DOI: 10.1002/chem.201000392

The Crystal Structure of Non-Modified and Bipyridine-Modified PNA Duplexes

Joanne I. Yeh,*[a] Ehmke Pohl,[b] Daphne Truan,[c] Wei He,[d] George M. Sheldrick,[e] Shoucheng Du,[a] and Catalina Achim*[d]

Abstract: Peptide nucleic acid (PNA) is a synthetic analogue of DNA that commonly has an *N*-aminoethyl glycine backbone. The crystal structures of two PNA duplexes, one containing eight standard nucleobase pairs $(GGCATGCC)_2$, and the other containing the same nucleobase pairs and a central pair of bipyridine ligands, have been solved with a resolution of 1.22 and 1.10 Å, respectively. The non-modified PNA duplex adopts a P-type helical structure similar to that of previously characterized PNAs. The atomic-level resolution of the structures allowed us to observe for the first time specific modes of interaction between the terminal lysines of the PNA and the backbone and the nucleobases situated in the vicinity of the lysines, which are considered an important factor in the induction of a preferred handedness in PNA duplexes. Our results support the notion that whereas PNA typically adopts a P-type helical structure, its flexibility is relatively high. For example, the base-pair rise in the bipyridine-containing PNA is the largest measured to date in a PNA homoduplex. The two bipyridines bulge out of the duplex and are aligned parallel to the major groove of the PNA. In addition, two bipyridines from adjacent PNA duplexes form a π-stacked pair that relates the duplexes within the crystal. The bulging out of the bipyridines causes bending of the PNA duplex, which is in contrast to the structure previously reported for biphenyl-modified DNA duplexes in solution, where the biphenyls are π stacked with adjacent nucleobase pairs and adopt an intrahelical geometry. This difference shows that relatively small perturbations can significantly impact the relative position of nucleobase analogues in nucleic acid duplexes.

Keywords: bipyridines · nucleic acid bending · nucleic acids · PNA · X-ray diffraction

[a] Prof. J. I. Yeh, Dr. S. Du
Department of Structural Biology
University of Pittsburgh Medical School
Pittsburgh, PA 15260 (USA)
Fax: (+1412-648-9008
E-mail: jiyeh@pitt.edu

[b] Dr. E. Pohl
Department of Chemistry and
School of Biological and Biomedical Sciences
Durham University, South Road, Durham DH1 3LE (UK)

[c] D. Truan
Swiss Light Source, Paul Scherrer Institute
52323 Villigen, PSI (Switzerland)

[d] Dr. W. He, Prof. C. Achim
Department of Chemistry, Carnegie Mellon University
4400 5th Ave., Pittsburgh, PA 15213 (USA)
Fax: (+1)412-268-1061
E-mail: achim@cmu.edu

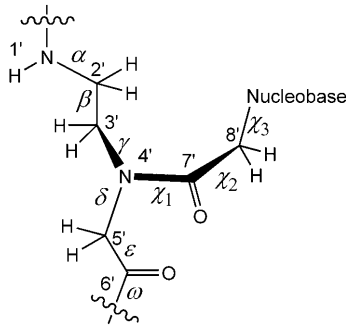
[e] Prof. G. M. Sheldrick
Institut of Inorganic Chemistry, University of Göttingen
Tammanstr. 4, 37077 Göttingen (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.201000392.

Introduction

Peptide nucleic acid (PNA) is a synthetic analogue of DNA that has a pseudo-peptide backbone instead of the sugar diphosphate backbone of the DNA (Scheme 1).[1–3] PNA forms homoduplexes by Watson–Crick hybridization. To date, crystal structures have been reported for a palindromic 6-base pair (bp) PNA duplex,[4] for another 6-bp PNA duplex with the same sequence but with a lysine at the C terminus,[5] as well as for a partly self-complementary single-stranded (ss) PNA that formed both short duplex and triplex motifs within the crystal.[6] These studies have shown that PNA homoduplexes adopt a P-type helix structure, which is more unwound and has a larger diameter than helices corresponding to A-, B- or Z-DNA duplexes.

Our research focuses on the incorporation of metal ions at specific locations within PNA duplexes to create hybrid inorganic–nucleic acid nanostructures, which have potential applications in nanotechnology and biology. The substitution of pairs of nucleobases with metal-binding ligands enables



Scheme 1. Chemical structure of PNA and definitions of torsion angles: α = C6′-N1′-C2′-C3′, β = N1′-C2′-C3′-N4′, γ = C2′-C3′-N4′-C5′, δ = C3′-N4′-C5′-C6′, ε = N4′-C5′-C6′-O1′, 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). The angle ε defines the orientation of the backbone carbonyl group. An alternative definition of ε as N4′-C5′-C6′-N1′ has been used in the literature,[4] which corresponds to the definition of the dihedral angle of ε in DNA (py = pyrimidine base, pu = purine base).

the incorporation of metal ions into PNA or other nucleic acid duplexes, and it allows control of the number of metal ions and of their position in the duplexes. Numerous DNA and PNA duplexes incorporating one or multiple metal complexes have been studied in the last decade.[7–11] In most of these studies, the ligands incorporated in the nucleic acid duplex were aromatic and the metal complexes were square-planar, 1) to make possible π-stacking between the complex and adjacent base pairs, and 2) to minimize structural distortions of the duplex. Whereas the thermal stability of the ligand- and metal-containing nucleic acids has been routinely studied, structural information about the duplexes before and after the incorporation of ligands and metal ions has been obtained only in a few cases. Such information is important for the rational design of metal-containing, nucleic acid-based nanostructures.

The crystal structure of a modified DNA duplex that had an alternating purine–pyrimidine sequence of natural nucleobases and contained two [dipic•Cu^{2+}•py] complexes, where dipic and py stand for pyridine-2,6-dicarboxylate and pyridine, respectively, revealed a Z-type conformation for the duplex.[12] Each of the two Cu^{2+} ions situated in the duplex had a distorted octahedral coordination with four equatorial coordination sites occupied by donor atoms from the dipic and the py ligands, and two axial sites occupied by monodentate nucleobases from the base pairs adjacent to the metal complex. The very recent crystal structure of a DNA duplex that contained three central, adjacent [imidazole•Ag^{+}•imidaz ole] complexes has showed that the duplex adopts a B-like structure, in which the Ag^{+} ions are aligned on the helical axis with Ag–Ag distances of 3.9 Å.[13] Based on molecular modeling, Brostchi et al. proposed a "zipper-like" model for the stacking of 2,2′-bipyridine and biphenyl in DNA duplexes.[14] In this model, the distal pyridyl rings of bipyridine ligands situated in complementary positions in the DNA are alternatively stacked, with the DNA backbone being axially stretched to accommodate the pair of stacked bipyridines. A recent NMR spectroscopy and molecular dynamics study of a 10-base pair DNA duplex containing a central pair of asymmetric biphenyls showed that the biphenyls are intercalated in the duplex, with each of the two biphenyls being π stacked head-to-tail between the other biphenyl and a neighboring nucleobase pair.[15] The overall structure of the DNA duplex containing the biphenyl displayed only minor differences when compared with B-DNA.

Although the crystal structures of several PNA duplexes have been published,[4–6] no crystal structure of a ligand-modified PNA has been reported to date. In this paper, we describe the crystal structure of an 8-base pair PNA duplex with a palindromic sequence (8-bp PNA, see Scheme 2), and of a PNA duplex with the same sequence of nucleobases but with a pair of bipyridine (bipy) ligands situated in the middle of the duplex (bipy PNA). The high resolution of these structures, 1.22 and 1.10 Å, respectively, allows us to identify atomic details of the structures, which have not been observed in previously reported crystal structures of PNA homoduplexes that had resolutions of 1.70–2.60 Å.[4–6]

12345678
8-bp PNA: H-GGCATGCC-Lys-NH_2
H_2N-Lys-CCGTACGG-H

1234 5 6789
Bipy PNA: H-GGCA**Bipy**TGCC-Lys-NH_2
H_2N-Lys-CCGT**Bipy**ACGG-H

Scheme 2. Sequences of non-modified and bipyridine-modified PNA duplexes.

Results

With more than 950 non-hydrogen atoms in the asymmetric unit, including more than 200 non-hydrogen atoms from ordered solvent and water molecules, the bipy PNA structure represents the second largest novel equal-atom structure solved by ab intio methods, with the largest being the antibiotic feglymycin, which has 1033 non-hydrogen atoms.[16] Although the 8-bp PNA structure is smaller than the one of bipy PNA, with more than 500 atoms in the asymmetric unit, it is still one of the largest all-equal atom structures solved to date by direct methods.

The crystal structures of the 8-bp PNA and bipy PNA have a resolution of 1.22 and 1.10 Å, respectively. This resolution made possible the observation of geometric and conformational details possible. Four PNA strands comprise the asymmetric unit in both structures (Figure 1 and Figure S1 in the Supporting Information). These strands form two antiparallel duplexes, one of which has a left-handed (L) configuration and the other a right-handed (R) configuration. The duplexes are stacked head-to-tail within the crystal leading to a continuous, curvilinear network of alternating right- and left-handed helices (Figure S1 b in the Supporting Information). The only disordered region is found for the C-terminal lysines of the 8-bp PNA, which cannot be traced beyond the C_β position in any of the four strands. However,

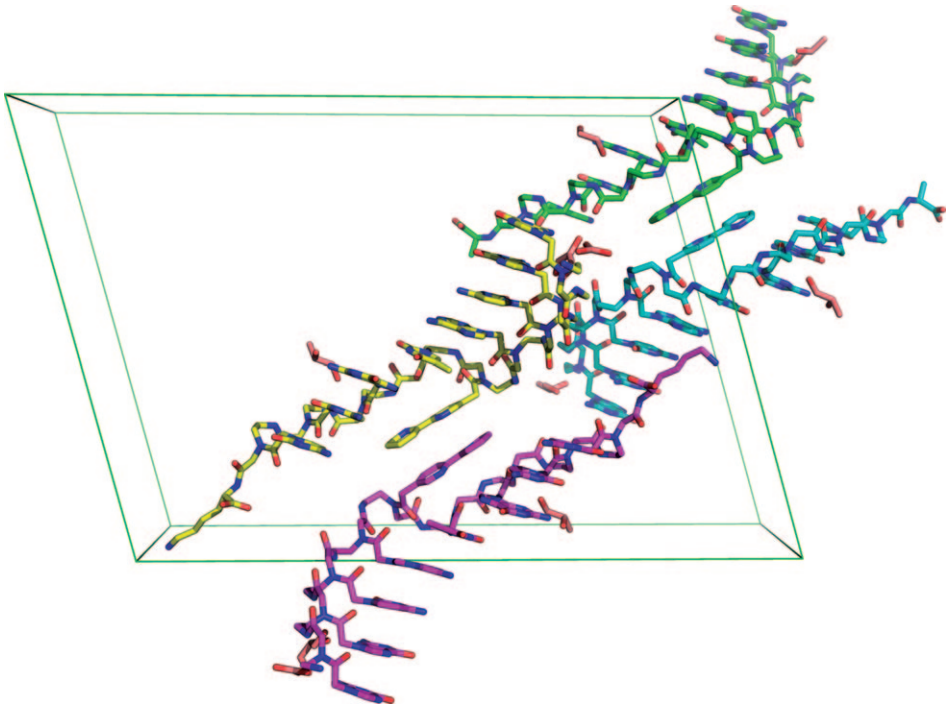


Figure 1. The four single-stranded PNAs in the asymmetric unit of the bipy PNA crystal structure shown in ball-and-sticks representation, with oxygen atoms in red, nitrogen in blue, and carbon atoms of each strand differentiated by color (yellow, green, cyan, and purple). Several well-ordered ethylene glycol molecules (carbon atoms in orange) bound to the PNA are also visible.

in the bipy PNA, two of the four lysines within the unit cell are well ordered, with low thermal factors (23–28 $Å^2$) and strong electron densities (Figure 2 and Figure S2 in the Supporting Information).

The nucleobases and the PNA backbone are very well resolved in the electron density maps of both 8-bp PNA and bipy PNA. All nucleobases participate in Watson–Crick hydrogen bonding. The carbonyl groups of the linkers that connect the nucleobases and the PNA backbone point towards the C terminus of the PNA (Figure 3). The amide groups of the backbone are in *trans* configuration and the majority of the backbone carbonyl groups are directed towards the C terminus of the PNA and into the solution (Figure 3 and Figure S3 in the Supporting Information). In 8-bp PNA, two backbone carbonyl groups, specifically those of the terminal cytosine C8 and of the thymidine T5 have opposite orientation, namely towards the N terminus of the PNA and the minor groove of the duplex (Figure S3 in the Supporting Information). For both nucleobases, a well-ordered water molecule is located at the position that would be occupied by the oxygen atom of the backbone carbonyl group if they were oriented away from the duplex (Figure 4b). The same properties are observed for the backbone carbonyls of bipy PNA, with only those of the nucleobases A4 and T6 being oriented towards the duplex and with a water molecule found at the position that would be occupied by the oxygen atom of the carbonyl group in the "away from the duplex" configuration (Figure 4).

Within the asymmetric unit of the bipy PNA, the two L-lysines situated in the oligomers with *R* configuration are well resolved (Figure 2 and Figure S2 in the Supporting Information). They have an extended conformation stabilized by several hydrogen bonds formed between atoms in the lysine and atoms of the nucleobases or the backbone of adjacent duplexes. One of these bonds is formed between the peptide carbonyl group of the lysine and the H8 proton of a terminal G, and two other hydrogen bonds are formed between the C=O and NH_2 groups of the $CONH_2$ end of the lysine, and the terminal NH_2 and the last peptide C=O groups, respectively, of the same duplex to which G belongs (Figure 2). Two more hydrogen bonds anchor the side chain amino group of the lysine to C=O groups of two backbone peptide groups of another PNA duplex.

The central bipy ligands of the bipy PNA are situated outside the duplex and form inter-duplex, π-stacked bipy–bipy pairs that bridge adjacent duplexes with identical handedness in the unit cell. The orientation of the bipy rings is almost parallel to the major grooves (Figure 4). The backbone of the protruding bipy is compressed, so that the A–T and T–A base pairs flanking the bipy are π stacked and form a base pair step with a geometry that is similar to that of the rest of the base pairs in the duplex (Figure 4c and Table S1 in the Supporting Information).

Discussion

The crystal structures of PNA homoduplexes reported to date had resolutions of 1.70–2.60 Å.[4–6] The two structures of non-modified and bipyridine-modified PNAs reported in this paper have higher resolution, namely 1.22 and 1.10 Å, respectively. Consequently, we have obtained additional details of the structures, specifically in the area of the terminal lysines and the solvent situated in closest proximity of the PNA.

The major features of the non-modified 8-bp PNA helix and of the two halves of the bipy PNA duplex that flank the bipyridines coincide with those observed for other PNAs. For example, helical parameters and the torsion angles of the 8-bp PNA duplex are similar to those characteristic for a P-form helix (Table 1 and Table S3 in the Supporting Infor-

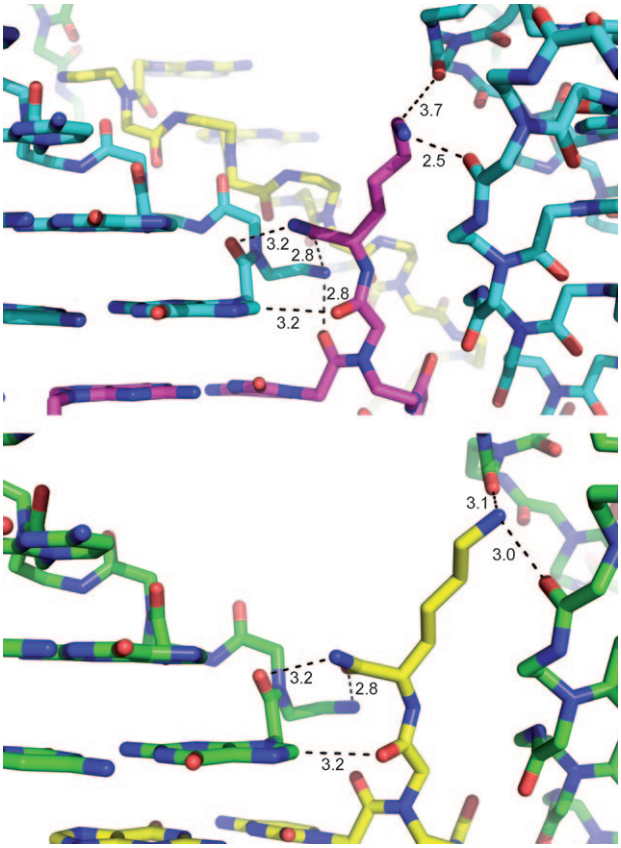


Figure 2. Structural details of the hydrogen bonds that involve two of the four terminal lysines in the unit cell of bipy PNA and atoms from the G1 nucleobase and the PNA backbone. Hydrogen bonds are shown by dashed lines and their bond length [Å] is written next to the dashed lines.

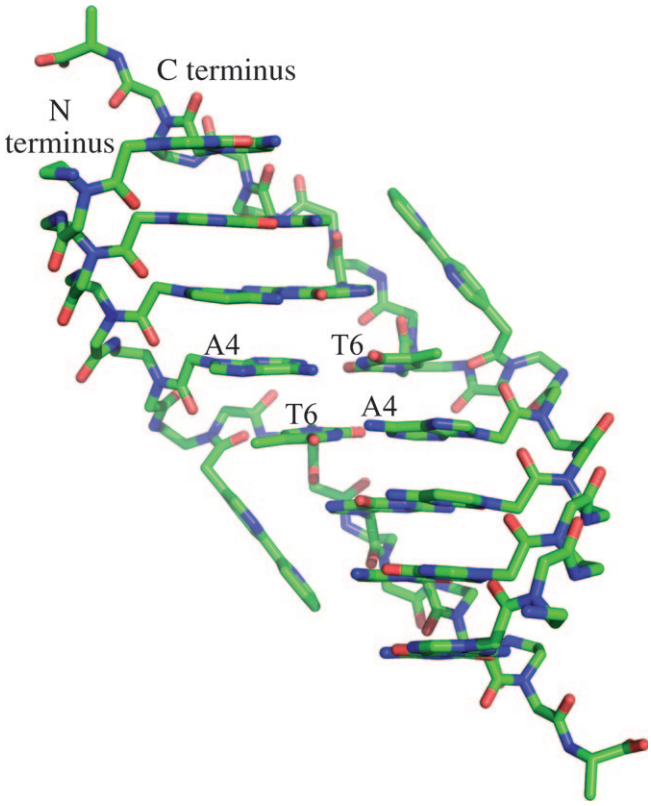


Figure 3. Structural features of the bipy PNA duplex. Note the out-of-duplex position of the central bipyridines. The figure shows the N and C termini of the two PNA strands at one of the ends of the PNA duplex and the central base pairs of the PNA duplex.

mation), with each of the torsion angles α–χ_3 of the non-terminal bases (i.e., G2–C7), having a relatively narrow distribution similar to that observed in other PNA duplexes (Figure S4 in the Supporting Information). The 8-bp PNA duplexes have large base pair displacement from the helical axis and small twist angles (Table 1 and Table S3 in the Supporting Information). Consequently, they are wider and more unwound than B-DNA duplexes. As it was the case in previously reported PNA structures, the carbonyl groups of the linkers that connect the nucleobase to the backbone are oriented towards the C terminus of the PNA oligomer, and the majority of backbone carbonyls are oriented towards the C terminus of each strand and towards the solution. Only the backbone carbonyl groups of the terminal cytosine and central thymine monomers point towards the N terminus and into the duplex. It is interesting to note that in the crystal structure of a 6-bp PNA duplex, there were also the terminal cytosines and the middle thymines that had their backbone carbonyls oriented towards the N terminus and into the duplex in 50% of the PNA molecules.[4] The flipping of the carbonyl group in a minority of the PNA monomers is unlikely to be a consequence of lattice–packing interactions because it occurs in PNAs that crystallize in different space groups. Furthermore, analysis of the interactions between symmetry-related molecules in crystals of the 8-bp PNA showed that the region surrounding both types of carbonyls have similar steric constraints.

In previous crystal structures of PNA, the conformation of the terminal lysines could not be determined due to the flexibility of the side chain of this amino acid. For example, in the crystal structure of a 6-bp PNA duplex containing a C-terminal lysine in each oligomer, the lysines had B-factors higher than the rest of the PNA.[5] Also, the side chains of the lysines pointed into solvent channels within the crystal and there were no contacts between the lysine and the PNA. On the other hand, molecular mechanics calculations identified low-energy conformations containing hydrogen bonds between the ammonium group of the L-lysine side chain and the linker carbonyl oxygen atom or the backbone amide NH, besides cation–π interactions between RNH_3^+ and guanine.[5] In the current structure of bipy PNA, both the backbone and the side chain of two terminal L-lysines within the asymmetric unit are involved in a relatively large hydrogen-bonding network with atoms of other PNA duplexes and consequently are well ordered (Figure 2). Whereas it is possible that the hydrogen bonds in which the lysines are involved are due to packing interactions, they confirm the ability of lysine to participate in hydrogen bonds similar to those identified in the molecular mechanics calculations.

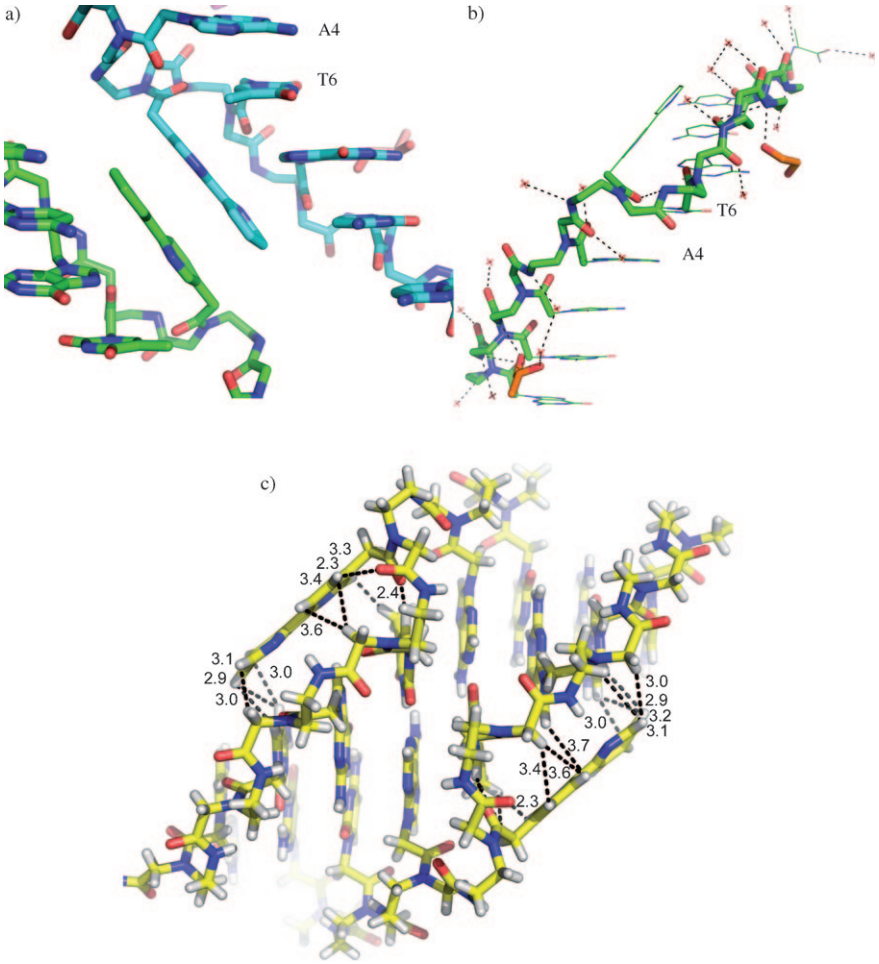


Figure 4. a) Pair of π-stacked bipyridine ligands in the crystal structure of bipy PNA. The view of the backbone of bipy PNA shows the orientation of the backbone carbonyl group of T6 towards the N terminus of the PNA and the minor groove of the duplex in contrast to that of the carbonyl groups of A4 and G7, which are oriented towards the C terminus and away from the duplex. b) PNA backbone for the central base pairs in bipy PNA and water molecules (red crosses) that are situated closest to the PNA backbone. One solvent molecule is also represented in orange/red. c) Relative orientation of bipyridines with respect to the bipy PNA duplex from which they are extruded. Shortest atom–atom distances are identified by dashed lines and their bond lengths are given in Å.

Based on molecular dynamics modeling, Brotschi and coworkers suggested that two or more bipyridines situated in complementary positions in a DNA duplex form a zipper-like motif in which the distal pyridine rings of bipyridines from opposite strands participate in π-stacking interactions.[14] Recent NMR studies demonstrated the validity of this model for a pair of biphenyl groups situated in a DNA duplex.[15] We found that the bipy ligands of the bipy PNA are bulged out of the PNA duplex and form π-stacked pairs that bridge PNA duplexes that are adjacent within the crystal (Figure 4). The backbone of the extruding bipy is compressed so that the A–T and T–A base pairs flanking the bipy are π stacked; the rise between these two base pairs is ≈3 Å, which is comparable to that of the other base pair steps (Table S1 in the Supporting Information). The torsion angles of the central bipy5 monomer and some of the torsion angles of the PNA monomers adjacent to bipy5 have values distinct from those of the other PNA monomers of typical PNA duplexes (Table S4 in the Supporting Information). The difference in the bipy arrangement in the PNA and the arrangement of biphenyl in DNA duplexes may be due to the differences between the chemical nature of the two nucleic acids or to the packing effects manifested in the crystal.

The two duplexes 8-bp PNA and bipy PNA have the same melting temperature, namely 63 °C. This result suggests that the two bipyridines of the bipy PNA are outside the duplex in solution too, and thus, do not have an effect on the thermal stability of the bipy PNA duplex. In previous studies,

Table 1. Helical parameters of PNA duplexes.[a]

Duplex		Helical sense	Disp. [Å]	Rise [Å]	Inclination [°]	Tilt [°]	Twist [°]	Slide [°]	Roll [°]	Bp/turn	Ref.
H-GGCATGCC-Lys-NH_2[b]	X-ray	R	−8.2	3.3	1.0	−0.3	18.8	−0.3	−0.2	19	this paper
		L	8.4	3.2	1.8	0.3	−18.8	−0.2	0.4	19	
H-GGCATGCC-Lys-NH_2[b]	NMR	L	7.9	3.7	−5.0	−0.2	−17.3	−0.22	−0.60	21	[17]
H-CGTACG-NH_2	X-ray	R	8.3	3.2	0.3	1.0	19.8	0.06	1.81	18	[4]
H-CGTACG-L-Lys-NH_2	X-ray	L	6.6	3.4	−5.02	0	−19	−0.74	−5.12	18	[5]
		R	7.7	3.2	−5.47	0	20.5	−0.41	5.17	18	
H-GTAGATCACT-L-Lys-NH_2	X-ray	L	7.9	3.3	−8.9	0	−18	–	–	20	[6]
		R	3.5	3.4	−26.7	0	19.1	–	–	19	

[a] Global helical parameters calculated by using Curves[18] with respect to a linear helical axis.

we have found that in a 10-bp PNA duplex formed from the complementary PNA oligomers GTAG**A**TCACT-Lys-NH_2 and AGTGA**T**CTAC-Lys-NH_2 , the substitution of the A–T base pair shown in bold with a pair of bipyridine ligands produced a strong destabilization of the PNA duplex, for which the melting temperature changed from 66 to 47 °C. The melting temperature of a 9-bp PNA duplex with the same sequence as the 10-bp PNA but lacking the central A–T base pair was 63 °C. The difference between the effects of the bipyridines on the thermal stability of PNA duplexes with different sequences indicates that the effect of the ligand is modulated by the duplex sequence and suggests that the two bipyridines can adopt different orientations with respect to the duplex, depending on the PNA sequence.

The two half-duplexes situated on each side of the pair of bipyridines have helical parameters identical to each other, except for the tilt angles, which have the same absolute value but opposite signs (Table 2). The displacement of the base pairs from the helical axis is smaller and the tilt is larger for these half duplexes than in most P-type helices, but the values of these parameters are similar to those observed for the short duplex motif formed within a crystal of a single-stranded PNA.[6] The base-pair rise in each half duplex of the bipy PNA is the largest observed to date in a PNA structure. The half duplexes have relatively small twist angles and thus, are unwound, as it is typical for P-type duplexes.

Table 2. Helical parameters[a] for the half-duplexes situated on each side of the bipyridines in bipy PNA.

Duplex	X-Disp. [Å]	Inclination [°]	Rise [Å]	Tilt [°]	Twist [°]	Slide [°]	Roll [°]
R-AB	−3.04	−24.67	4.08	5.70	20.20	−0.01	8.88
	−3.04	−24.67	4.08	−5.70	20.20	−0.01	8.88
R-CD	−2.82	−24.70	4.07	4.89	20.82	−0.07	9.01
	−2.82	−24.70	4.07	−4.89	20.82	−0.07	9.01
L-EF	3.80	−21.75	3.99	6.16	−19.54	−0.03	−7.80
	3.80	−21.75	3.99	−6.16	−19.54	−0.03	−7.80
L-GH	3.75	−21.41	3.97	5.78	−19.53	−0.10	−7.35
	3.74	−21.42	3.97	−5.78	−19.53	−0.10	−7.35

[a] Global helical parameters were calculated by using Curves[18] with respect to a linear helical axis determined for each half of the PNA duplex. AB, CD, EF and GH identify the four slightly different PNA duplexes in the unit cell.

The apparent persistence length of a DNA duplex, which is the distance over which a segment of the duplex is linear, is estimated to be ≈50 nm (≈150 bp). DNA bending is important for the packaging of genetic material, the regulation of gene expression and the interaction of nucleic acids with proteins.[19,20] Short tracts of adenines (A_n tracts) cause macroscopic curvature of the DNA molecule by additive contributions of the helically-phased A_n tracts leading to overall curvature angles of up to 110°.[21,22] On the other hand, repair of damaged DNA, including DNA containing modified bases, abasic sites or mismatches, involves recognition of the damaged site by DNA repair proteins. Crystal structures and spectroscopic studies of complexes formed between the DNA and the repair proteins have shown that DNA is bent and that the bases to be repaired or the abasic site can be extrahelical.[23,28] Molecular dynamics studies of DNA have shown that kinks of 20–30° may occur at abasic sites in the DNA and that the abasic site is particularly flexible.[29,30] A recent NMR spectroscopic study of the effect of nucleotides with a conformationally-locked, bicyclo-[3.1.0]hexane pseudosugar backbone on the structure of DNA duplexes showed that such a modification leads to duplex bending towards the major groove.[31] The comparative study of the crystal structure of bipy PNA and 8-bp PNA shows that the effect of bulging out of the PNA duplex of the two bipyridines causes a ≈53° bending angle for the duplex, which is significantly larger than the corresponding angle for 8-bp PNA (≈25°) (Table S5 in the Supporting Information) and is similar to that observed for a DNA duplex containing a synthetic *cis-syn*-cyclobutane-pyrimidine-dimer-like lesion in its complex with DNA photolyase after in situ repair.[32]

Conclusion

The crystal structures of non-modified and bipyridine-modified PNAs reported in this paper have the highest resolution reported to date for a PNA crystal structure. This resolution allowed us to identify specific modes of interaction between the terminal lysines of the PNA and the backbone and nucleobases situated in the vicinity of the lysines, which are considered as an important factor in the induction of a preferred handedness in PNA duplexes. The results described in this paper support the notion that, whereas PNA typically adopts a P-type helical structure, its flexibility is relatively high. For example, the base-pair rise in bipy PNA is the largest measured to date in a PNA homoduplex (Tables 1 and 2). The significant bending of the PNA duplex that contains a pair of bipyridines situated in extrahelical positions is in contrast to the situation in DNA duplexes in solution, in which these ligands are π stacked with adjacent nucleobase pairs and adopt an intrahelical geometry. These observations show that relatively small perturbations, such as packing forces, can significantly impact the relative position of the ligands in the nucleic acid duplex. It will be of interest to examine the effect of the PNA sequence and metal coordination to the bipyridine ligands on the relative orientation of the ligands with respect to the PNA duplex. Studies aimed to address these questions are in progress in our laboratories.

Experimental Section

PNA monomer and oligomer synthesis: The bipyridine-containing PNA monomer was synthesized according to a previously published method.[33] Briefly, PNA monomers A, T, C, and G were purchased from Applied Biosystems and used without further purification. PNA oligomers were

synthesized on lysine-preloaded *p*-methylbenzhydrylamine resin by using the Boc-protection strategy.[34] After cleavage by using trifluoromethanesulfonic acid/trifluoroacetic acid (TFMSA/TFA), the PNA was precipitated by using ethyl ether and was purified by reverse-phase HPLC by using a C18 silica column on a Waters 600 System (Waters, Milford, Massachusetts (USA)). Absorbance was measured with a Waters 2996 Photodiode Array Detector. Oligomers were characterized by MALDI-TOF mass spectrometry, by using an Applied Biosystems Voyager Biospectrometry Workstation with delayed extraction by using a (*R*)-cyano-4-hydroxycinnamic acid matrix (10 mg mL^{-1} in 1:1 water/acetonitrile, 0.1% TFA). MALDI-TOF calcd/found for $[PNA+H]^+$ *m*/*z*: 8-bp PNA: 2315.24/2314.90, bipy PNA: 2611.57/2610.73. After HPLC purification, the PNA was lyophilized and then stored at −20 °C. PNA stock solutions were prepared by dissolving solid PNA in HEPES buffer (10 mM pH 7.0). The concentration of the stock solution was determined by UV/Vis spectroscopy by using ε values for nucleobases at 260 nm of 13700 M^{-1} for A, 8600 M^{-1} for T, 11700 M^{-1} for G, 6600 M^{-1} for C,[34] and 9770 M^{-1} for bipy. The concentration of the stock solutions was 2.16 mM ss PNA for 8-bp PNA and 1.91 mM for bipy PNA, which corresponds to 5 mg PNA per mL. To form the PNA duplexes, the solutions were annealed by heating to 95 °C for 5 min, followed by slow cooling to room temperature.

Crystallization

8-bp PNA: Single crystals of the non-modified 8-bp PNA duplex were grown from a crystallization (reservoir) solution comprised of 40% (v/v) ethylene glycol and Na/K phosphate buffer (0.1 M, pH 6.2). The 5 mg mL^{-1} stock solution of the 8-bp PNA duplex in HEPES buffer (10 mM, pH 7.0) was used for screening and final optimization. All crystallization setups were vapor diffusion, in which 1 μL of PNA stock solution was added to 1 μL of reservoir solution on a cover slip, then suspended over 0.7 mL of reservoir solution. All trays were maintained at 20 °C. Nucleation was observed after five days and crystals continued to grow over two weeks to yield plate-like crystals, the largest with final dimensions of up to 0.08 × 0.08 × 0.04 mm. The ethylene glycol in the crystallization solution served dual roles of precipitant and cryoprotectant, which permitted the crystals to be mounted in a rayon loop and flash-cooled to 100 K for data collection.[35]

Bipy PNA: Single crystals of the bipyridine-containing PNA duplex were grown from a crystallization (reservoir) solution comprised of 40% (v/v) 1,2-propanediol, citrate (0.1 M, pH 5.5), and NaCl (0.2 M). A stock solution of bipy PNA at 5 mg mL^{-1} and crystallization setups were prepared as described above for the non-modified PNA Nucleation was observed after about one week, with continued crystal growth over three weeks before reaching final crystal dimensions of approximately 0.1 × 0.05 × 0.05 mm with rod-like morphology. Immediately before crystal mounting, ethylene glycol was added to the well solution to a final concentration of 10% (vol/vol), to supplement the 1,2-propanediol in the crystallization solution, permitting the crystals to be flash-cooled to 100 K for data collection.

Structure determination: A combination of direct methods in conjunction with maximum likelihood residual refinement and/or density modification was used to solve and iteratively build and refine the crystal structures of 8-bp PNA and bipy PNA.

Bipy PNA: Complete data sets were collected from two crystals on the undulator beam line X06SA at the Swiss Light Source equipped with a mar165 CCD detector with the parameters shown in Table 3. Since each data set alone could not easily be solved by direct methods, a more complete data set (overall completeness 97%, 90.7 in the last shell) was obtained by merging the two data sets by using XPREP (P9-merged in Table 4).[36] The merged data set was solved by SHELXD[37] by using the dual space strategy.[38] Density modification as implemented in SHELXE[39] significantly improved the resulting electron density and allowed us to build the first model comprising of the PNA backbone and all bases with COOT.[40] The structure determination was completed by iterative cycles of model building and refinement with SHELXL.[41] Local NCS symmetry restraints were only used in the initial phases of refinement in order to improved convergence.[42] The course of the refinement was carefully monitored by using 5% of the reflections to calculated free R factors,[43] and completed by restrained anisotropic refinement against all data by using a conjugate gradient method as implemented in SHELXL.[41] All non-hydrogen atoms were refined anisotropically by using standard bond lengths and angles and applying suitable rigid-bond and similarity restraints. Additional isotropic restraints were applied to solvent molecules and parts of the molecule that showed unreasonable ADP ellipsoids. Additional crystallographic data are summarized in Table 4.

Table 3. Data collection phasing and refinement statistics for two bipy PNA crystals.

Data set	P9-I	P9-II
beam line	X10SA, SLS	X10SA, SLS
temperature [K]	100	100
wavelength [Å]	0.8	0.8
oscillation range [°]	0.5	0.5
detector distance [mm]	75	75
a [Å]	38.20	38.1
b [Å]	25.73	25.7
c [Å]	53.05	52.90
α [°]	90	90
β [°]	105.3	105.3
γ [°]	90	90
space group	*P*2	*P*2
resolution range [Å]	50–1.05	50–1.05
reflections	277897	249854
unique reflections	42731	43541
redundancy	6.5	5.7
completeness[a]	92.1 (71.9)	93.4 (66.7)
R_{merge}	0.06 (0.43)	0.05 (0.35)
$I/\sigma(I)$	23.3 (1.7)	29.0 (2.4)

[a] Last shell 1.05–1.09.

Table 4. Structure refinement parameters for 8-bp PNA and bipy PNA.

	P9-merged[a]	P8-III
resolution [Å]	50–1.1	50–1.2
no. of data	38601	29399
no. of restraints	11775	9536
no. of parameter	8683	7484
no. of non-hydrogen atoms	963	831
no. of solvent molecules[b]	8	5
no. of carbonate	2	0
no. of water	171	189
w*R*2 (all data)	0.356	0.454
$R1[F>4\sigma]$	0.127	0.179
$R1_{free}$ $[I>4\sigma]$	0.169	0.220

[a] P9-merged is for the merged data from the datasets of Table 3. [b] Ethylene glycol, glycerol, or propanediol.

The final model contains four PNA strands, two carbonate ions, and several solvent molecules, including ethylene glycol, glycerol, or propanediol from the crystallization solution. Water molecules were added by using COOT selected from peaks in the Fo–Fc map with a minimum height of 3.5 sigma and good hydrogen-bond geometry. The 184 water molecules were included in the refinement. The four PNA strands display a pseudo-inversion center. However, the terminal lysine residues, which were only partially visible in the electron density map, and the central bipyridines break this symmetry. Moreover, the backbone shows significant deviations from the pseudo-inversion center. As a consequence, the rmsds for the superposition of an inverted strand with its pseudo-symmetry mate are approximately 0.5 Å, which clearly justifies the refinement in the non-centrosymmetric space group *P*2. The refinement converged at *R* factors of $R=0.128$ and $R_{free}=0.168$.

8-bp PNA: Solving the structure of 8-bp PNA proved to be challenging, which is most likely due to the fact that the crystals were (at least partially) non-merohedrally twinned. The diffraction patterns clearly showed

Table 5. Data collection parameters for 8-bp PNA.

Data set	P8-I	P8-II	P8-III
beam line	X10SA, SLS	X10SA, SLS	ID23B, APS
temperature [K]	100	100	100
wavelength [Å]	0.8266	0.8266	1.0
oscillation range [°]	0.5	0.5	1.0
a [Å]	18.87	18.91	18.88
b [Å]	28.86	28.93	28.91
c [Å]	54.94	55.07	54.98
α [°]	87.9	88.0	88.1
β [°]	85.8	86.1	85.8
γ [°]	79.5	79.4	79.6
space group	$P1$	$P1$	$P1$
resolution range [Å]	50–1.1	50–1.1	50–1.27
reflections	156663	67947	81240
unique reflections	43264	43541	30089
redundancy	3.6	1.7	2.3
completeness	94.6 (91.4)[a]	88.6 (78.4)	91.7 (76.5)[b]
R_{merge}	0.11 (0.47)	0.06 (0.25)	0.06, (0.19)
$I/\sigma(I)$	11.8 (2.4)	12.4(2.7)	21.2 (2.2)

[a] Last resolution shell: 1.1–1.4 Å [b] Last resolution shell: 1.27–1.47 Å.

two (or more) components, which hampered the solution and refinement as described below. Two data sets, P8-I and P8-II in Table 5, were collected from two crystals on beam line X10SA at the Swiss Light Source (SLS) by using a marmosaic 225 CCD detector.[44] These two data sets were integrated with HKL2000.[45] Since the individual data proved difficult to solve, the two SLS data sets (P8-I and P8-II) were merged by using XPREP,[36] which did not affect the completeness significantly. However, this merged data set was finally solved with SHELXD.[46] The best solution resulted in a correlation coefficient of only 65%. Consequently, whereas the electron density map showed clearly-defined solvent regions and some helical features, it was not sufficient for model building. Density modification with SHELXE increased the correlation coefficient to 78% and the resulting map was easily interpretable.[36]

The course of the refinement was again monitored by using 5% of the reflections to calculated free R factors,[43] and completed by restrained anisotropic refinement against all data by using a conjugate gradient method as implemented in SHELXL.[39] All non-hydrogen atoms were refined anisotropically by using standard bond lengths and angles and applying suitable rigid-bond and similarity restraints. Additional isotropic restraints were applied to solvent molecules and parts of the molecule that showed unreasonable ADP ellipsoids. Nevertheless, the refinement was stuck at R factors around 28%, presumably due to the non-merohedral twinning observed during data collection.

In order to resolve this issue, a third data set for a crystal obtained from a slightly modified crystallization condition that contained additional ethanol (10%) in the crystallization solution was collected at Argonne National Laboratory, beam line ID23B (P8-III in Table 5). As the crystals suffered from severe radiation damage, a ten micrometer collimator was used and data were collected with one degree oscillation range per frame until the average intensity dropped per image to about 60% of the initial value, which on average related to approximately 20 degrees oscillation. Data collection was resumed after the crystal was translated by 50 micrometer until the entire crystal has been utilized. This data set was independently solved by using SHELXS[47] to identify the 50 strongest peaks as starting coordinates for the less flexible PNA peptide backbone atoms to be used in ACORN.[48]

The 8-bp PNA structure was refined following the same protocols as described above for the bipy PNA. The model contains four PNA strands—again related by a pseudo-inversion center. Since the lysine residues are for the most part disordered, the remaining PNA strands almost fulfill the pseudo-inversion center. The rmsds for the superposition of an inverted strand with its counterpart is only 0.3 Å. However, this value is still larger than the estimated uncertainties and hence, the structure was refined in the space group $P1$. The final model comprises four PNA strands, four ethylene glycol molecules and 197 well-defined water molecules. The refinement converged at $R=0.18$ ($R_{free}=0.22$). These relatively high R factors are presumably due to the fact that even these crystals suffered from partial non-merohedral twinning described above.

The structures have been deposited in the protein structure database and have numbers 3MBS and 3MBU.

Acknowledgements

This work was supported by the National Science Foundation (CHE-0347140 to C.A. and in part by CHE-9808188 to the Center for Molecular Analysis at Carnegie Mellon), Deutsche Forschungsgemeinschaft (SH14/5 to G.M.S.), The National Institutes of Health (GM066466 and AI76121 to J.I.Y. for X-ray facility support), and the Pennsylvania Department of Health (to J.I.Y.). W.H. acknowledges support from an Astrid and Bruce McWilliams Fellowship. We thank the staff members at the General Medicine and Cancer Institutes Collaborative Access Team (GM/CA-CAT) and the Southeast Regional Collaborative Access Team, both at the Advanced Photon Source, Argonne National Laboratory, and the Paul Scherrer Institute (X10SA) at the Swiss Light Source for access and technical assistance. GM/CA CAT has been funded in whole or in part by the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104). Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

[1] P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science* **1991**, *254*, 1497.
[2] P. E. Nielsen, M. Egholm, *Curr. Issues Mol. Biol.* **1999**, *1*, 89.
[3] C. Achim, B. A. Armitage, D. H. Ly, J. W. Schneider in *Wiley Encyclopedia of Chemical Biology*, Wiley, New York, **2008**.
[4] H. Rasmussen, J. S. Kastrup, J. N. Nielsen, J. M. Nielsen, P. E. Nielsen, *Nat. Struct. Biol.* **1997**, *4*, 98.
[5] H. Rasmussen, T. Liljefors, B. Petersson, P. E. Nielsen, J. S. Kastrup, *J. Biomol. Struct. Dyn.* **2004**, *21*, 495.
[6] B. Petersson, B. B. Nielsen, H. Rasmussen, I. K. Larsen, M. Gajhede, P. E. Nielsen, J. S. Kastrup, *J. Am. Chem. Soc.* **2005**, *127*, 1424.
[7] W. He, R. M. Franzini, C. Achim, *Prog. Inorg. Chem.* **2007**, *55*, 545.
[8] J. Müller, *Eur. J. Inorg. Chem.* **2008**, 3749.
[9] K. Tanaka, M. Shionoya, *Coord. Chem. Rev.* **2007**, *251*, 2732.
[10] G. H. Clever, C. Kaul, T. Carell, *Angew. Chem.* **2007**, *119*, 6340; *Angew. Chem. Int. Ed.* **2007**, *46*, 6226.
[11] M. Shionoya, K. Tanaka, *Curr. Opin. Chem. Biol.* **2004**, *8*, 592.
[12] S. Atwell, E. Meggers, G. Spraggon, P. G. Schultz, *J. Am. Chem. Soc.* **2001**, *123*, 12364.
[13] S. Johannsen, N. Megger, D. Böhme, R. K. O. Siegel, J. Müller, *Nat. Chem.* **2010**, *2*, 229.
[14] C. Brotschi, A. Häberli, C. J. Leumann, *Angew. Chem.* **2001**, *113*, 3101; *Angew. Chem. Int. Ed.* **2001**, *40*, 3012.
[15] Z. Johar, A. Zahn, C. J. Leumann, B. Jaun, *Chem. Eur. J.* **2008**, *14*, 1080.
[16] G. Bunkóczi, L. Vértesy, G. M. Sheldrick, *Angew. Chem.* **2005**, *117*, 1364; *Angew. Chem. Int. Ed.* **2005**, *44*, 1340.
[17] W. He, E. Hatcher, A. Balaeff, D. N. Beratan, R. R. Gil, M. Madrid, C. Achim *J. Am. Chem. Soc.* **2008**, *130*, 13264.
[18] R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1988**, *6*, 63; R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *6*, 655.
[19] L. J. Maher III, *Curr. Opin. Chem. Biol.* **1998**, *2*, 688.
[20] K. K. Swinger, P. A. Rice, *Curr. Opin. Struct. Biol.* **2004**, *14*, 28.
[21] D. Strahs, T. Schlick, *J. Mol. Biol.* **2000**, *301*, 643.
[22] C. Gohlke, A. I. H. Murchie, D. M. J. Lilley, R. M. Clegg, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11660.
[23] N. C. Horton, J. J. Perona, *J. Mol. Biol.* **1998**, *277*, 779.
[24] J. L. Kim, D. B. Nikolov, S. K. Burley, *Nature* **1993**, *365*, 520.

[25] J. F. Kugel, *Biochem. Mol. Biol. Educ.* **2008**, *36*, 341.
[26] M. Lorenz, A. Hillisch, S. D. Goodman, S. Diekmann, *Nucleic Acids Res.* **1999**, *27*, 4619.
[27] M. Lorenz, A. Hillisch, D. Payet, M. Buttinelli, A. Travers, S. Diekmann, *Biochemistry* **1999**, *38*, 12150.
[28] G. Parkinson, C. Wilson, A. Gunasekera, Y. W. Ebright, R. E. Ebright, H. M. Berman, *J. Mol. Biol.* **1996**, *260*, 395.
[29] D. Barsky, N. Foloppe, S. Ahmadia, D. M. Wilson, III, A. D. MacKerell, Jr., *Nucleic Acids Res.* **2000**, *28*, 2613.
[30] J. Curuksu, K. Zakrzewska, M. Zacharias, *Nucleic Acids Res.* **2008**, *36*, 2268.
[31] Z. Wu, M. Maderia, J. J. Barchi, Jr., V. E. Marquez, A. Bax, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 24.
[32] A. Mees, T. Klar, P. Gnau, U. Hennecke, A. P. M. Eker, T. Carell, L.-O. Essen, *Science* **2004**, *306*, 1789.
[33] R. Franzini, R. M. Watson, D.-L. Popescu, G. K. Patra, C. Achim, *Polym. Prepr. Am. Chem. Soc. Div. Polym. Chem.* **2004**, *45*, 337.
[34] P. E. Nielsen, *Peptide Nucleic Acids: Protocols and Applications*, 2nd ed., Taylor and Francis, New York, **2004**.
[35] T. Y. Ten, *J. Appl. Crystallogr.* **1990**, *23*, 387.
[36] G. M. Sheldrick, XPREP, program for reciprocal space exploration, Bruker AXS, **2007**.
[37] I. Usón, G. M. Sheldrick, *Curr. Opin. Struct. Biol.* **1999**, *9*, 643.
[38] R. Miller, G. T. DeTitta, R. Jones, D. A. Langs, C. M. Weeks, H. A. Hauptman, *Science* **1993**, *259*, 1430.
[39] G. M. Sheldrick, *Z. Kristallogr.* **2002**, *217*, 644.
[40] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126.
[41] G. M. Sheldrick, T. R. Schneider, *Methods Enzymol.* **1997**, *277*, 319.
[42] I. Usón, E. Pohl, T. R. Schneider, Z. Dauter, A. Schmidt, H.-J. Fritz, G. M. Sheldrick, *Acta Crystallogr. Sect. D* **1999**, *55*, 1158.
[43] A. T. Brünger, *Nature* **1992**, *355*, 472.
[44] E. Pohl, C. Pradervand, R. Schneider, T. Tomizaki, A. Pauluhn, Q. Chen, G. Ingold, Z. Zimoch, C. Schulze-Briese, *Synch. Rad. News* **2006**, *19*, 24.
[45] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307.
[46] Bruker AXS, G. M. Sheldrick, **2007**.
[47] G. M. Sheldrick, *Acta Crystallogr. Sect. A* **1990**, *46*, 467.
[48] J. X. Yao, E. J. Dodson, K. S. Wilson, M. M. Woolfson, *Acta Crystallogr. Sect. D* **2006**, *62*, 901.

Received: February 13, 2010
Published online: September 21, 2010

