

# Continuous $\beta$ -turn fold of an alternating alanyl/homoalanyl peptide nucleic acid

Jose A. Cuesta-Seijo,<sup>a,b</sup> Jinhua Zhang,<sup>c</sup> Ulf Diederichsen<sup>c</sup> and George M. Sheldrick<sup>a\*</sup>

<sup>a</sup>Lehrstuhl für Strukturchemie, Institut für Anorganische Chemie, Georg-August-Universität Göttingen, Tammannstrasse 4, Göttingen, 37077, Germany, <sup>b</sup>Carlsberg Laboratory, Gamle Carlsberg Vej 10, Copenhagen V, 1799, Denmark, and <sup>c</sup>Institut für Organische und Biomolekulare Chemie, Georg-August-Universität Göttingen, Tammannstrasse 2, Göttingen, 37077, Germany

Correspondence e-mail:  
gsheldr@shelx.uni-ac.gwdg.de

The crystal structure of the PNA (peptide nucleic acid) oligomer H-Lys-HalG-AlaG-HalC-AlaG-HalC-AlaC-Lys-NH<sub>2</sub> (PNA1, amino acids with D-configuration are underlined, Ala = alanyl, Hal = homoalanyl) has been determined by *ab initio* direct methods and refined against 1.0 Å data. The asymmetric unit consists of a tetrameric cage with almost ideal Watson-Crick C-G base pairing of all the guanine and cytosine side-chain substituents. Each PNA strand has a 90°  $\beta$ -turn every second residue, stabilized by three hydrogen bonds between the backbone amides. The first, second, fifth and sixth bases stack on one side of the monomer and pair with the corresponding complementary bases of a second monomer to form a dimer. The two remaining bases on each side of the resulting dimer form Watson-Crick pairs with the complementary bases of a second dimer, leading to a unique cage structure. The extra methylene groups in the homoalanyl residues enable stacking of the bases with an optimal distance between base-planes but also with an appreciable lateral displacement (slide).

Received 8 March 2012

Accepted 10 May 2012

PDB Reference: PNA1, 3c1p.

## 1. Introduction

PNAs (peptide nucleic acids) consist of a polyamide backbone with side chains to which nucleobases are attached. They bind to DNA or other PNA strands by sequence-specific base pairing (Nielsen *et al.*, 1991). In contrast to DNA, the PNA backbone is uncharged, thermally stable and resistant to enzymatic degradation, which may make PNAs suitable for a number of medical and other applications (Ray & Nordén, 2000) and led to the suggestion (Nelson *et al.*, 2000) that similar molecules may have served as genetic material in a pre-RNA world. Crystal structures of a PNA duplex (Rasmussen *et al.*, 1997) and a PNA-DNA hetero-duplex (Menchise *et al.*, 2003) confirmed the formation of wide helices held together by Watson-Crick base pairing with the bases stacked perpendicular to the helix axis.

Diederichsen (1996) showed that a base-stacking distance close to the optimal value of 3.5 Å could also be achieved with a polypeptide backbone, provided that the alanyl or homoalanyl amino acids that carry the canonical purine and pyrimidine nucleobases have alternating D and L configurations so that the bases protrude on the same side of a strand. D/L-Alternating peptides with proteinogenic amino acids were shown to fold in a double-stranded right-handed antiparallel  $\beta$ -helix (Alexopoulos *et al.*, 2004). Alanyl peptide nucleic acids (alanyl-PNAs) were predicted to form a double strand by base

**Table 1**

Crystal data, data processing and refinement.

Data in parentheses are for the highest-resolution shell.

Space group	$P2_1$
Unit-cell dimensions (Å and °)	$a = 26.31, b = 30.73, c = 33.55, \beta = 99.53$
Resolution range (Å)	40.0–1.00 (1.09–1.00)
No. of crystals used	1
No. of observed reflections	357007
No. of unique reflections	28573
$R_{\text{merge}}(I)$	0.061 (0.413)
Completeness (%)	99.5 (98.7)
Average $I/\sigma(I)$	20.99 (5.16)
$R_{\text{work}}$	0.150
$R_{\text{free}}$	0.199
No. of non-H atoms	
PNA	396
Water	134
Average $B$ factors (Å <sup>2</sup> )	
PNA	17.9
Water	35.3
R.m.s. deviations from ideality	
Bond lengths (Å)	0.015
Bond angles (°)	1.67

pairing with a linear, stretched peptide backbone (Diederichsen, 1997). The duplex formation of alanyl-PNAs shows a clear sequence dependency and can be fine-tuned by replacing alternate alanyl units by homoalanyl, which has an extra methylene group between the  $\beta$ -carbon and the nucleobase (Diederichsen, 1997). In alternating alanyl/homoalanyl PNA sequences only alanyl/homoalanyl base pairs would be expected to have the correct relative orientation and distance between the bases (Diederichsen & Schmitt, 1996).

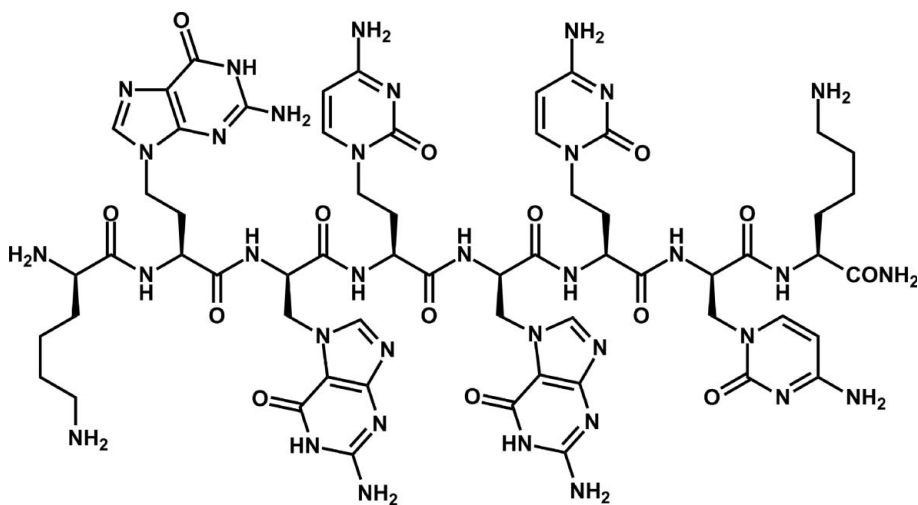
Since no crystal structure of a PNA with a peptide backbone appears to have been reported, we have determined the structure of a D/L-alternating alanyl/homoalanyl-PNA sequence to obtain information about the secondary structure and the nucleobase-pairing interactions. The PNA oligomer H-Lys-HalG-AlaG-HalC-AlaG-HalC-AlaC-Lys-NH<sub>2</sub> (PNA1,

amino acids with D-configuration are underlined, Ala = alanyl, Hal = homoalanyl) was synthesized by solid-phase peptide synthesis as previously described (Diederichsen *et al.*, 2005). Lysine was included to provide solubility in aqueous buffer. Although the pair of complementary alanyl-PNA sequences H-Lys-AlaG-AlaG-AlaC-AlaG-AlaC-AlaC-NH<sub>2</sub> and H-Lys-AlaG-AlaG-AlaC-AlaG-AlaC-AlaC-NH<sub>2</sub> interact to provide a clearly defined melting curve with double-strand stability of  $T_m = 331$  K (6  $\mu$ M, 0.1 M NaCl, 0.01 M Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub>, pH 7) as determined by temperature-dependent UV spectroscopy, the self-complementary alanyl/homoalanyl PNA1 gives a melting curve with two transitions at  $T_m = 283$  K and  $T_m = 329$  K (Zhang, 2004).

## 2. Materials and methods

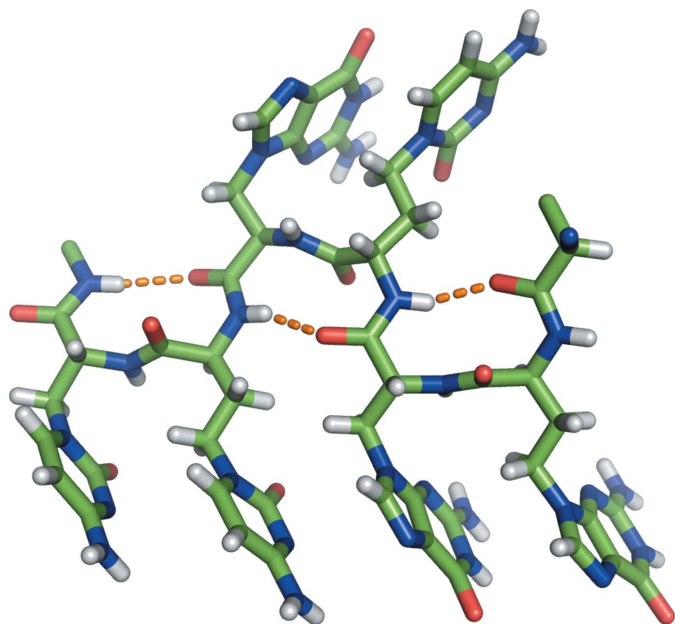
Oligomer PNA1 (H-Lys-HalG-AlaG-HalC-AlaG-HalC-AlaC-Lys-NH<sub>2</sub>, Fig. 1) was prepared by Boc solid-phase synthesis, purified with HPLC [RP C18 Jsphere ODS/H80 25 cm, H<sub>2</sub>O, 0.1% TFA (CF<sub>3</sub>COOH), gradient 5 to 30% CH<sub>3</sub>CN, H<sub>2</sub>O (10%), 0.1% TFA,  $t_R = 13.2$  min] and characterized by ESI mass spectrometry (759.1 [ $M + 2H$ ]<sup>2+</sup>, 759.1 [ $M + 3H$ ]<sup>3+</sup>). Crystals were grown by the hanging drop method. 4  $\mu$ l of a 2 mg ml<sup>-1</sup> aqueous solution of PNA1 was mixed with 4  $\mu$ l of reservoir solution consisting of 2.3 M 1,6-hexanediol, 0.2 M MgCl<sub>2</sub> and 0.1 M Tris-HCl buffer at pH 8.5. The hanging drop was incubated over 1 ml of the same reservoir solution at 283 K. Clusters of crystals formed but proved very difficult to manipulate because the 1,6-hexanediol also crystallized. Eventually, it was possible to separate and mount a 0.15 mm long and 0.10 mm thick crystal in a nylon loop with the help of an acupuncture needle immediately after transferring a drop to a 277 K room. The crystal was flash cooled in liquid nitrogen without the use of extra cryoprotectant. Data were collected at 100 K on beamline X10SA at the Swiss Light Source with a wavelength of 0.7999 Å and processed with XDS (Kabsch, 2010); details are given in Table 1.

The structure was solved by direct methods using SHELXD (Usón & Sheldrick, 1999; Sheldrick, 2010), which found 381 atoms. Instead of using these atoms directly to build a model, they were subject to density modification in SHELXE (Sheldrick, 2002, 2010), which also enabled inversion of the enantiomer (there is a 50% chance that direct methods, as in this case, produce a mirror image of the structure). The resulting map showed virtually every atom and had a weighted mean phase error against the final refined structure of 18.75°. An initial model was built into this density using XTALVIEW/XFIT (McRee, 1999). The structure was refined with SHELXL (Sheldrick, 2008) using standard geometrical restraints



**Figure 1**

The chemical structure of PNA1 (H-Lys-HalG-AlaG-HalC-AlaG-HalC-AlaC-Lys-NH<sub>2</sub>).



**Figure 2**  
The PNA1 monomer unit, showing the three N–H...O hydrogen bonds that link the main-chain peptides giving rise to  $\beta$ -turns.

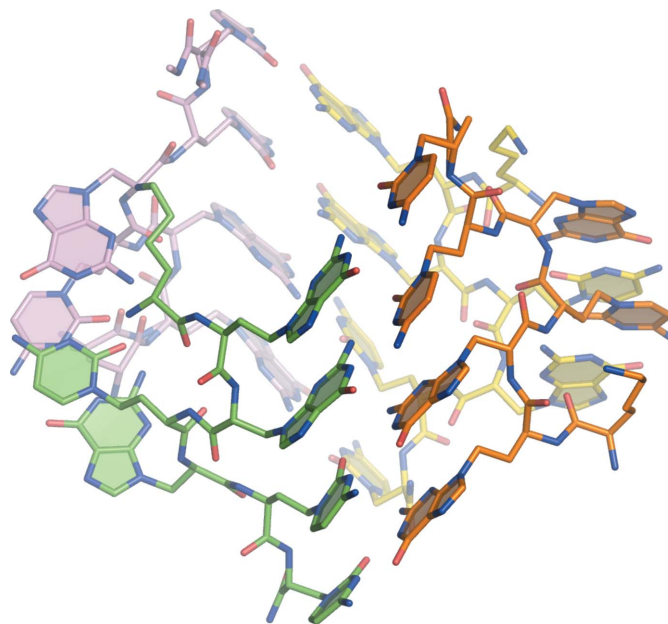
(Engh & Huber, 2001; Parkinson *et al.*, 1996), restrained anisotropic atomic displacement parameters (Sheldrick & Schneider, 1997) and riding hydrogen atoms. It has been deposited in the PDB with code 3c1p.

### 3. Results and discussion

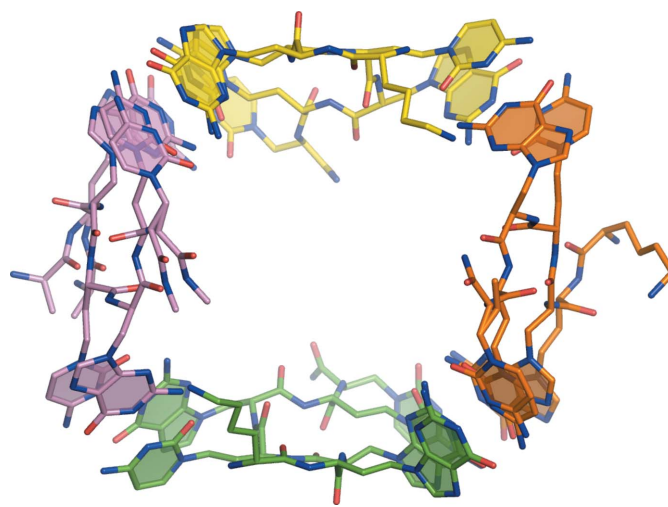
Instead of the expected extended antiparallel duplex with stacked Watson–Crick base pairs, a quite different intricate tetrameric cage structure was found for PNA1 in which all bases indeed form stacked Watson–Crick base pairs.

#### 3.1. Monomer units

Each PNA strand has a  $90^\circ$   $\beta$ -turn every second residue, with the result that the first, second, fifth and sixth bases stack on one side and bases 3 and 4 on the other side of the molecule. The Watson–Crick edges are approximately perpendicular for the two groups. The monomer (Fig. 2) is stabilized by three intramolecular N–H...O hydrogen bonds between the backbone peptides. The extra methylene group of the alternate homoalanyl moieties enables the bases to stack with an interplanar distance close to the optimal value of 3.5 Å, but with an appreciable displacement (slide) parallel to the base planes. The terminal lysines are appreciably disordered and only three of the eight could be completely localized in the electron density. Except for the lysines, the four monomers in the asymmetric unit have very similar geometries, with root-mean-square (r.m.s.) deviations in the range 0.5 to 0.7 Å after superposition. The backbone dihedral angles are about  $\varphi = -60^\circ$ ,  $\psi = 130^\circ$  for the L-amino acids and  $\varphi = 100^\circ$ ,  $\psi = -15^\circ$  for the D-amino acids; allowing for the different



**Figure 3**  
The complete tetramer cage of PNA1, which also constitutes the asymmetric unit.



**Figure 4**  
An alternative view of the tetramer, rotated by  $90^\circ$  from Fig. 3.

chiralities these correspond to well populated regions of the Ramachandran diagram.

#### 3.2. The PNA1 tetramer

The monomers combine to form antiparallel dimers by Watson–Crick pairing of the two base quartets so that HaIG pairs with AlaC and AlaG with HaIC. The HaIC and AlaG bases that remain on each side of the dimer can then pair with the complementary bases of another dimer so that a tetrameric cage results that is open at the top and bottom (Figs. 3 and 4). The base pairing is rather regular; all 36 hydrogen

bonds fall in the range 2.75 to 2.94 Å, with tilt angles between the base-planes in the range  $-1.5$  to  $2.5^\circ$  and twist angles between  $9.8$  and  $20.2^\circ$ . Full details of the hydrogen-bonding and base-pairing geometry may be found in Cuesta-Seijo (2005). The tetramer structure provides a plausible explanation for the two transitions in the melting curve mentioned at the end of §1. Below 283 K PNA1 may well be tetrameric as in the crystal, and between the transitions at 283 and 329 K the predominant oligomer is probably a dimer held together by four base pairs formed by the bases on the same side of each monomer. This behaviour may also help to explain why the best crystals grew at exactly 283 K.

#### 4. Conclusions

This rather unique and stable cage structure is probably specific for the chosen pattern of homoalanyl and alanyl residues and alternate D- and L-amino acids. However, it might be possible to replace the base pairs with other complementary base pairs without appreciable distortion of the rest of the structure. It suggests the possibility of building a stable anti-parallel PNA double helix in which only residues 1, 2, 5, 6, 9, 10, ... are alternately L-homoalanyl and D-alanyl with attached nucleobases, and the bases are chosen to be complementary. The remaining residues could be any alternate D- and L-amino acids, which would provide considerable control over solubility and interactions with other biological molecules. Based on the fairly consistent base-pair twist angles observed for PNA1, such a helix would be right-handed with about 26 bases per helix turn and a helix pitch of about 87 Å. However, the hand could be inverted by inverting the order of D- and L-amino acids.

Generous support from the Deutsche Forschungsgemeinschaft (grant Nos. DI 542/7-1 and SH 14/5-1) is gratefully

acknowledged. GMS thanks the Volkswagenstiftung for a Niedersachsenprofessur.

#### References

- Alexopoulos, E., Küsel, A., Sheldrick, G. M., Diederichsen, U. & Usón, I. (2004). *Acta Cryst.* **D60**, 1971–1980.
- Cuesta-Seijo, J. A. (2005). PhD thesis. Georg-August-Universität Göttingen, Germany. [http://webdoc.sub.gwdg.de/diss/2005/cuesta\\_seijo/](http://webdoc.sub.gwdg.de/diss/2005/cuesta_seijo/).
- Diederichsen, U. (1996). *Angew. Chem. Int. Ed. Engl.* **35**, 445–448.
- Diederichsen, U. (1997). *Angew. Chem. Int. Ed. Engl.* **36**, 1886–1889.
- Diederichsen, U. & Schmitt, H. W. (1996). *Tetrahedron Lett.* **37**, 475–478.
- Diederichsen, U., Weicherding, D. & Diezemann, N. (2005). *Org. Biomol. Chem.* **3**, 1058–1066.
- Engh, R. A. & Huber, R. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 382–392. Dordrecht: Kluwer Academic Publishers.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
- Menchise, V., De Simone, G., Tedeschi, T., Corradini, R., Sforza, S., Marchelli, R., Capasso, D. & Pedone, C. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 12021–12026.
- Nelson, K. E., Levy, M. & Miller, S. L. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3868–3871.
- Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. (1991). *Science*, **254**, 1497–1500.
- Parkinson, G., Vojtechovsky, J., Clowney, L., Brünger, A. T. & Berman, H. M. (1996). *Acta Cryst.* **D52**, 57–64.
- Rasmussen, H., Kastrup, J. S., Nielsen, J. N., Nielsen, J. M. & Nielsen, P. E. (1997). *Nat. Struct. Biol.* **4**, 98–101.
- Ray, A. & Nordén, B. (2000). *FASEB J.* **14**, 1041–1060.
- Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Sheldrick, G. M. (2010). *Acta Cryst.* **D66**, 479–485.
- Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
- Usón, I. & Sheldrick, G. M. (1999). *Curr. Opin. Struct. Biol.* **9**, 643–648.
- Zhang, J. (2004). PhD thesis. Georg-August-Universität Göttingen, Germany. *Synthesis of alanyl and homoalanyl peptide nucleic acids as a model system for base pairing and use as a catalyst*. Göttingen: Cuvillier Verlag.