Amino/guanidino-functionalized *N*-(pyrrolidin-2-ethyl)glycine-based pet-PNA: Design, synthesis and binding with DNA/RNA†

Sachin S. Gokhale and Vaijayanti A. Kumar*

Received 9th March 2010, Accepted 21st May 2010
First published as an Advance Article on the web 10th June 2010
DOI: 10.1039/c004005c

The *N*-(pyrrolidin-2-ethyl) glycine-based PNA (pet-PNA) backbone, with 4-amino or 4-guanidino-functionalized pyrrolidine ring, confers constrained conformational flexibility on aegPNA. The oligomers bind to the target DNA and RNA sequences with increased sequence specificity and antiparallel *versus* parallel orientation selectivity. The easy post-synthetic guanidination gives very good access to the positively charged PNA oligomers.

Introduction

Natural as well as unnatural positively charged oligopeptides such as Tat and other guanidine rich peptides exhibit very good cell penetrating properties.[1] The conjugation of such peptides as vehicles[2] to the aminoethylglycyl peptide nucleic acids (aegPNAs)[3] has widely opened the possibility of therapeutic applications of these highly competent nucleic acid mimics because of the cargo carrying capacity of the amino/guanidine substituted peptide vehicles. The optimum utility of PNAs is stymied by the low water solubility, ambiguity in directional selectivity as well as negligible capacity to cross cellular membrane.[4] Although the conjugation of positively charged peptides improves the cellular uptake of PNAs, the optimum utility for RNA targeting could still not be achieved. The conjugates remain entrapped in the endosomal cavities and release of PNAs for the antisense activity is not very efficient.[5] These charged peptides often are cytotoxic, show off-target effects and may not be useful for *in vivo* studies.[6] As an alternative to peptide conjugation, the PNAs could be designed to be intrinsically positively charged. We and others[7–9] have embarked to address this problem and have devised several interesting structural analogues of PNA. Marchelli and coworkers have studied the effect of chirality-induced handedness and charge in the backbone of aegPNA on binding with complementary DNA.[8] A recent promising example is that of GPNA which was developed by Ly *et al.*[9] The guanidino derived GPNA, obtained by substituting the glycyl segment in aegPNA by D-arginine, was shown to bind as efficiently to the target RNA as aegPNA and was also found to be equipped to cross the cell-membrane barriers. Similarly the γGPNA had a guanidino group inserted at γ-position in the aminoethyl segment of aegPNA, which was thought to be a helical director due to conformational pre-organization.[9d] It was also envisaged that the less amphipathic character of such intrinsically guanidinated PNA could be less toxic to cells.[10]

Introduction of chirality and cyclic structures in the PNA backbone are known to reduce the possible conformational states of acyclic, achiral single stranded PNAs.[7] The cyclic pyrrolidine/pentose based PNAs with extended backbone have been previously reported by us and others to bind sequence specifically only to RNA.[11] Our present design of extended backbone PNA converges Ly's guanidinylated γGPNA and the *N*-(pyrrolidinyl-2-methyl)glycine-based pmg-PNA (Fig. 1), earlier described by Slaitas and Yeheskiely.[12] It was shown that the pmg-PNA units, when incorporated in a mixed purine-pyrimidine decamer sequence, considerably destabilized the duplex formed, either with complementary RNA or DNA. Complexes of pmg-PNA were less destabilized with RNA as compared to those with DNA. We envisaged that the destabilization due to the constraint of the pyrrolidine ring in pmg-PNA observed in this work may be compensated by the additional flexibility of an extended backbone. Further, the amino/guanidino functionalized pyrrolidine ring would serve as constrained analogues of lysine[8] or arginine[9] in these PNAs. Thus, we describe in this paper the synthesis and properties of *N*-(pyrrolidin-2-ethyl)glycine based PNA, petPNA (Fig. 1) and its amino (Am-petPNA) and guanidino (Gu-petPNA) analogues.

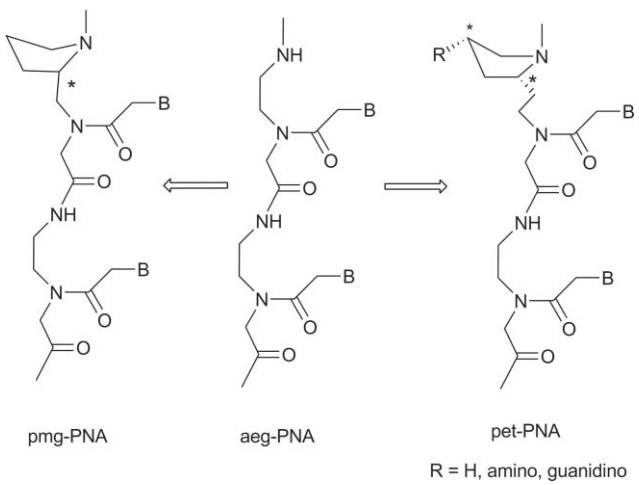


Fig. 1 pet-PNA as backbone extended pmg-PNA.

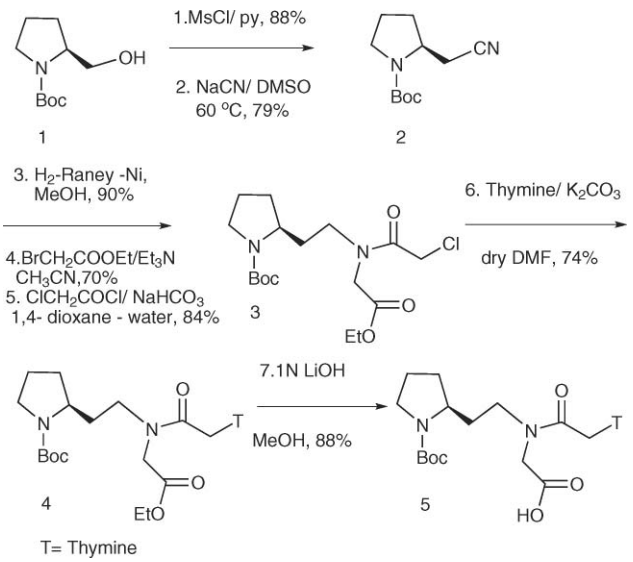
The post-synthetic conversion of amino to guanidino group simplifies the synthetic procedures as compared to the previously reported GPNA or γGPNA.[9] The amino/guanidino functionality

Division of Organic Chemistry, National Chemical Laboratory, Pune 411008, India. E-mail: va.kumar@ncl.res.in; Fax: +91 20 25602623

† Electronic supplementary information (ESI) available: ^{1}H, ^{13}C NMR and mass spectral data for all new compounds, HPLC and MALDI-TOF of PNA sequences and UV-T_m studies of all the complexes. See DOI: 10.1039/c004005c

is now in a one carbon extended, but conformationally constrained cyclic structure that may be an equivalent to the γ-position in the ethyl segment of aegPNA. The synthesis could be achieved from easily accessible naturally occurring amino acids, L-proline and 4-hydroxy-L-proline.

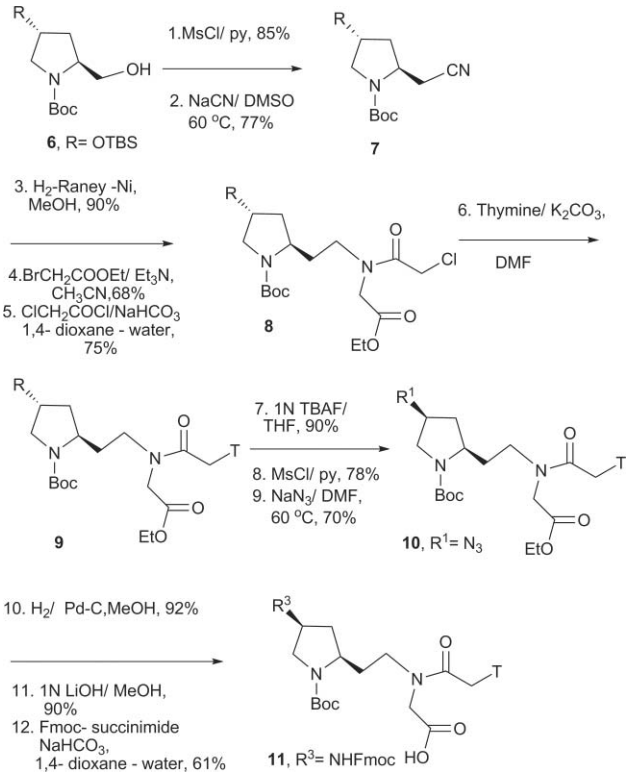
The synthesis of *N*-(pyrrolidin-2-ethyl)glycine based thyminyl and *N*-(4-Fmoc-amino-pyrrolidin-2-ethyl)glycine based thyminyl monomers is outlined in Scheme 1 and Scheme 2, starting from *N*-Boc-prolinol **1** and *N*-Boc-4-*O*-*tert*.butyldimethylsilyl-4-hydroxy prolinol **6**, respectively. The primary hydroxyl group in **1**/**6** was converted to its *O*- mesyl derivative followed by substitution with cyano nucleophile to give cyano derivative **2**/**7**. Hydrogenation of **2**/**7** over Ra–Ni gave 2-aminoethyl pyrrolidine derivatives.



Scheme 1 Synthesis of H-petPNA monomer

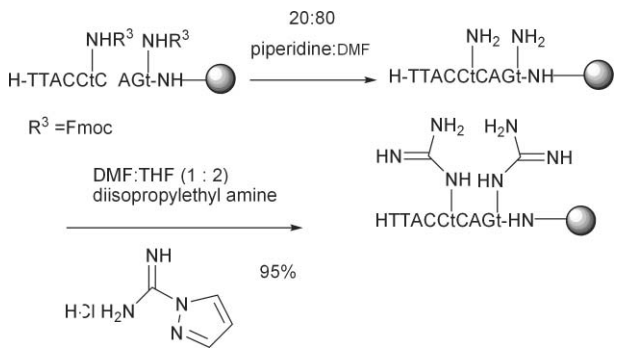
Alkylation of the 2-aminoethyl-pyrrolidine derivatives thus obtained, with ethyl bromoacetate, followed by acylation with chloroacetyl chloride gave the intermediate compounds **3**/**8**. The replacement of chloro group in **3**/**8** by thymine nucleophile in presence of K_2CO_3 gave *N*-1 alkylated products **4**/**9** in good yields. The intermediates **3**/**8** can also be used to synthesize other natural and unnatural nucleobase-containing analogues.[13] Compound **4** was hydrolyzed to get **5** as protected *N*-(pyrrolidin-2-ethyl)glycine (H-petPNA)-thymine monomer unit for solid phase PNA synthesis (Scheme 1).

Compound **9** was desilylated and converted to 4-azido derivative **10** *via* mesylation and displacement with azide nucleophile with inversion of configuration at C-4 of the pyrrolidine ring. Hydrogenation of azide over Pd–C, ester hydrolysis and Fmoc- protection of the 4-amino substituent gave protected *N*-(4-Fmoc-amino-pyrrolidinyl-2-ethyl)glycine, (Am-petPNA)-thymine monomer **11** in good yield (Scheme 2). The pyrrolidine ring nitrogen was protected with Boc whereas the exocyclic amine of the pyrrolidine ring was protected with an orthogonal Fmoc protecting group. Solid phase PNA synthesis was carried out using standard Boc chemistry on MBHA resin using the modified monomers **5** or **11** at predefined positions in the PNA sequences. The exocyclic amine in the modified monomer was deprotected at the end of solid phase synthesis. For the synthesis of Gu-petPNAs the free amino groups



Scheme 2 Synthesis of protected amino-petPNA monomer

was converted to free guanidino groups by global guanidination using 1-*H*-pyrazole-1-carboxamidine hydrochloride reagent on solid support.[14] (Scheme 3).



Scheme 3 Post-synthetic conversion of amino to guanidino functionality

Incorporation of the pyrrolidine unit with R = H, R = amino and R = guanidino, in aeg**PNA-1** in the center of the aeg**PNA-1** sequence gave three sequences **H-petPNA-1(s)**, **Am-petPNA-1(s)** and **Gu-petPNA-1(s)** respectively. Incorporation at the C-terminus of the sequence gave three sequences **H-petPNA-1(s*)**, **Am-petPNA-1(s*)** and **Gu-petPNA-1(s*)** respectively. A single pyrrolidine unit in mixed purine-pyrimidine sequence aeg-**PNA-2**, gave modified sequence **H-petPNA-2(s)**. Two units with R = H, R = amino and R = guanidino in aeg**PNA-2** gave three doubly modified sequences **H-petPNA-2(d)**, **Am-petPNA-2(d)** and **Gu-petPNA-2(d)** respectively. We also synthesized a mixed cytosin-thyminyl sequence **PNA-3**, and by incorporating four modified units at predefined positions in **PNA-3** obtained tetra-amino or guanidino substituted PNA oligomers **Am-petPNA-3(t)**

Table 1 PNA sequences and their MALDI-TOF mass analysis

PNA sequences	Mass Calc.	Mass Obsd.
1. H-TTTTTTTT-Lys-NH_2 **PNA-1**	2273.9	2274.9
H-TTTTt^RTTT-Lys-NH_2		
2. R=H, **H-petPNA-1(s)**	2328.98	2328.3
3. R=NH_2, **Am-petPNA-1(s)**	2343.99	2343.3
4. R = –NH–C(=NH)NH_2, **Gu-petPNA-1(s)**	2386.02	2382.3
H-TTTTTTTt^R-Lys-NH_2		
5. R=H, **H-petPNA-1(s*)**	2328.98	2329.2
6. R=NH_2, **Am-petPNA-1(s*)**	2343.99	2366.1 ($M+Na^+$)
7. R = –NH–C(=NH)NH_2, **Gu-petPNA-1(s*)**	2386.02	2386.57
8. H-TTACCTCAGT-Lys-NH_2 **PNA-2**	2805.2	2827.0 ($M+Na^+$)
H-TTACCt^RCAGT-Lys-NH_2		
9. R=H, **H-petPNA-2(s)**	2859.3	2859.4
H-TTACCt^RCAGt^R-Lys-NH_2		
10. R=H, **H-petPNA-2(d)**	2912.4	2910.4
11. R = NH_2, **Am-petPNA2(d)**	2942.4	2940.6
12. R = –NH–C(=NH)NH_2, **Gu-petPNA-2(d)**	3026.0	3024.5
13. H-CTTCTTCCTT-Lys-NH_2 **PNA-3**	2746.13	2741.15
H-Ct^RCTt^RCC$t^R$$t^R$-Lys-$NH_2$		
14. R=NH_2, **Am-petPNA-3(t)**	3026.4	3034.5
15. R = –NH–C(=NH)NH_2, **Gu-petPNA-3(t)**	3194.3	3201.4

t^R denotes petPNA units

and **Gu-petPNA-3(t)**. The PNA sequences were then cleaved from the resin by the standard TFA-TFMSA protocol.[15] The highly water-soluble oligomers were purified by reverse phase HPLC. The purity of oligomers was checked by HPLC analysis on RP-C18 column. The oligomers were characterized by MALDI-TOF mass spectrometry. The oligomer sequences thus obtained are listed in Table 1.

UV-melting studies

The synthesized PNA sequences were examined for their binding affinity with complementary DNA/RNA sequences by employing UV-thermal denaturation studies. Oligothyminyl PNA is known to form triplex structures with 2 : 1 PNA : DNA stoichiometry[3] that shows a single transition in temperature-dependent UV-T_m measurements. The results for thymine octamers sequences are summarised in Table 2.

The modified thyminyl sequences, **H-petPNA-1(s)**, **Am-petPNA-1(s)** and **Gu-petPNA-1(s)**, in which the modified unit is at the centre of the PNA sequence, formed triplexes with both complementary DNA and RNA. The destabilization was 2–5 °C except in the case of **H-petPNA-1(s)**: DNA (entry 2, column 2). The modified thyminyl sequences **H-petPNA-1(s*)**, **Am-petPNA-1(s*)** and **Gu-petPNA-1(s*)**, in which the modified unit is at the C-terminus of the PNA sequence, stabilized the triplexes formed with complementary DNA by 4–10 °C (entry 6–7, column 2), except in the case of **H-petPNA-1(s*)**: DNA (entry 5, column 2). In the case of RNA complement, presence of C-terminal modified units had a large stabilizing effect (entry 5–7, column 3). The mismatch discrimination was found to be more pronounced as compared to **PNA-1** when modified units **H-petPNA-1(s), Am-petPNA-1(s)** or **Gu-petPNA-1(s)** were present in the backbone in spite of amino/guanidino substitution for all the sequences. This could mean that the electrostatic interactions of the amino/guanidino groups with phosphate groups of DNA/RNA that might confer a stabilizing effect, particularly when present at the C-terminus do not overrule the WC or HG hydrogen bonding interactions in triplexes.

Table 2 UV-T_m (°C) studies with unmodified and modified thyminyl octamers

Sequence	**DNA-1(DNA- 3** with TT mismatch)[a]	**RNA-1 (RNA -3** with TU mismatch)[a]
1. **PNA-1**	41.6 (30.0)	41.2 (29.8)
2. **H-petPNA-1(s)**	27.2 (nd)	37.4 (16.5)
3. **Am-petPNA-1(s)**	40.7 (19.4)	36.8 (18.2)
4. **Gu-petPNA-1(s)**	36.0 (21.4)	36.4 (18.9)
5. **H-petPNA-1(s*)**	26.7 (nd)	48.7(20.9)
6. **Am-petPNA-1(s*)**	45.7 (26.8)	49.9(28.5)
7. **Gu-petPNA-1(s*)**	51.6 (32.6)	58.7(31.5)

DNA-1 : 5′GCAAAAAAAACG3′ RNA-1 : 5′ GCAAAAAAAACG 3′
DNA-3: 5′GCAAATAAAACG3′ RNA-3 5′ GCAAAUAAAACG 3′
[a] Values in the parentheses indicate melting temperatures of PNA: DNA/RNA complexes with a single mismatch

The mixed purine-pyrimidine (pu/py) sequences that form duplex structures with complementary DNA/RNA were then studied in comparison to unmodified **PNA-2**. The results are summarized in Table 3.

The **H-petPNA-2(s)** having a single pyrrolidine unit, slightly stabilized the complex with complementary antiparallel (*ap*) **DNA-2** but destabilized the complex with **RNA-2** by 2 °C. It is interesting to note that pmgPNA reported earlier[12] in contrast had a large destabilizing effect (–14 to –16 °C) on the stability of duplexes. The present pet-PNA, having one additional atom than the pmgPNA, was found to favourably compensate for the constrained ring structure. This improved stability of complexes of petPNA with DNA is in contrast to the recently reported RNA-selective backbone extended PNAs.[11] The doubly modified sequences **H-petPNA-2(d)** further stabilized the duplex with **DNA-2**, likewise destabilizing the duplex with **RNA-2**. The stabilizing effect of the second modified unit could result in significant conformational pre-organization, favourable for DNA binding, thus leading to the increased stability of the complexes with DNA. The doubly modified amino or guanidino substituted PNA sequences **Am-petPNA-2(d)** and **Gu-petPNA-2(d)** further stabilized the complexes with both complementary DNA as well as RNA by 4–6 °C as compared to **H-petPNA-2(d)**. Binding with complementary RNA sequence was as good as control PNA when two units of guanidino substituted pyrrolidine PNA monomers were present. In the earlier reported GPNA, as many as five modified units were needed for a similar result.[9a] The UV-T_m values in parentheses in Table 2 (column 2 and column 3) and Table 3 (column 2 and column 3) also suggest that the pyrrolidine modified PNAs with or without amino/guanidino substitution show as good as or better single mismatch discrimination with either TT(DNA-4) to TT(DNA-3)/UT(RNA-3)/CT(RNA-4) mismatch.

To see the effect of chirality of the petPNA monomer on the antiparallel(*ap*) *versus* parallel(*p*) orientation selectivity, binding studies were carried out with *p***DNA-5** and the results are summarized in Table 3, column 4. Figures in the parentheses in column 4 indicate the difference in T_m values (*ap*DNA-*p*DNA). It is clear that the introduction of petPNA units induced better binding in *ap* orientation and the difference in (*ap*-*p*) binding was maximum

Table 3 UV-T_m (°C) studies with unmodified and modified mixed pu/py oligomers

Sequences	*ap*DNA-2 (DNA 4 with TT mismatch)[a]	RNA-2 (RNA 4 with CT mismatch)[a]	*p*DNA-5 ($\Delta T_m = ap\text{-}p$)[b]
1. **PNA-2**	54.2 (47.5)	54.5 (46.5)	53.9(+0.3)
2. **H-petPNA-2(s)**	57.6 (51.2)	52.5 (38.8)	55.4 (+2.2)
3. **H-petPNA-2(d)**	63.2 (50.8)	48.1 (34.5)	59.3 (+3.9)
4. **Am-petPNA-2(d)**	66.9 (58.1)	52.1 (43.8)	59.5 (+7.4)
5. **Gu-petPNA -2(d)**	68.9 (60.0)	54.2 (46.1)	59.9 (+9.0)

DNA-2 ACTGAGGTAA 3′ DNA-4 ACTGTGGTAA 3′
DNA-5 AATGGAGTCA 3′ RNA-2 UGUAACUGAGGUAAGAGG 3′
RNA-4 UGUAACUGCGGUAAGAGG 3′
[a] Values in the parentheses indicate melting temperatures of PNA: DNA/RNA complexes with a single mismatch [b] Values in parentheses indicate differences in melting temperature between antiparallel and parallel binding modes.

Table 4 UV-T_m (°C) studies with unmodified and modified thymine-cytosine

Sequence	**DNA-6** (**DNA-7** with TT mismatch)[a]	**RNA-6** (**RNA-7** with TU mismatch)[a]
1. **PNA-3**	52.0(42.6)	55.6 (45.2)
2. **Am-petPNA-3(t)**	63.7(49.9)	65.5(57.6)
3. **Gu-petPNA-3(t)**	64.8(54.5)	68.3(59.2)

DNA6: AAGGAAGAAG RNA6: AAGGAAGAAG
DNA7: AAGGTAGAAG RNA7: AAGGUAGAAG
[a] values in the parentheses indicates melting temperatures of PNA:DNA/RNA complexes with a single mismatch

with **Gu-petPNA-2-(d)**. This could be because of the structural pre-organization of modified PNAs compared to the unmodified aegPNA sequences. CD data presented in Fig. 2 shows that doubly modified pet-PNA sequences show a prominent negative band at 270 nm indicating a base stacked conformation.[11e,16,17]

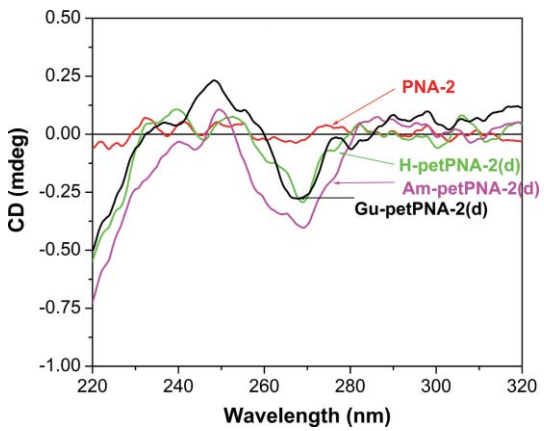


Fig. 2 CD spectra of unmodified aegPNA compared with doubly modified PNA sequences.

We further studied the effect of tetra-substituted amino **Am-petPNA-3-(t)** or guanidino **Gu-petPNA-3-(t)** substituted petPNA on the stability of complexes of mixed thymine-cytosine sequence **PNA-3**. The results are summarized in Table 4.

In this case, the complexes of petPNAs, **Am-petPNA-3(t)** and **Gu-petPNA-3(t)** with complementary **DNA-6** as well as **RNA-6** were highly stabilized (ΔT_m 11–13 °C) as compared to the unmodified **PNA-3**. The mismatch discrimination with single base mismatch in **DNA-7/RNA-7** was found to be comparable (8–14 °C) with unmodified **PNA-3** (10 °C), in spite of the presence of high population of amino or guanidino substitution on the modified PNA oligomers.

Improved binding with the target DNA as well as RNA sequences, improved mismatch discrimination and improved orientation binding selectivity by the petPNA sequences, lead us to believe that the structural features acquired by this PNA modification are in tune with the helical geometry of PNA: DNA/RNA complexes.

Conclusions

This paper presents the synthesis of the backbone extended petPNA and Am-petPNA monomers and their incorporation in aegPNA sequences. The conversion of amino to guanidino functionality was achieved on solid support and could be a method of choice for obtaining guanidinated PNA sequences. The amino and guanidino petPNA sequences were thus obtained in high yield and purity. The chirality and amino/guanidino functionality on PNA proved to be synergistic in improving the binding preferences of the modified PNA. The extended backbone in petPNA also compensated the constraint caused by the introduction of the cyclic pyrrolidyl unit. The work presented in this paper gives a new dimension to the pyrrolidine based PNAs reported so far.

Experimental

General

All the reagents were purchased from Sigma-Aldrich and used without purification. DMF, pyridine were dried over KOH and 4 Å molecular sieves. THF was passed over basic alumina and dried by distillation over sodium. Ethanol was dried over Mg/Iodine. TLCs were run on Merck 5554 silica 60 aluminium sheets. All reactions were monitored by TLC and usual workup implies sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulfate and evaporation under vacuum. Column chromatography was performed for purification of compounds on silica gel (100–200 mesh, LOBA Chemie). TLCs were performed using dichloromethane–methanol or petroleum ether–ethyl acetate solvent systems for most compounds. Compounds were visualized with UV light and/or by spraying with ninhydrin reagent subsequent to Boc-deprotection (exposing to HCl vapours) and heating. ^{1}H (200 MHz) and ^{13}C (50 MHz) NMR spectra were recorded on a Bruker ACF 200 spectrometer fitted with an Aspect 3000 computer and all the chemical shifts (δ (ppm))

are referred to internal TMS for ^{1}H and chloroform-*d*, DMSO-d_6 for ^{13}C NMR. ^{1}H NMR data are reported in the order of chemical shift, multiplicity (s, singlet; d, doublet; t, triplet; br, broad; br s, broad singlet; m, multiplet and/or multiple resonance), number of protons. Optical rotations were measured on a Bellingham-Stanley Ltd, ADP220 polarimeter. Mass spectra were recorded on a AP-QSTAR spectrometer, while MALDI-TOF spectra were obtained from a Voyager-De-STR (Applied Biosystems). UV experiments were performed on a Varian cary 300 UV-VIS spectrophotometer fitted with a peltier temperature programmer and a Julabo water circulator. The DNA oligomers were synthesized on CPG solid support using Applied Biosystems ABI 3900 High Throughput DNA Synthesizer by β-cyanoethyl phosphoramidite chemistry followed by ammonia treatment and their purities checked by HPLC prior to the use.

Synthesis of *N*-Boc-2-cyanomethyl pyrrolidine 2.

Step 1(Scheme 1): Synthesis of *N*-Boc-*O*-mesyl-prolinol.

Compound **1** (3.6 g, 18.32 mmol) was dissolved in 20 mL dry pyridine and mesyl chloride (1.7 mL, 21.98 mmol) was added at 0 °C. Reaction was stirred for 2 h at RT. After completion of the reaction, pyridine was removed under reduced pressure. Residue was dissolved in 100 mL ethyl acetate and organic layer was washed with 50 mL 10% $NaHCO_3$ solution. Organic layer was dried over anhydrous sodium sulfate and concentrated to give the crude product. It was used further for next reaction without purification. Crude yield = (4.5 g, 88%).

Step-2(Scheme 1): Synthesis of *N*-Boc-2*R*-cyanomethyl pyrrolidine **2**.

NaCN (6.3 g, 128.8 mmol) was added to the solution of the crude mesylate obtained as above (4.5 g, 16.1 mmol) in 20 mL dry DMSO, reaction was heated to 60 °C for 6 h. After completion of reaction DMSO was removed under reduced presssure and residue was dissolved in 100 mL ethyl acetate. Organic layer was washed with 50 mL water followed by a wash with 20 mL brine. Organic layer was concentrated under reduced presssure; crude compound was purified by column chromatography using gradient ethyl acetate–petroleum ether. Yield (2.8 g, 79%).

UV λ_{max}(nm)(CH_3CN) = 197(3.2), IR, ν/cm^{-1} ($CHCl_3$) 3128, 3019,2210,1724,1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 1.42 (s, 9H, BocCH_3), 1.77–2.14 (m, 4H, *H*3,3′, *H*4, 4′), 2.54–2.77 (m, 2H, CH_2CN), 3.35–3.38 (m, 2H, *H*5, 5′), 3.94–4.05 (m, 1H, *H*2).^{13}CNMR (50 MHz, $CDCl_3$) δ = 21.8, 23.3 (*C*3, *C*4), 28.0 (Boc (*C*H$_3$)$_3$), 30.9 (*C*H$_2$CN), 46.3 (*C*5), 53.4 (*C*2), 79.5 (Boc*C*(CH$_3$)$_3$), 117.6 (*C*N), 154.1 (Boc *C*O). *m*/*z*(ESI) 233.44 (M+ Na$^+$) ($C_{11}H_{18}N_2O_2$ requires 210.13).

Synthesis of Ethyl-[–*N*-(*N*-Boc-pyrrolidin-2*R*-ethyl)-*N*-chlorocetyl]-glycinate 3.

Step-3 (Scheme 1): Synthesis of *N*-Boc-2*R*-aminoethyl-pyrrolidine: The cyano derivative **2** obtained (2.5 g, 11.9 mmol) was dissolved in 20 mL methanol and was hydrogenated over RANEY®-Ni catalyst in presence of dry Et_3N (2.0 mL, 23.88 mmol) with 60 psi hydrogen pressure. After the completion of the reaction, methanol was removed under reduced presssure to give product 2.3 g (yield 90%) which was used further without purification.

Step 4 (Scheme 1): Ethyl-[–*N*-(*N*-Bocpyrrolidin-2*R*-ethyl)-glycinate: above amino compound (2.2 g, 10.47 mmol) was dissolved in 25 mL dry MeCN and triethyl amine (2.6 mL, 20.94 mmol) was added with stirring. The reaction mixture was cooled to 0 °C and ethyl bromoactate (0.81 mL, 12.56 mmol) diluted with 10 mL MeCN was added slowly. The reaction was stirred for 2 h at RT. After completion of the reaction, MeCN was removed under reduced presssure; and the residue was dissolved in 150 mL ethyl acetate. Organic layer was washed with 50 mL 10% $NaHCO_3$ solution, dried over anhydrous sodium sulfate and concentrated under reduced pressure to give ethyl-[–*N*-(*N*-Boc–2*R*-pyrrolidinethyl)-glycinate as crude product which was purified by column chromatography using gradient ethyl acetate–petroleum ether (Yield = 2.2 g, 70%).

UV λ_{max}(nm)(CH_3CN) = 211(3.4), IR, ν/cm^{-1} ($CHCl_3$) 3128, 3019, 1724, 1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 1.24–1.31 (t, 3H *J* = 7.2 Hz, COOCH$_2$C*H*$_3$), 1.46 (s, 9H, Boc(C*H*$_3$)$_3$), 1.52–1.94 (m, 6H, *H*3, 3′, *H*4, 4′, NHCH$_2$C*H*$_2$), 2.59–2.67 (m, 2H, NHC*H*$_2$CH$_2$), 3.32–3.4 (m, 4H, *H*5, *H*5′, NHC*H*$_2$COOEt), 3.85 (m, 1H, *H*2), 4.2–4.24 (m, 2H, *J* = 7.2 Hz, COOC*H*$_2$CH$_3$). ^{13}C NMR (50 MHz, $CDCl_3$) δ = 14.1 (COOCH$_2$*C*H$_3$), 28.4 (Boc(*C*H$_3$)$_3$), (30.4, 30.8), (*C*3, *C*4), 34.9 (NCH$_2$*C*H$_2$), 46.2 (N*C*H$_2$CH$_2$), 46.7 (NH*C*H$_2$COOC$_2$H$_5$), 50.8 (*C*5), 55.1 (*C*2), 60.6 (COO*C*H$_2$CH$_3$), 78.9 (BocC(*C*H$_3$)$_3$), 154.6 (Boc*C*O), 172.2 (*C*OOC$_2$H$_5$). Mass(ESI) 301.34 (M+H$^+$), 323.34 (M+Na$^+$) ($C_{15}H_{28}N_2O_4$ requires 300.20)

Step-5 (Scheme 1): Ethyl-[–*N*-(*N*-Boc-pyrrolidin–2*R*-ethyl)-*N*-chlorocetyl]-glycinate **3**.

Chloroacetyl chloride (4.3 mL, 34 mmol) was added in 2–3 portions to the solution of the compound obtained in the previous step (2.04 g, 6.8 mmol) and $NaHCO_3$ (9.9 g, 68 mmol) in 20 mL (1 : 1 water : 1,4-dioxane) at 0 °C. pH of reaction mixture was maintained at 8–9. After complete consumption of the starting material, reaction mixture was concentrated under reduced presssure. Product was extracted in 2 x 50 mL dichloromethane. The organic layer was dried over anhydrous sodium sulfate, concentrated and the product was purified by column chromatography using gradient ethyl acetate–petroleum ether to get pure product. Yield product (2.2 g, 84%).

UV λ_{max} (nm) (CH_3CN) = 205(3.7), IR, ν/cm^{-1} ($CHCl_3$) 3128, 3019, 1710, 1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 1.23–1.31 (m, 3H, COOCH$_2$C*H*$_3$), 1.46 (s, 9H, Boc (C*H*$_3$)$_3$), 1.76–1.99 (m, 6H, *H*3, *H*3′, *H*4, *H*4′, NCH$_2$C*H*$_2$), 3.33–3.41 (m, 4H, *H*2, *H*5, NC*H*$_2$CH$_2$), 3.79 (m, 1H, *H*5′), 4.01–4.25 (m, 6H, COOC*H*$_2$CH$_3$, NC*H*$_2$COOEt, COC*H*$_2$Cl)

^{13}C NMR (50 MHz, $CDCl_3$) δ = 14.0 (COOCH$_2$*C*H$_3$), 28.4 (Boc(*C*H$_3$)$_3$), 30.8 (*C*4), 33.7 (*C*3), 40.6 (NCH$_2$*C*H$_2$), 46.9(*C*5), 47.7 (N*C*H$_2$CH$_2$), 49.4 (NH*C*H$_2$COOC$_2$H$_5$), 54.9 (*C*2), 61.2 (CO*C*H$_2$Cl), 61.8 (CO*C*H$_2$CH$_3$), 79.3 (Boc*C*(CH$_3$)$_3$), 154.5 (BocCO), 166.6 (N*C*OCH$_2$Cl), 168.8 (*C*OOC$_2$H$_5$). Mass (ESI) 377.33 (M+H$^+$), 399.28(M+Na$^+$), 401.31 ($C_{17}H_{29}ClN_2O_5$ requires 376.17).

Synthesis of Ethyl-[–*N*-(*N*-Boc-pyrrolidin-2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate 4.

Step-6 (Scheme 1): The chloroacetyl derivative (2.05 g, 5.50 mmol), K_2CO_3 (0.97 g, 6.61 mmol) and thymine (0.68 g, 6.32 mmol) was

suspended in 5 mL dry DMF and stirred at RT for 6 h. DMF was removed under reduced presssure and compound extracted in 50 mL ethyl acetate. The crude product was purified by column chromatography using gradient dichloromethane–methanol to give the pure product **4**. (1.9 g yield = 74%).

UV λ_{max} (nm) (CH_3CN) = 206(3.6), 267(3.5), IR, ν/cm^{-1} ($CHCl_3$) 3128, 3019, 1720, 1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 1.23–1.32 (m, 3H, $COOCH_2CH_3$), 1.46 (s, 9H, Boc (CH_3)$_3$), 1.85–2.07 (m, 6H, *H*3, *H*3′, *H*4, *H*4′, NCH_2CH_2), 1.91 (s, thymine CH_3), 3.29–3.43 (m, 4H, *H*2, *H*5, NCH_2CH_2), 3.73–3.85 (m, 1H, *H*5′), 4.08–4.63 (m, 6H, $COOCH_2CH_3$, NCH_2COOEt, $COCH_2$ thymine), 7.03 (s, 1H *H*6 thymine). ^{13}C NMR (50 MHz, $CDCl_3$) δ = 12.1 (thymineCH_3), 13.9($COOCH_2CH_3$), 28.3 (Boc(CH_3)$_3$), 29.9 (*C*4), 30.6 (*C*3), 33.4 (NCH_2CH_2), 45.9 (*C*5), 47.5 (NCH_2CH_2), 49.0 ($NHCH_2COOC_2H_5$), 54.7 (*C*2), 61.2 ($COCH_2$thymine), 61.9 ($COCH_2CH_3$), 79.3 (Boc*C*(CH_3)$_3$), 110.4 (*C*5thymine), 140.9 (*C*6thymine), 151.0 (*C*2thymine), 154.4 ((BocCO), 164.4 (*C*4thymine), 166.9 ($NCOCH_2$thymine), 168.7 ($COOC_2H_5$). Mass (ESI) 489.60 (M+Na$^+$) ($C_{22}H_{34}N_4O_7$ requires 466.24)

[–*N*-(*N*-Boc-pyrrolidin-2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycine 5

Step-7 (Scheme 1): The compound **4** (1.6 g, 3.41 mmol) was dissolved in 2 mL methanol, to which 2 mL 1 M LiOH solution in water was added. Reaction was complete in 30 min. Methanol was removed, aqueous layer was neutralized with Dowex H$^+$ resin and resin was separated by filtration. The aqueous solution was washed with ethyl acetate, concentrated to give thymine monomer **5** (1.4 g yield = 88%).

UV λ_{max} (nm) (CH_3OH) = 201 (4.1), 268(3.8), IR, ν/cm^{-1} (CH_3OH) 3126, 3022, 1720, 1670.

^{1}H NMR (200 MHz, DMSO d_6) δ = 1.38 (s, 9H, Boc (CH_3)$_3$), 1.62–2.01 (m, 6H, *H*3, *H*3′, *H*4, *H*4′, NCH_2CH_2), 1.77(s, thymine CH_3), 3.23–3.35 (m, 4H, *H*2, *H*5, NCH_2CH_2), 3.62–3.86 (rotamers, m, 2H, *H*5′, *H*4), 3.96–4.06 (rotamers, s, 2H, NCH_2COOH), 4.48–4.63 (m, rotomers, $COCH_2$thymine) 7.31–7.43 (rotamers, d, 1H, *H*6 thymine).^{13}C NMR (50 MHz DMSO d_6) δ = 12.1 (thymineCH_3), 28.3 (Boc(CH_3)$_3$), 30.5 (*C*4), 32.7 (*C*3), 35.0 (NCH_2CH_2), 45.1(*C*5), 46.1 (NCH_2CH_2), 47.7 ($NHCH_2COOC_2H_5$), 54.8 (*C*2), 59.7 ($COCH_2$thymine), 78.6 (Boc*C*(CH_3)$_3$), 108.2 (*C*5thymine), 142.5 (*C*6thymine), 151.2 (*C*2thymine), 153.8 (BocCO), 164.7 (*C*4thymine), 167.3 ($NCOCH_2$thymine), 170.6 (*C*OOH). $[\alpha]_D^{20}$ = −34.07 (*c* = 1.2.MeOH).

Mass obsd. 445.08 (M+Li$^+$), 461.06 (M+Na$^+$) ($C_{20}H_{30}N_4O_7$ requires 438.22).

Synthesis of *N*-Boc-4-*R*-*O*-TBS-2*S*-cyanomethyl-pyrrolidine 7:

Step-1 (Scheme 2): *N*-Boc-4*R*-*O*-TBS-2*S*-*O*-mesyl-prolinol:

Compound **6** (8.6 g, 26.13 mmol) was dissolved in 30 mL dry pyridine and mesyl chloride (2.5 mL,31.36 mmol) was added at 0 °C. Reaction was stirred for 2 h at RT. After completion of the reaction, pyridine was removed under reduced presssure. Residue was dissolved in 100 mL ethyl acetate and organic layer was washed with 50 mL 10% $NaHCO_3$ solution. Organic layer was dried over anhydrous sodium sulfate and concentrated to give the crude product. It was used further for next reaction without purification. Crude yield: (8.7 g, 85%)

Step-2 (Scheme 2): Synthesis of *N*-Boc-4*R*-*O*-TBS-2*S*-cyanomethyl-pyrrolidine **7**:

NaCN (8.7 g, 177.8 mmol) was added to the solution of the crude mesylate obtained as above (8.5 g, 22.2 mmol) in 25 mL dry DMSO, reaction was heated to 60 °C for 6 h. After completion of reaction DMSO was removed under reduced presssure and residue was dissolved in 200 mL ethyl acetate. Organic layer was washed with 50 mL water followed by a wash with 20 mL brine. Organic layer was concentrated under reduced presssure; crude compound was purified by column chromatography using gradient ethyl acetate–petroleum ether. Yield (5.6 g, 77%).

UV λ_{max}(nm)(CH_3CN) = 194(3.6), IR, ν/cm^{-1} ($CHCl_3$) 3128, 3019,2208,1724,1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 0.07 (s, 6H, Si–(CH_3)$_2$), 0.86 (s, 9H, Si–C(CH_3)$_3$), 1.46 (s, 9H, Boc (CH_3)$_3$), 1.94–2.18 (m, 2H, *H*3,3′), 2.62–3.11 (m, 2H, CH_2CN), 3.42 (m, 2H, *H*5,5′), 4.12–4.39 (m, 2H, *H*4, *H*2).^{13}C NMR (50 MHz, $CDCl_3$) δ = −5.0 (Si(CH_3)$_2$), 17.7(Si*C*(CH_3)$_3$), 25.5 (SiC(CH_3)$_3$), 28.2 (BocCH_3), 39.8 (*C*3),40.8 (CH_2CN), 52.6 (*C*2), 55.6 (*C*5), 69.6 (*C*4), 79.9 (Boc*C*(CH_3)$_3$), 117.3 (*C*N), 154.5 (Boc*C*O). Mass (ESI) 363.35 (M+Na$^+$) ($C_{17}H_{32}N_2O_3Si$ requires 340.21).

Synthesis of Ethyl-[–*N*-(*N*-Boc-4*R*-*O*-TBS-pyrrolidin–2*R*-ethyl)-*N*-chlorocetyl]-glycinate 8.

Step-3 (Scheme 2): Synthesis of *N*-Boc-4*R*-*O*-TBS–2*R*-aminoethyl- pyrrolidine: The cyano derivative **7** (5 g, 14.74 mmol) was dissolved in 20 mL methanol and was hydrogenated over RANEY®-Ni catalyst in presence of dry Et_3N (4.09 mL, 47.18 mmol) with 60 psi hydrogen pressure. After the completion of the reaction, raction mixture was filtered and methanol was removed under reduced presssure to give product 4.5 g (yield 90%) which was used further without purification.

Step-4 (Scheme 2): Synthesis of ethyl –*N*-(*N*-Boc-4*R*-*O*-TBS-pyrrolidin–2*R*-ethyl)-glycinate: above compound (4.2 g, 12.28 mmol) was dissolved in 40 mL dry MeCN and Et_3N (3.14 mL, 24.56 mmol) was added with stirring. The reaction mixture was cooled to 0 °C and ethyl bromoactate (0.95 mL, 14.73 mmol) diluted with 10 mL MeCN added slowly. The reaction was stirred for 2 h. at RT. After completion of the reaction, CH_3CN was removed under reduced presssure and the residue was dissolved in 200 mL ethyl acetate. Organic layer was washed with 50 mL 10% $NaHCO_3$ solution, dried over anhydrous sodium sulfate and concentrated under reduced presssure to give ethyl –*N*-(*N*-Boc-4-*R*-*O*-TBS-pyrrolidin–2*R*-ethyl)-glycinate as crude product which was purified by column chromatography using gradient ethyl acetate–petroleum ether. 3.6 g (yield 68%).

UV λ_{max}(nm)(CH_3CN) = 203(3.8), IR, ν/cm^{-1} ($CHCl_3$) 3128, 3019,2110,1724,1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 0.05 (s, 6H, Si–(CH_3)$_2$), 0.86 (s, 9H, Si–C(CH_3)$_3$), 1.24–1.31 (m, 3H, $COOCH_2CH_3$), 1.46 (s, 9H, Boc (CH_3)$_3$), 1.56–2.01 (m, 4H, *H*3, 3′, NCH_2CH_2), 2.57–2.64 (m, 2H, NCH_2CH_2), 3.34–3.39 (m, 4H, *H*2, *H*5, NCH_2COOEt), 3.84–3.94(m, 1H, H5′), 4.13–4.24 (q, J, $COOCH_2CH_3$), 4.28–4.35 (m, 1H, *H*4). ^{13}C NMR (50 MHz, $CDCl_3$) δ = −4.9(Si-(CH_3)$_2$), 14.1($COOCH_2CH_3$), 17.8(Si*C*(CH_3)$_3$), 25.7(SiC(CH_3)$_3$), 28.3(BocCH_3), 29.5 (CH_2CH_2NH), 35.5 (*C*3), 40.3 ($HNCH_2CH_2$),

46.4 (*C*5), 50.8 (NH*C*H$_2$COOC$_2$H$_5$), 54.1 (*C*2), 60.6 (COO*C*H$_2$CH$_3$), 69.7(*H*4), 79.1(Boc*C*(CH$_3$)$_3$), 155.0(Boc*C*O), 172.2(*C*OOC$_2$H$_5$). Mass (ESI) 431.60 (M+H) ($C_{21}H_{42}N_2O_5Si$ requires 430.28)

Step-5 (Scheme 2): Synthesis of Ethyl-[–*N*-(*N*-Boc-4*R*-*O*-TBS-pyrrolidin–2*R*-ethyl)-*N*-chlorocetyl]-glycinate **8**: chloroacetyl chloride (5.2 mL, 40.65 mmol) was added in 2–3 portions to the solution of the compound obtained in the previous step (3.5 g, 8.13 mmol) and NaHCO$_3$ (6.8 g, 46.51 mmol) in 20 mL (1 : 1 water : 1,4-dioxane) at 0 °C, pH of reaction mixture was maintained at 8–9. After complete conversion of the starting material, reaction mixture was concentrated under reduced presssure. Product was extracted in 2 x 50 mL dichloromethane. The organic layer was dried over anhydrous sodium sulfate, concentrated and the product was purified by column chromatography using gradient ethyl acetate–petroleum ether to get pure product **8**. 3.0 g (yield 75%).

UV λ_{max}(nm)(CH$_3$CN) = 205(3.6), IR, ν/cm^{-1} (CHCl$_3$) 3128, 3019, 2110, 1724, 1670.

^{1}H NMR (200 MHz, CDCl$_3$) δ = 0.06 (s, 6H, Si–(CH$_3$)$_2$), 0.86 (s, 9H, Si–C(CH$_3$)$_3$), 1.24–1.27 (m, 3.0 H, COOCH$_2$C*H*$_3$), 1.48 (s, 9H, Boc (CH$_3$)$_3$), 1. 56–2.14 (m, 4H, *H*3, 3′, NCH$_2$C*H*$_2$), 3.25–3.46 (m, 4H, *H*2, *H*5, NC*H*$_2$CH$_2$), 4.00–4.33 (m, 7H, *H*4, OC*H*$_2$CH$_3$, NCOC*H*$_2$T, NC*H*$_2$COOEt).

^{13}C NMR (50 MHz, CDCl$_3$) δ = –4.9 (Si(*C*H$_3$)$_2$), 14.0 (COOCH$_2$*C*H$_3$), 17.8 (Si*C*(CH$_3$)$_3$), 25.5 (SiC(*C*H$_3$)$_3$), 28.3 (BocCH$_3$), 29.5 (*C*H$_2$CH$_2$N), 33.7 (*C*3), 41.1 (N*C*H$_2$CH$_2$), 46.5 (*C*5), 47.7 (N*C*H$_2$COOC$_2$H$_5$), 53.4 (*C*2), 54.9 (NCO*C*H$_2$Cl), 61.3 (COO*C*H$_2$CH$_3$), 70.0 (*C*4), 79.5 (Boc*C*(CH$_3$)$_3$), 155.1 (Boc*C*O), 155.3 (*C*OCH$_2$Cl), 166.6 (*C*OOC$_2$H$_5$). Mass (ESI) 507.59 (M+H$^+$), 509.58 ($C_{23}H_{43}ClN_2O_6Si$ requires 506.25)

Synthesis of Ethyl-[–*N*-(*N*-Boc-4*R*-*O*-TBS-pyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate 9.

Step-6 (Scheme 2): The chloroacetyl derivative **8** (3.0 g, 6.32mmol), K$_2$CO$_3$ (0.87 g, 5.92 mmol) and thymine (0.79 g, 6.32 mmol) were suspended in 5 mL dry DMF and stirred at RT for 6 h. After completion of reaction DMF was removed under reduced pressure and compound extracted in 50 mL ethyl acetate. The organic layer was dried over anhydrous sodium sulfate, concentrated to give crude product. The crude product was purified by column chromatography using gradient dichloromethane–methanol to give the pure product. 2.6 g (yield 72%).

UV λ_{max}(nm)(CH$_3$CN) = 267 (3.9) IR, ν/cm^{-1} (CHCl$_3$) 3128, 3019, 1724, 1670.

^{1}H NMR (200 MHz, CDCl$_3$) δ = 0.06 (s, 6H, Si–(CH$_3$)$_2$), 0.86 (s, 9H, Si–C(CH$_3$)$_3$), 1.26–1.45 (m, 3H, COOCH$_2$C*H*$_3$), 1.45 (s, 9H, Boc (CH$_3$)$_3$), 1.66–2.05 (m, 4H, NCH$_2$C*H*$_2$, *H*3, *H*3′), 1.92 (s, 3H, thymine-C*H*$_3$), 3.25–3.41 (m, 4H, *H*2, *H*5, NC*H*$_2$CH$_2$), 3.86–3.98 (m, 1H, *H*5′), 4.07–4.62 (m, 7H, *H*4, OC*H*$_2$CH$_3$, NCOC*H*$_2$T, NC*H*$_2$COOEt), 7.03 (s, 1H, *H*6 thymine).^{13}C NMR (50 MHz, CDCl$_3$) δ = –4.9 (Si(*C*H$_3$)$_2$), 12.2 (thymine*C*H$_3$), 14.0 (COOCH$_2$*C*H$_3$), 17.8 (Si*C*(CH$_3$)$_3$), 25.5 (SiC(*C*H$_3$)$_3$), 28.3 (BocCH$_3$), 29.5 (*C*H$_2$CH$_2$N), 33.9 (*C*3), 40.7 (N*C*H$_2$CH$_2$), 45.7 (*C*5), 47.5 (N*C*H$_2$COOC$_2$H$_5$), 53.6 (*C*2), 54.9 (NCO*C*H$_2$thymine), 61.3 (COO*C*H$_2$CH$_3$), 70.0 (*C*4), 79.6 (Boc*C*(CH$_3$)$_3$), 110.4 (*C*5thymine), 140.9 (*C*6thymine), 151.1 (*C*2thymine), 155.1 (Boc*C*O), 164.3 (*C*OCH$_2$thymine), 167.0 (*C*4thymine), 168.7 (*C*OOC$_2$H$_5$).

Mass (ESI) 597.68 (M+H), 619.71 (M+Na$^+$) ($C_{28}H_{48}N_4O_8Si$ requires 596.3241).

Synthesis of Ethyl-[–*N*-(*N*-Boc-4*S*-azidopyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate 10.

Step-7 (Scheme 2): Synthesis of ethyl-[–*N*-(*N*-Boc-4*R*-hydroxy-pyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate: Compound **9** (2.5 g, 3.35 mmol) was dissolved in 10 mL dry THF, 1M TBAF in THF (5.0 mmol) was added. After 30 min, THF was removed under reduced presssure, product extracted in 50 mL dichloromethane, organic layer washed with 20 mL water. After evaporation of the solvent, the crude product was purified by column chromatography using gradient methanol–dichloromethane to give the pure ethyl-[–*N*-(*N*-Boc-4*R*-hydroxy-pyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate yield (1.75 g, 90%)

Step-8 (Scheme 2): Synthesis of ethyl-[–*N*-(*N*-Boc-4*R*-*O*-mesyl-pyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate: The pure compound (1.7 g, 3.52 mmol) obtained in the previous step was dissolved in 20 mL dry pyridine and mesyl chloride (0.4 mL, 5.2 mmol) was added drop wise to this solution at 0 °C while stirring. After 2 h., pyridine was removed under reduced presssure; residue was dissolved in 50 mL ethyl acetate. Organic layer washed with 10% NaHCO$_3$ and dried over anhydrous sodium sulfate. The organic layer was concentrated to give, the crude product. This was purified by column chromatography using gradient methanol–dichloromethane to give ethyl-[–*N*-(*N*-Boc-4*R*-*O*-mesyl-pyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate 1.5 g (yield 78%)

Step-9 (Scheme 2): Synthesis of Ethyl-[–*N*-(*N*-Boc-4*S*-azido-pyrrolidin-2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate **10**: The mesylate obtained in the previous step (1.4 g, 2.5 mmol) and NaN$_3$ (1.65 g, 25.0 mmol) were suspended in 10 mL dry DMF and stirred at 60 °C for 6 h. After completion of the reaction, DMF was removed under reduced pressure and residue was dissolved in 100 mL ethyl acetate. The organic layer was washed with water, dried over anhydrous sodium sulfate, concentrated and the crude product was purified by column chromatography using gradient methanol–dichloromethane to give **10**. 0.88 g (yield 70%)

UV λ_{max}(nm)(CH$_3$CN) = 194(3.2), 266(3.8), IR, ν/cm^{-1} (CHCl$_3$) 3128, 3019,2110,1724,1670.

^{1}H NMR (200 MHz, CDCl$_3$) δ = 1.24–1.29 (m, 3H, COOCH$_2$C*H*$_3$), 1.46 (s, 9H, Boc (CH$_3$)$_3$), 1.92 (s, 3H, thymine CH$_3$), 1.84–2.35 (m, 4H, *H*3, 3′, NCH$_2$C*H*$_2$), 3.32–3.74 (m, 5H, *H*2, *H*5, 5′, NC*H*$_2$CH$_2$), 4.10–4.64 (m, 7H, *H*4, COOC*H*$_2$CH$_3$, NC*H*$_2$COOEt, COC*H*$_2$thymine), 7.03 (s, 1H, thymine *H*6).

^{13}C NMR (50 MHz, CDCl$_3$) δ = 12.3 (thymineCH$_3$), 14.0 (COOCH$_2$*C*H$_3$), 28.4 (BocC(*C*H$_3$)$_3$), 29.6 (NCH$_2$*C*H$_2$), 33.7 (*C*3), 36.0 (N*C*H$_2$CH$_2$), 46.0 (*C*5), 47.7 (N*C*H$_2$COOC$_2$H$_5$), 54.3 (*C*2), 59.4 (*C*4), 61.4 (COOCH$_2$CH$_3$), 80.4 (Boc*C*(CH$_3$)$_3$), 110.7 (*C*5thymine), 141.1 (*C*6thymine), 151.1 (*C*2thymine), 154.2 (BocCO), 164.3 (*C*4thymine), 167.2 (N*C*OCH$_2$thymine), 168.8 (*C*OOC$_2$H$_5$). Mass (ESI) 508.58 (M+H$^+$), 530.64 (M+ Na$^+$) ($C_{22}H_{33}N_7O_7$ requires 507.24)

Synthesis of [–*N*-(*N*-Boc-4*S*-Fmoc-aminopyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycine 11.

Step-10 (Scheme 2): Synthesis of ethyl [–*N*-(*N*-Boc-4*S*-aminopyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate

Compound **10** (0.8 g 1.6 mmol) and 10% Pd–C (80 mg) were suspended in 10 mL methanol and hydrogenation reaction was carried out at 40 psi hydrogen pressure. After completion of the reaction, the reaction mixture was filtered and solvent was evaporated to give amino derivative (0.69 g, 92%) which was used without further purification in the next step.

Step-11 (Scheme 2): Synthesis of [–*N*-(*N*-Boc-4*S*-aminopyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycine: The amino derivative obtained in the previous step (0.6 g, 1.24 mmol) was dissolved in 2 mL methanol and to which was added 2 mL 1 M LiOH solution in water. Reaction was complete in 30 min, methanol was removed and aqueous layer was neutralized with Dowex H^+ resin and resin was separated by filtration. The aqueous solution was washed with ethyl acetate, concentrated to give free amino acid (0.5 g, 90%).

Step-12 (Scheme 2): Synthesis of [–*N*-(*N*-Boc-4*S*-Fmoc-aminopyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycine **11**: The amino acid obtained in the previous step (0.5 g, 1.1 mmol) was dissolved in 3 mL 1 : 1 water-1,4-dioxan and $NaHCO_3$ (0.23 g, 2.75 mmol) and Fmoc –succinimide (0.44 g, 1.32 mmol) was added to the reaction mixture. The reaction was stirred for 6 h at RT. After completion of reaction 1,4-dioxane was removed under reduced presssure and aqueous layer was washed with ethyl acetate until impurity of Fmoc succinimide is removed. Then aqueous layer was neutralized with H^+ dowex resin. The compound was then extracted in 40 mL 5% methanol–ethyl acetate. Organic layer was dried over anhydrous sodium sulfate, concentrated and the product was purified by column chromatography using gradient methanol–dichloromethane to give thymine monomer **11** (0.46 g, 61%).

UV λ_{max}(nm)(CH_3OH) = 205(4.4), 265(4.03), 299(3.3)), IR, ν/cm^{-1} (CH_3OH) 1724,1670

^{1}H NMR (200 MHz, DMSO d_6) δ = 1.41 (s, 9H, Boc (CH_3)$_3$), 1.55–2.36 (m, 4H, *H*3,3′, NCH_2CH_2),1.76 (s, 3H, thymine CH_3), 2.69 –3.92 (8H, m, *H*2, *H*5,5′, NCH_2CH_2, Fmoc CH_2, *CH*), 3.16–3.75 (moisture from DMSO), 4.23–4.63 (m, 5H, *H*4, $COCH_2$thymine, $NHCH_2COOH$) 7.31–7.92 (m, 9H, Fmoc, Thymine *H*5).

^{13}C NMR (50 MHz, DMSO d_6) δ = 12.3 (thymineCH_3), 28.5 (Boc$C(CH_3)_3$), 29.4 (NCH_2CH_2), 31.6 (*C*3), 41.2 (NCH_2CH_2) 44.4 (*C*5), 47.1 (Fmoc*C*HCH_2OCON), 48.1 (NCH_2COOH), 50.8 (*C*2), 54.1 (*C*4), 65.8 (CH_2OFMOC), 79.0 (Boc$C(CH_3)_3$), 108.4 (*C*5 thymine), (FmocAr110.1, 120.5, 121.8, 125.6, 127.5, 128.0, 129.9, 137.9, 139.8, 141.2, 144.3) 142.7 (*C*6 thymine), 151.5 (*C*2thymime), 153.8 (BocCO), 156.2 (FmocCO), 164.2 (*C*4thymine), 167.5 ($NCOCH_2$), 174.4 (*C*OOH).

Mass (ESI) 698.68 (M+ Na^+), 720.69(M+2Na^+) ($C_{35}H_{41}N_5O_9$ requires 675.29) $[\alpha]_D^{20}$ = 16.7(c = 1.2, MeOH)

Oligomer Synthesis. The PNA oligomers were synthesized by standard Boc-solid phase peptide strategy.[3] Purification of all the PNA oligomers was carried out by Varian dual pump PROSTAR model No.210 HPLC on RP-C18 column with Water: CH_3CN-0.1% TFA system. PNA oligomers were characterized by MALDI-TOF mass spectrometry by using Voyager-DE-STR (Applied Biosystems) MALDI-TOF. A nitrogen laser (337 nm) was used for desorption and ionization. Spectra were acquired in linear mode. The matrix used for analysis was CHCA (α-cyano hydroxyl cinnamic acid).

Post synthetic conversion of Am-petPNA to Gu-petPNA. Am-petPNA synthesized MBHA resin (5mg) was suspended in dry DMF. Fmoc protected amino groups in Am-pet units in the Am-petPNA deprotected by 20% piperidine-DMF solution and resin was then subjected for guanidinilation by suspending in the solution of pyrazole-1-caboxamidine hydrochloride (10 equivalents) DIPEA in DMF–THF (1 : 2) mixture. After 6–7 h, the resin was washed with dry DMF and treated with the same guanidination mixture once again to ensure completion conversion from amino to guanidino functionality. The resin was washed with dry DMF and dry DCM and was then dessicated. Gu-petPNA was cleaved from this resin with standard conditions employing TFMSA-TFA and purified by reverse phase HPLC. (yield of conversion of amino to guanidino is 95%).

UV-*Tm* Measurements. The concentration was calculated on the basis of absorbance from molar extinction coefficients of the corresponding nucleobases of DNA/RNA/PNA.[18] The experiments were performed at 1 μM concentrations. The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (10 mM) and EDTA (0.1mM) and were annealed by keeping the samples at 90 °C for 5 min followed by slow cooling to room temperature. Absorbance *versus* temperature profiles were obtained by monitoring at 260 nm with Varian cary 300 spectrophotometer scanning from 10–85 °C at a ramp rate of 0.5 °C per minute. The data were processed using Microcal Origin 6.1 and T_m (°C) values were derived from the first derivative curves. All values are an average of at least 3 experiments and accurate to within ± 1.0 °C.

CD Analysis. Single strand PNA at 2 μmolar concentration were suspended in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (10 mM) and EDTA (0.1 mM). CD was scanned for PNA oligomers from 320 nm–220 nm wavelength range and the data were processed using Microcal Origin 6.1.

Acknowledgements

SSG thanks UGC, New Delhi for senior research fellowship. VAK thanks CSIR (NWP0036), New Delhi, for research grants.

Notes and references

1 (*a*) M. Zorko and U. Langel, *Adv. Drug Delivery Rev.*, 2005, **57**, 529; (*b*) E. L. Snyder and S. F. Dowdy, *Pharm. Res.*, 2004, **21**, 389.
2 (*a*) S. El-Andaloussi, T. Holm and U. Langel, *Curr. Pharm. Des.*, 2005, **11**, 3597; (*b*) B. Gupta, T. S. Levchenko and V. P. Torchilin, *Adv. Drug Delivery Rev.*, 2005, **57**, 637.
3 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497.
4 (*a*) U. Koppelhus, S. K. Awasthi, V. Zachar, H. U. Holst, P. Ebbesen and P. E. Nielsen, *Antisense Nucleic Acid Drug Dev.*, 2002, **12**, 51–63; (*b*) N. Bendifallah, F. W. Rasmussen, V. Zachar, P. Ebbesen, P. E. Nielsen and U. Koppelhus, *Bioconjugate Chem.*, 2006, **17**, 750; (*c*) S. Abes, J. J. Turner, G. D. Ivanova, D. Owen, D. Williams, A. Arzumanov, P. Clair, M. J. Gait and B. Lebleu, *Nucleic Acids Res.*, 2007, **35**, 4495–4501; (*d*) E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew. Chem., Int. Ed.*, 1998, **37**(20), 2796–2823.

5 (*a*) Y. Wolf, S. Pritz, S. Abes, M. Bienert, B. Lebleu and J. Oehlke, *Biochemistry*, 2006, **45**, 14944; (*b*) S. Abes, D. Williams, P. Prevot, A. Thierry, M. J. Gait and B. Lebleu, *J. Controlled Release*, 2006, **110**, 595; (*c*) J. J. Turner, M. Fabani, A. A. Arzumanov, G. Ivanova and M. J. Gait, *Biochim. Biophys. Acta, Biomembr.*, 2006, **1758**, 290.

6 S. W. Jones, R. Christison, K. Bundell, C. J. Voyce, S. M. V. Brockbank, P. Newham and M. A. Lindsay, *Br. J. Pharmacol.*, 2005, **145**, 1093.

7 (*a*) J. Micklefield, *Curr. Med. Chem.*, 2001, **8**, 1157; (*b*) J Kurreck, *Eur. J. Biochem.*, 2003, **270**, 1628; (*c*) V. A. Kumar and K. N. Ganesh, *Curr. Top. Med. Chem.*, 2007, **7**, 715; (*d*) V. A. Kumar, *E. J. Org. Chem.*, 2002, 2012–2021; (*e*) V. A. Kumar and K. N. Ganesh, *Acc. Chem. Res.*, 2005, **38**, 404–413.

8 (*a*) S. Sforza, G. Haaima, R. Marchelli and P. E. Nielsen, *Eur. J. Org. Chem.*, 1999, 197–204; (*b*) T. Tedeschi, S. Sforza, A. Dossena, R. Corradini and R. Marchelli, *Chirality*, 2005, **17**, S196–S204; (*c*) S. Sforza, T. Tedeschi, R. Corradini and R. Marchelli, *Eur. J. Org. Chem.*, 2007, 5879–5885.

9 (*a*) P. Zhou, M. Wang, L. Du, G. W. Fisher, A. Waggoner and D. H. Ly, *J. Am. Chem. Soc.*, 2003, **125**, 6878; (*b*) A. Dragulescu-Andrasi, P. Zhou, G. He and D. H. Ly, *Chem. Commun.*, 2005, 244–246; (*c*) P. Zhou, A. Dragulescu-Andrasi, B. Bhattacharya, H. O'Keefe, P. Vatta, J. Jens, Hyldig-Nielsen and D. H. Ly, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 4931–4935; (*d*) B. Sahu, V. Chenna, K. L. Lathrop, S. M. Thomas, G. Zon, K. J. Livak and D. H. Ly, *J. Org. Chem.*, 2009, **74**, 1509–1516.

10 M. Zasloff, *Nature*, 2002, **415**, 389–395.

11 (*a*) P. S. Pallan, P. von Matt, C. J. Wilds, K.-H. Altmann and M. Egli, *Biochemistry*, 2006, **45**, 8048–8057; (*b*) C. J. Wilds, G. Minasov, F. Natt, P. von Matt, K.-H. Altmann and M. Egli, *Nucleosides, Nucleotides Nucleic Acids*, 2001, **20**, 991–994; (*c*) T. Govindaraju and V. A. Kumar, *Chem. Commun.*, 2005, 495–497; (*d*) T. Govindaraju and V. A. Kumar, *Tetrahedron*, 2006, **62**, 2321–2330; (*e*) K. Gogoi and V. A. Kumar, *Chem. Commun.*, 2008, 706–708; (*f*) R. J. Worthington, N. M. Bell, R. Wong and J. Micklefield, *Org. Biomol. Chem.*, 2008, **6**, 92–103.

12 A. Slaitas and E. Yeheskiely, *Eur. J. Org. Chem.*, 2002, 2391.

13 (*a*) B. P. Gangamani, V. A. Kumar and K. N. Ganesh, *Tetrahedron*, 1996, **52**, 15017; (*b*) B. P. Gangamani, V. A. Kumar and K. N. Ganesh, *Chem. Commun.*, 1997, 1913–1914.

14 M. S. Bernatowicz, Y. L. Wu and G. R. Matsueda, *J. Org. Chem.*, 1992, **57**, 2497.

15 L. Christensen, R. Fitzpatrick, B. Gildea, K. H. Petersen, H. F. Hansen, T. Koch, M. Egholm, O. Buchardt, P. E. Nielsen, J. Coull and R. H. Berg, *J. Pept. Sci.*, 1995, **1**, 175.

16 S. Bagmare, M. D'Costa and V. A. Kumar, *Chem. Commun.*, 2009, 6646–6648.

17 C. Glemarec, A. Nyilas, C. Sund and J. Chattopadhaya, *J. Biochem. Biophys. Methods*, 1990, **21**, 311–332.

18 P. Wittung, P. Nielsen and B. Norden, *Biochemistry*, 1997, **36**, 7973–7979.

