3] was modified with VMD[36] as follows. First, an extra base pair was added to each end of the original hexamer, then the N- and C-backbone termini were capped with acetyl and *N*-methylamide groups, respectively, and finally, the nucleic bases of the original sequence were mutated to obtain the desired sequence (GGCATGCC). The charged lysine residues at the C′-termini were not included for simplicity of modeling, since no interaction was observed between the lysines and the main body of the duplex by NMR. Moreover, a previous crystallographic study of a single L-lysine-tethered PNA sequence showed the existence of both right- and left-handed helices alternatively despite the presence of the tethered L-lysine,[5] thus supporting the lack of a significant interaction between the lysine moiety and the PNA. The thus constructed structure was subjected to 10 independent 250 ps runs of restrained simulated annealing. Each of the annealing runs has been started with different initial random velocities in order to improve the coverage of the conformational space. A time step of 1 fs was employed. The temperature was increased to 600 K during the first 80 ps and then slowly decreased to 300 K during the rest of the simulation.

The simulated annealing runs employed 36 distance restraints obtained from the NMR experiments and 36 Watson−Crick base pairing restraints. The potential function for the restraints has a flat square bottom between the minimum and maximum interproton distances estimated from the NOESY cross peaks and parabolic sides for 0.5 Å beyond that and grows linearly thereafter. The force constant for the restraints was 20 kcal/mol. The weight of this force constant was gradually increased from 0.1 to 1 during the first 80 ps and then kept at 1 for the remainder of the simulation. In left- and right-handed PNA duplexes, the torsion angle β takes the ideal values of −60° and +60°, respectively.[25] Indeed, with the exception of one PNA·DNA duplex whose structure was determined by NMR, β had values between +60° and +90° in all crystal structures of right-handed homo- or heteroduplexes that contain at least one PNA strand. Therefore, during the MD simulations, β was restrained to values between −70° and −50°, because the CD spectrum of the PNA 8-mer duplex is indicative of a left-handed structure.

The 10 structures resulting from the simulated annealing runs were solvated in a truncated octahedral TIP3P water box, such that the distance between the walls of the box and the closest PNA atom was 9 Å. After energy minimization, the solvated structures were subjected to 5 ns of MD using the same restraints as above. The simulations employed an NPT ensemble (T = 300 K, P = 1

(32) Nielsen, P. E., Eds. *Peptide Nucleic Acids: Protocols and Applications*, 2nd ed.; Horizon Bioscience: Wymondham, U.K., 2004.
(33) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Structure, Function and Bioinformatics* **2006**, *65*, 712–725.
(34) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz Jr., K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
(35) Frisch, M. J. et al. *Gaussian 03*, rev. C.02, Gaussian, Inc.: Wallingford, CT, 2004.
(36) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14.1*, 33–38.

atm) in periodic boundary conditions. The Particle Mesh Ewald method was used to compute full electrostatic interactions.[26]

Average structures for each of the 10 MD simulations were generated from the last 400 ps of the trajectories. These structures and the experimental restraints have been deposited in the Protein Data Bank (www.wwpdb.org/),[37] code 2K4G. The convergence of the simulations was inferred using RMSDs from either the initial structure or the time-averaged structures. The RMSDs were computed after the alignment of either the backbone atoms N1′, C2′, C3′, N4′, and C5′ or the base pair atoms N1, N3, C5, N7, and N9. Helical parameters were calculated with the program Curves,[30] using only the six central base pairs of the PNA duplex, because the terminal G−C pairs were not regularly stacked with the central base pairs. A linear global helical axis was chosen for the duplex, and standard nucleotide geometry was applied in the calculation (using command settings MINI = true, LINE = true, FIT = true).[30,31]

Two more MD trajectories were based on initial structures derived from A- and B-DNA, generated with the module nucgen of AMBER.[26] The DNA backbone and sugar atoms were replaced by the corresponding PNA backbone atoms, according to the one-to-one mapping scheme between PNA and DNA.[18] The resulting structures were solvated in truncated octahedral TIP3 water boxes and minimized, as described above. MD was performed for 380 ps, during which the bases were harmonically restrained with a force constant which was slowly reduced from 100 to 0.2 kcal/mol Å^2. Finally, the harmonic restraints were released, and the solvated PNA was subjected to a 5-ns-long MD run using the same MD protocol and NOE, Watson−Crick, and torsion angle restraints as those above.

In order to test the effect of the experimental restraints on the simulated structure, unrestrained MD simulations were also performed. Approximately 5- and 4-ns-long unrestrained MD simulations were conducted for both A-PNA and B-PNA, respectively (Figure S4, c and d). The structures obtained from these two simulations exhibit a large number of violations of the restraints (Table S1), indicating that the restraints have been effective in determining structural parameters during the simulation.

Acknowledgment. M.M. and D.N.B. acknowledge support from the U.S. National Science Foundation (CHE 0628169). C.A. acknowledges support from the U.S National Science Foundation (CHE 0347140), the Sloan Foundation, and the Camille and Henry Dreyfus Foundation. We thank the NSF (CHE-0130903 and CHE-9808188) for partially supporting our NMR spectroscopy and MALDI-TOF mass spectrometry instrumentation. Most simulations were performed at the Pittsburgh Supercomputing Center (TG-MCB070070N).

Supporting Information Available: Complete refs 26 and 35, NMR spectrum of the PNA in D_2O, figures containing rmsd for MD trajectories, tables and figures containing global and local helical parameters and torsion angles for A-, B-, and X-PNA and for several PNAs for which crystal structures have been previously reported, structure statistics for MD simulations, and minor groove widths. This material is available free of charge via the Internet at http://pubs.acs.org.

JA800652H

(37) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.