Backbone-extended pyrrolidine peptide nucleic acids (*bep*PNA): design, synthesis and DNA/RNA binding studies†

T. Govindaraju and Vaijayanti A. Kumar*

Received (in Cambridge, UK) 3rd September 2004, Accepted 12th October 2004
First published as an Advance Article on the web 2nd December 2004
DOI: 10.1039/b413542c

One-carbon extended conformationally constrained pyrrolidine PNA monomer (*bep*PNA) has been synthesized, incorporated into PNA sequences at predefined positions, and showed selective RNA binding properties.

Among several nucleic acid analogues, Peptide Nucleic Acids (*aeg*PNA, Fig. 1) have emerged as important DNA/RNA binding molecules.[1] PNAs bind to DNA and RNA with high thermal stability and specificity by forming duplexes (WC base pair) and triplexes (WC base pair and HG hydrogen bonding).[1] The use of PNAs as gene-targeted drugs and as tools in molecular biology is a developing field.[2] However, weaknesses of PNAs such as limited aqueous solubility, poor cellular uptake and binding to cDNA/RNA in either parallel or antiparallel orientation[1b] need to be addressed before they can become viable drug candidates. Our approach[3] so far has been to create chiral, conformationally restricted PNA backbones that reduce the otherwise possible conformational states of the flexible open chain single strand PNAs to those appearing in the complex form. The pyrrolidine PNA (Fig. 1) is the outcome of one of the ways to lock the PNA conformation for this purpose. The mixed results surfacing from different groups[4] infer that the conformational constraint imposed onto the PNA backbone in the form of ring structure may not be uniform for various nucleobases or with sequence context. The conformational locking of the backbone probably needs to be reasonably moderated for fine tuning the internucleobase distance complementarity for uniform effective recognition *via* WC/HG hydrogen bonding. There are many examples in the literature that suggest that a five-atom amide leading to a seven-atom repeating backbone may be more useful because of the reduced conformational flexibility of the amide relative to the six-atom phosphodiester backbone.[5] This postulate has also been supported by X-ray studies.[5a] Some of these analogues with an extended seven-bond backbone cross-pair with RNA with high affinity compared to DNA.[5a,d] In contrast, Lowe *et al*. very effectively accomplished preferential DNA binding by replacing glycine with one-atom extended β-amino acids in their prolyl-glycyl PNA analogs.[6] We had earlier reported pyrrolidine based positively charged PNAs and found that the ring system in either 2*R*,4*R* or 2*S*,4*S* *cis* stereochemistry as a polyamide or chimeric amide-sugar phosphate backbone destabilized the complexes with either DNA or RNA.[4d,e] Our preliminary results on the chimeric phosphate–amide extended backbone revealed that 2*R*,4*R* pyrrolidine–amide chimera were accommodated better in triplex forming sequences.[4f] In view of the above discussion, studies of chiral pyrrolidine PNA with an extended backbone (Fig. 1) that may have potential for selective targetting to DNA or RNA was undertaken. The positive charge and chirality in the backbone could have additional advantages such as increased water solubility and orientational selectivity binding.

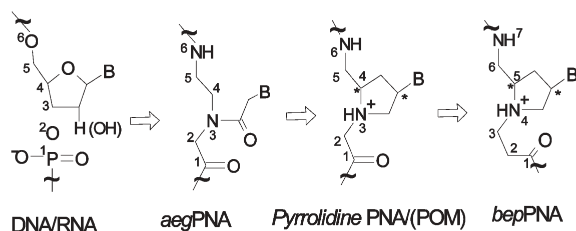
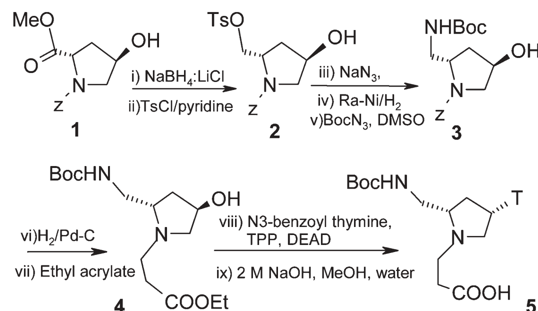


Fig. 1 Designed *bep*PNA as DNA/RNA and PNA mimic.

The synthesis of monomer **5** was accomplished as shown in Scheme 1. Protection of the ring nitrogen of the naturally occurring *trans*-4-hydroxy-L-proline and esterification with MeOH–$SOCl_2$ gave methyl ester **1**.[4] The methyl ester was reduced to the diol using LiCl/$NaBH_4$ and selective tosylation of the primary hydroxy group by controlled dropwise addition of *p*-TsCl in pyridine gave monotosylate **2** (ditosylate side product was removed by column chromatography), which was treated with excess NaN_3 to obtain the azido compound. The selective hydrogenation of the azide function using Raney-Ni in methanol gave the free amine and Boc protection of free amine yielded compound **3**. This was then subjected to hydrogenation over Pd–C to remove the *N*-benzyloxy carbonyl group and subsequent alkylation of the free ring nitrogen with ethyl acrylate gave *N*-alkylated product **4**. The nucleobase thymine was introduced at C-4 to get the protected monomer ethyl ester by treatment with



Scheme 1 Synthesis of protected *bep*PNA monomer.

† Electronic supplementary information (ESI) available: Experimental procedures for the synthesis of compounds, ^{1}H, ^{13}C NMR, mass spectral data, HPLC profiles, pH titration data and melting curves for triplexes. See http://www.rsc.org/suppdata/cc/b4/b413542c/

*vakumar@dalton.ncl.res.in

This journal is © The Royal Society of Chemistry 2005

*N*3-benzoylthymine under Mitsunobu conditions. The ester hydrolysis and deprotection of *N*3 of thymine using 2 M NaOH in methanol–water yielded the required monomer [(2*S*,4*S*)-2-(*tert*-butoxycarbonylaminomethyl)-4-(thymin-1-yl)pyrolidin-1-yl] propanoic acid **5** in good yield. All new compounds were characterized by NMR and mass spectrometry. The pK_a of the ring nitrogen was found to be 7.7 and partial protonation at physiological conditions may be expected.

The monomer **5** was incorporated into oligothymine PNA sequences at *C*-/*N*- terminals, in the center, alternative positions and through the entire sequence by solid phase peptide synthesis on L-lysine derivatized MBHA resin using Boc chemistry.[2b] The oligomers were cleaved from the solid support using the TFA–TFMSA method[2b] to yield *bep*PNAs **7**–**11**. For control studies *aeg*PNA T_8 octamer **6** was also synthesized. These were purified by RP HPLC on a semi-preparative C18 column and characterized by MALDI-TOF mass spectral analysis.[7]

The T_m values of homopyrimidine PNAs **6**–**11**, hybridized with complementary DNA and RNA were obtained from temperature dependent UV-absorbance data (Table 1). UV and CD Job plots[6,8] suggested the formation of 2:1 *bep*PNA_2:DNA and *bep*PNA_2:RNA triplexes and hence all the complementation studies were performed with 2:1 PNA:DNA/RNA stoichiometry. The *C*-terminal modified *bep*PNA **7** binds to DNA with slight decrease in T_m ($\Delta T_m = -1$ °C) whereas *bep*PNA **9** modified at the *N*-terminal stabilizes the complex ($\Delta T_m = +2$ °C) compared to control *aeg*PNA **6**. Surprisingly, *bep*PNA **8**, modified unit at the center did not show any complexation with DNA. Alternate and homooligomeric *bep*PNAs (**10** and **11**) did not form complexes with DNA as shown in UV-melting curves. A linear increase in absorbance was observed in these cases that corresponded to the PNA and DNA single strand melting (ESI†). UV-T_m data of PNA:DNA complexation is supported by the gel shift assay (Fig. 2). The *bep*PNA **8** exhibited very weak binding interaction at lower temperature, though it was not seen during UV-T_m thermal denaturation (Fig. 2, lane 6). When the complexation studies were performed with RNA, chimeric PNAs with a single *bep*PNA unit were found to bind with approximately the same T_m but slightly lower than that of control *aeg*PNA **6**. *bep*PNA **10** with alternating *aeg*-*bep* units exhibited a very high binding affinity (ΔT_m = + 4.5 °C/mod) (Table 1, Fig. 3). The observed transitions were very sharp with RNA compared to those with DNA. The sequence **11** comprised of only *bep*PNA backbone also recognized only RNA but with reduced strength compared to the alternating *aeg*-*bep* PNA.

Table 1 UV-T_m values in °C of PNA_2:DNA/RNA triplexes[a]

Entry	Sequence	DNA	RNA
1	*aeg*PNA **6**, H-TTTTTTTT-LysNH_2	51.5	65.8
2	*bep*PNA **7**, H-TTTTTTT**t**-LysNH_2	49.0	59.9
3	*bep*PNA **8**, H-TTT**t**TTTT-LysNH_2	[b]	59.2
4	*bep*PNA **9**, H-**t**TTTTTT-LysNH_2	53.0	59.0
5	*bep*PNA **10**, H-T**t**T**t**T**t**T**t**-LysNH_2	[b]	84.4
6	*bep*PNA **11**, H-**ttttttt**-LysNH_2	[b]	58.9

[a] T_m = melting temperature (measured in buffer: 10 mM sodium phosphate, pH 7.0 with 100 mM NaCl and 0.1 mM EDTA). Measured from 10 to 90 °C at ramp 0.2 °C min^{-1}. UV-absorbance measured at 260 nm. All values are an average of three independent experiments and accurate to within ±0.5 °C. DNA = $CGCA_8CGC$, RNA = polyrA. T = *aeg*PNA and **t** = *bep*PNA monomers. [b] Did not form complex with DNA.

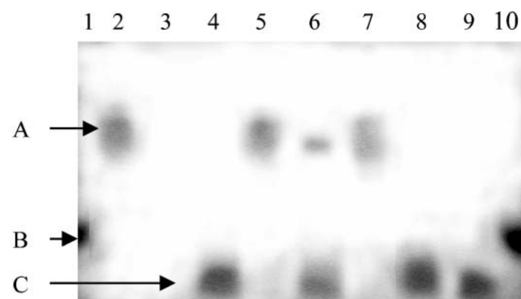


Fig. 2 PNA–DNA complexation. Lane 1, BPB; lane 2, (*aeg*PNA **6**)$_2$:DNA ; lane 3, ss*bep*PNA **7**; lane 4, ssDNA ; lane 5, (*bep*PNA **7**)$_2$ + DNA; lane 6, (*bep*PNA **8**)$_2$ + DNA; lane 7, (*bep*PNA **9**)$_2$:DNA; lane 8, (*bep*PNA **10**)$_2$ + DNA; lane 9, (*bep*PNA **11**)$_2$ + DNA; lane 10, BPB (bromophenol blue). DNA = CGCAAAAAAAACGC. **A**. PNA–DNA complex, **B**. BPB, **C**. Single stranded DNA.

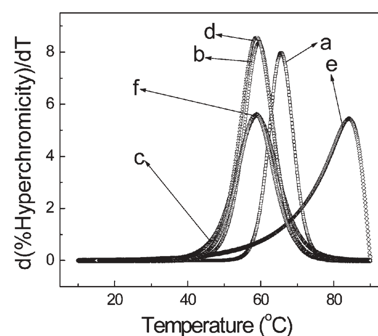


Fig. 3 First derivatives of melting curves of (a) *aeg*PNA **6** and chimeric *aeg*-*bep*PNA, (b) **7**, (c) **8**, (d) **9**, (e) **10** and (f) *bep*PNA **11** with RNA (poly rA).

These results suggest that *bep*PNA monomer in chimeric and homooligomeric PNAs induced binding selectivity for RNA over DNA. Incorporation of the modified units at the terminals (*C*-/*N*-) seems to exert only a very weak effective preorganized conformation and allowed binding with DNA as well as RNA. When in the centre of the sequence, the induced conformation allows recognition of RNA but that of DNA is suppressed. The high affinity binding of alternating *aeg*-*bep*PNA **10** with RNA suggests that the alternating *aeg*-*bep* units are uniformly spaced such that a balanced optimum conformation may be reached for recognition of RNA. The fully modified backbone in **11** binds to RNA but with reduced strength compared with the alternating sequence **10**. This could be because of overpreorganization of the single strand as suggested for fully modified LNA[9] or high positive charge concentration of two *bep*-homooligomers in 2:1 binding mode. The 2:1 binding stoichiometry for **11**:RNA was confirmed by a UV-Job plot (ESI†). The charge–charge repulsions could therefore be a possible reason for the observed reduced T_m. The inclusion of an extra atom in the backbone was easily achieved by using conjugate addition to ethyl acrylate as compared to cumbersome methods in the sugar–phosphate backbone.[5] The selective RNA recognition by the chiral, cationic PNA analogue is important for application perspectives. Further studies on the mixed

 This journal is © The Royal Society of Chemistry 2005

purine–pyrimidine sequences and other diastereomeric *bep*PNAs are in progress.

V. A. K. thanks the Department of Science and Technology, New Delhi for research grants. T. G. thanks CSIR, New Delhi for a fellowship. We thank Dr K. N. Ganesh for support and encouragement.

T. Govindaraju and Vaijayanti A. Kumar*
Division of Organic Chemistry (Synthesis), National Chemical Laboratory, Pune, 411008, India. E-mail: vakumar@dalton.ncl.res.in; Fax: 91 20 589 3153; Tel: 91 20 589 3153

Notes and references

1 (*a*) M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, **365**, 566–568; (*b*) K. N. Ganesh and P. E. Nielsen, *Curr. Org. Chem.*, 2000, **4**, 1931–1943.

2 (*a*) D. A. Braasch and D. R. Corey, *Biochemistry*, 2002, **41**, 4503–4510; (*b*) P. E. Nielsen and M. Egholm, in *Peptide nucleic acids (PNA), Protocols and Applications*, ed. P. E. Nielsen and M. Egholm, Horizon Scientific, Norfolk, CT, 1999.

3 V. A. Kumar, *Eur. J. Org. Chem.*, 2002, 2021–2032.

4 (*a*) D. T. Hickman, P. M. King, M. A. Cooper, J. M. Slater and J. Mickelfield, *Chem. Commun.*, 2000, 2251–2252; (*b*) T. H. S. Tan, D. T. Hickman, J. Morral, I. G. Beadham and J. Micklefield, *Chem. Commun.*, 2004, 516–517; (*c*) A. Puschl, T. Tedeschi and P. E. Nielsen, *Org. Lett.*, 2000, **2**, 4161–4163; (*d*) V. Kumar, P. S. Pallan, Meena and K. N. Ganesh, *Org. Lett.*, 2001, **3**, 1269–1272; (*e*) Meena and V. A. Kumar, *Nucleosides, Nucleotides Nucleic Acids*, 2003, **22**, 1285–1288; (*f*) V. A. Kumar and Meena, *Nucleosides, Nucleotides Nucleic Acids*, 2003, **22**, 1101–1104.

5 (*a*) C. J. Wilds, G. Minasov, F. Natt, P. Von Matt, K.-H. Altamann and M. Egli, *Nucleosides, Nucleotides Nucleic Acids*, 2001, **20**, 991–994; (*b*) G. V. Petersen and J. Wengel, *Tetrahedron*, 1995, **51**, 2145–2154; (*c*) A. Chur, B. Holst, O. Dahl, P. Valentin-Hansen and E. B. Pedersen, *Nucleic Acids Res.*, 1993, **21**, 5179–5183; (*d*) A. Lauritsen and J. Wengel, *Chem. Commun.*, 2002, 530–531.

6 T. Vilaivan and G. Lowe, *J. Am. Chem. Soc.*, 2002, **124**, 9326–9327.

7 LC-TOF-MS: *aeg*PNA **6**, 2275.90 [M + H^+]; *bep*PNA **7**, 2288.0 [M + $2H^+$]; *bep*PNA **8**, 2288.0 [M + $2H^+$]; *bep*PNA **9**, 2288.04, [M + $2H^+$]; *bep*PNA **10**, 2326.06 [M + $3H^+$]; *bep*PNA **11**, 2374.43 [M + $4H^+$].

8 P. Job, *Ann. Chim. (Paris)*, 1928, **9**, 113–203.

9 C. J. Leumann, *Bioorg. Med. Chem.*, 2002, **10**, 841–854.