PNA forms an i-motif†

Yamuna Krishnan-Ghosh,[b] Elaine Stephens[a] and Shankar Balasubramanian*[a]

Received (in Cambridge, UK) 22nd July 2005, Accepted 23rd August 2005
First published as an Advance Article on the web 23rd September 2005
DOI: 10.1039/b510405j

A C-rich PNA hexanucleotide, p(C_5T), has been shown to form an i-motif by nanoelectrospray ionization mass spectrometry coupled with H/D exchange, to have thermal stability comparable with its DNA analogue, but to exist over a much narrower pH range.

At moderately acidic pH, cytosine-rich DNA and RNA strands can associate both inter- and intramolecularly to form an i-motif.[1] The i-motif consists of two parallel duplexes, comprising hemiprotonated C–C^+ base pairs,[2] that are intercalated in an antiparallel orientation.[3] The i-motif has potential as a nanoscale structural element for the construction of functional assemblies.[4] We reasoned that an unnatural i-motif analog that lacks the negatively charged backbones may have properties that are distinct from a DNA or RNA i-motif. Towards this end we chose to investigate C-rich peptide nucleic acids (PNAs). These are synthetic analogues of DNA where the sugar-phosphate backbone of DNA has been replaced by a neutral polyamide backbone.[5] Moreover, the PNA backbone has proven to be a good structural mimic of the sugar-phosphate backbone forming duplexes,[6] triplexes[7] and G-quadruplexes.[8] Although chiral, modified alanyl PNA forms i-motifs, i-motif formation has not been demonstrated by the more widely used achiral, glycyl aminoethyl PNA backbone.[9]

We investigated the self-assembly of the PNA sequence p(C_5T) (See Chart 1) comprising five tandem cytosines (C). A lysine was appended to the C-terminus in order to enhance solubility in water. p(C_5T) was synthesized using solid-phase synthesis using Fmoc chemistry (see electronic supplementary information (ESI)†).

To form a PNA complex, a solution of 2 mM p(C_5T) in 50 mM NaOAc buffer, pH 4.5, was heated at 90 °C for 10 min, then cooled at a rate of 0.5 °C min^{-1} to 5 °C and incubated at 5 °C for 24 h. An aliquot of the 2 mM solution was diluted 20 fold with Milli-Q water and subjected to nano-electrospray ionisation-mass spectrometry (Nano-ESI-MS) to study the molecularity of the PNA species and seek evidence for i-motif formation. ESI-MS has been used to observe non-covalent complexes of PNA[8a] and PNA–DNA hybrids.[10] Analysis at a cone voltage of 60 V and source temperature of 30 °C showed peaks at *m*/*z* 1336.18, 1252.5 and 1113.3 corresponding to M_4^{5+}, M_3^{4+} and M_2^{3+} (Table 1) respectively (ESI).†

The symmetrically charged tetramer M_4^{6+} and M_2^{3+} have indistinguishable *m*/*z* values.[11] The i-motif consists of two identical, dimeric subunits that are intercalated. The unsymmetrical M_4^{5+} peak is comprised of two non-identical intercalated subunits, one of which corresponds to the dimer, M_2^{3+}, as evidenced by MS-MS on M_4^{5+} (ESI†). Therefore for clarity, we have chosen to present the analysis on M_2^{3+} rather than M_4^{5+} to elucidate hydrogen bonding in the component subunit.[12] Chart 1, **1** shows the putative

[a]*University Chemical Laboratories, Department of Chemistry, Lensfield Road, University of Cambridge, Cambridge, UK CB2 1EW. E-mail: sb10031@cam.ac.uk; Fax: +44 1223 336913; Tel: +44 1223 336347*
[b]*National Centre for Biological Sciences, TIFR, UAS-GKVK Campus, Bellary Road, Bangalore, 560 065, India. E-mail: yamuna@ncbs.res.in; Fax: +91 80 23636675; Tel: +91 80 23636421*

† Electronic supplementary information (ESI) available: Synthesis, mass spectra, CD spectra. See http://dx.doi.org/10.1039/b510405j

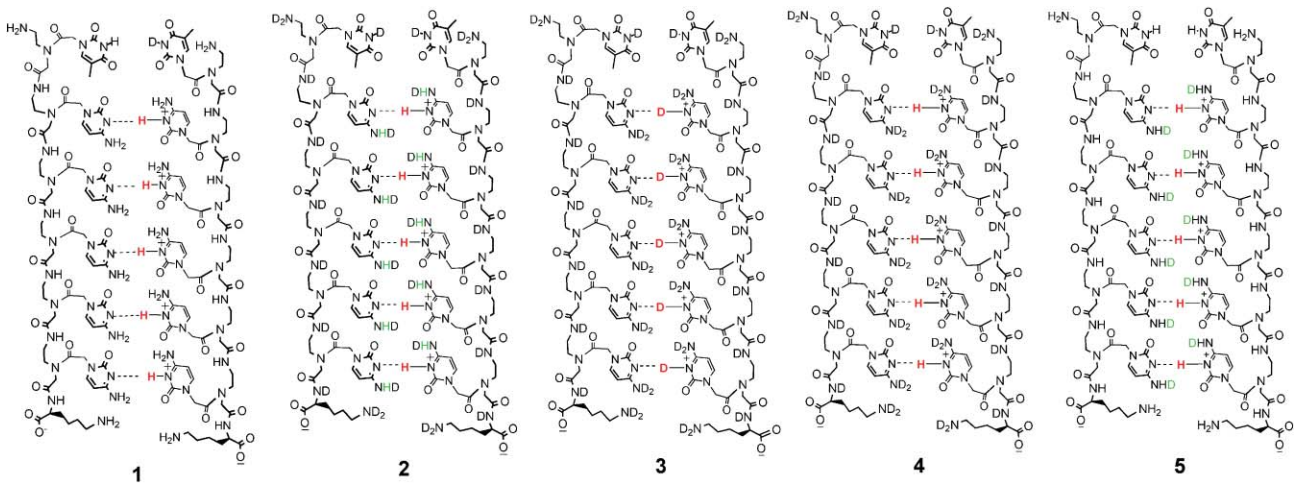


Chart 1 Proposed structure of M_2^{3+} formed by p(C_5T) where all the exchangeable sites are indicated by H. Buried protons involved in charged H-bonds are shown in red; Exchange-inert H/D sites involved in C–C^+ base pairs *via* neutral H-bonds are shown in green. (**1**): Undeuterated M_2^{3+}; (**2**): H/D exchange on (**1**) at 25 °C; (**3**): Fully deuterated M_2^{3+}; (**4**): H/D exchange upon heating (**2**); (**5**): D/H exchange on (**4**)

Table 1 Observed *m*/*z* and associated molecular weight (M_w) of p(C_5T) as determined by Nano-ESI-MS[a] before and after H/D exchange

Trial	Observed *m*/*z* (M_2^{3+}, M_3^{4+}, M_4^{5+})	Expected[b] *m*/*z* of M_2^{3+}, M_4^{5+}	M_w/Da of M_2^{3+} Calcd (Expected)[b]	Sites exchanged in M_2^{3+}
1	1113.3, 1252.5 , 1336.18	1113.44, 1335.93	3339.9 ± 0.53 (3340.33)	0
2	1124.07, 1264.32, 1349.6	1124.17, 1349.20	3372.36 ± 0.78 (3372.53)	32
3[c]	1129.13, 1269.58, 1354.35	1129.21, 1354.64	3387.21 ± 0.68 (3387.62)	47[c]
4	1127.53, 1268.08, 1352.72	1127.53, 1352.83	3383.42 ± 0.86 (3382.59)	42
5	1116.54, 1255.8, 1339.43	1116.80, 1339.55	3349.40 ± 0.76 (3350.39)	32[d]

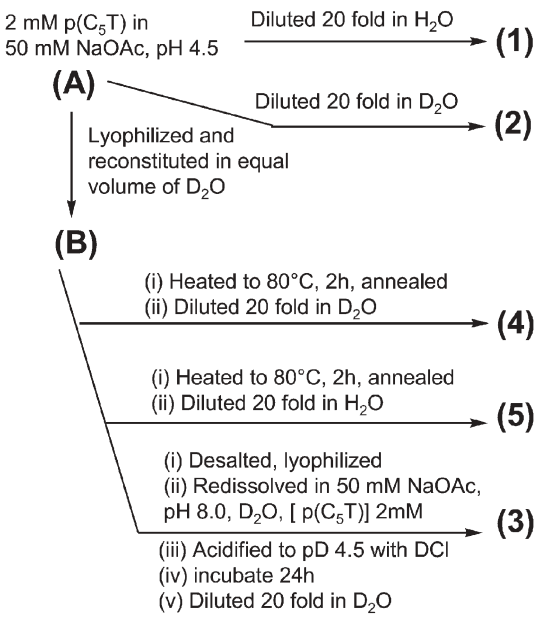
[a] Nano-ESI-MS used a cone voltage of 45 eV, source temperature 30 °C, analyzer pressure 10^{-5} bar; [b] MW calculated assuming the number of exchanges listed in right-most column ; [c] Obtained by incubation at pH 8.0 prior to complexation in D_2O; [d] Number of sites that underwent back exchange (D/H).

structure of the triply charged dimer, M_2^{3+}. It incorporates five buried protons while the C-termini in both strands are present as carboxylates giving a net charge of +3.[13]

The observation of non-covalent complexes by ESI-MS can sometimes be the result of a non-specific interaction, due to the nature of the electrospray process.[14] In order to address whether the complex of interest was formed in solution, we combined solution phase H/D exchange with ESI-MS.

H/D exchange coupled with ESI-MS can provide insight into the structure of the complex and the associated non-covalent interactions.[10a,15] A 2 mM solution of p(C_5T) tetramer in buffer, was lyophilised and re-suspended in an equal amount of D_2O to deuterate exchangeable protons as indicated in Scheme 1. An aliquot of this sample was diluted 20 fold in D_2O to give sample **2**. Nano-ESI-MS of **2** showed peaks shifted to *m*/*z* 1349.75 and 1124.74 corresponding to M_4^{5+} and M_2^{3+} respectively (Fig. 1). The computed average mw of M_2^{3+} is 3372.36 ± 0.8 Da, revealing that on average, H/D exchange occurs at 32 out of a possible 47 sites in the dimer (Chart 1, **2**). This suggests the remaining 15 sites per dimer subunit are present in a protected environment and are exchange inert, possibly due to H-bonding.

In order to obtain the fully deuterated dimer **3**, the partially exchanged complex (B), (2 mM p(C_5T)) was heated to 80 °C for 2 h. An aliquot of (B) was diluted 20 fold with D_2O to give sample **4**, which was subjected to Nano-ESI-MS (Fig. 1). Fully deuterated p(C_5T) (**3**), should show an *m*/*z* of 1129.21. However, we did not obtain **3** by merely heating (B) to 80 °C. Sample **4** instead showed an *m*/*z* 1127.54, indicating that five protons in M_2^{3+} resisted exchange even when heated (Table 1). This is entirely consistent with structure shown in Chart 1(**4**), where the H-bonded amine protons on the cytosine residues have exchanged, but the 5 central, charged hydrogen bonds have not. To confirm this model, back-exchange experiments were carried out on **4**. 2 mM p(C_5T) in D_2O was lyophilized and resuspended in H_2O. An aliquot was diluted 20 fold in H_2O to give sample **5**. Back-exchanged sample **5**, showed an *m*/*z* of 1116.53 (calc. *m*/*z* 1116.527) (Fig. 1). This indicates D/H exchange of (a) the backbone amide deuterons (b) amide deuterons on the thymine N3 sites and (a) one amino deuteron per cytosine residue (See Chart 1(**5**)). This consistent pattern observed in the forward and reverse exchange experiments reaffirms that p(C_5T) forms a specific tetrameric complex in solution. Heating to 80 °C was not sufficient to disrupt the complex formed by p(C_5T) at a strand concentration of 2 mM, consistent with the melting experiments (Fig. 2A). It was therefore possible to observe further differentiation by H/D exchange even



Scheme 1

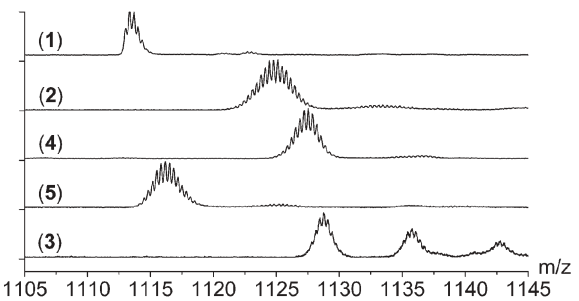


Fig. 1 Partial Nano-ESI-MS spectrum of p(C_5T), showing M_2^{3+}, under various conditions of H/D exchange, for samples **1–5**.

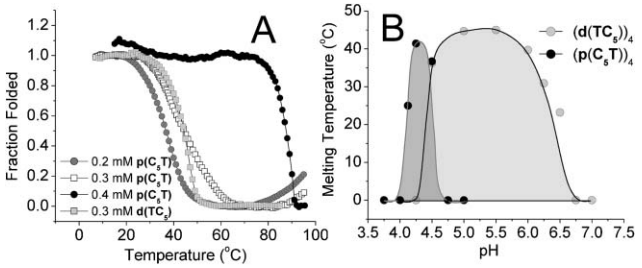


Fig. 2 (A) UV-melting profiles at 295 nm of $(p(C_5T))_4$ at various concentrations in 30 mM Acetate buffer, pH 4.5.(B) Plot of the melting temperatures of $(p(C_5T))_4$ and $(d(TC_5))_4$ at 300 μM as a function of pH.

among those H-bonded sites reluctant to exchange. The five protons on the cytosine N3 sites involved in the C–C^+ base pairs are more shielded than the 10 H-bonded amino protons of the C–C^+ base pairs. Furthermore, N–H $^+$N bonds are much stronger ($\sim$13 kcal mol^{-1})[16a] than amine N–H O=C hydrogen bonds ($\sim$1.4–2.4 kcal mol^{-1}).[16b] These five centrally located protons could be replaced by incubating the sample at pD 8.0 to disrupt the complex and facilitate H/D exchange, consistent with the pH dependence study (Fig. 2B), followed by reforming the i-motif of fully exchanged, monomeric p(C_5T) at pD 4.5 to give fully deuterated p(C_5T)$_4$, shown in Chart 1(**3**).

UV-melting experiments were then performed on complexed p(C_5T). DNA duplexes and triplexes show a positively sloped sigmoidal curve when the UV absorbance at 260 nm (A_{260}) is plotted against temperature (*T*). The corresponding i-motif melting transition is characterised by a distinctive *inverse* sigmoidal melting curve observed at 295 nm.[17] Complexed p(C_5T) at 2mM strand concentration did not evidence a melting profile. However, when aliquots of this solution were diluted to strand concentrations of 0.4, 0.3 and 0.2 mM, melting profiles characteristic of i-motifs were obtained. The UV-melting profile of 300 μM p(C_5T) at 295 nm in 30 mM NaOAc buffer pH 4.5, showed an inverse sigmoidal curve with a $T_{1/2}$ of 40 $\pm$ 1 °C. The melting temperature decreased with decreasing strand concentration confirming the intermolecular nature of the complex (Fig. 2A). Because C–C^+ base pairs require hemi-protonation of the cytosine nucleobases, i-motif stability is pH dependent. Melting studies on p(C_5T) and its DNA analog, d(TC_5), were conducted and the corresponding melting temperatures were plotted as a function of pH (Fig. 2B). At strand concentrations of 300 μM, both the DNA and PNA tetraplexes exhibited bell-shaped curves. In our hands, d(TC_5) forms an i-motif over a pH range 4.5–6.5 showing a maximum stability at pH 4.7–5.7 ($T_{1/2}$ = 44.7 °C).[18] However, p(C_5T) forms a tetraplex over a much narrower pH range 4.1–4.5 with maximum stability at pH 4.3 ($T_{1/2}$ = 41.4 °C). Given that N3 on cytidine monophosphate has a pK_a of 4.4–4.5,[19] this reaffirms that p(C_5T) forms a tetraplex that is held together by hemiprotonated C–C^+ base pairs. We attribute the difference in response to pH to the polyanionic nature of the DNA backbone which probably shields pH changes more effectively than the neutral PNA backbone. CD studies (ESI†) suggested an antiparallel strand arrangement consistent with DNA based i-motifs.[3]

The PNA_4 i-motif offers the gross structural features of the DNA_4 i-motif, but without the negatively charged backbone. We believe that this motif has potential as a structural element in the construction of nucleic acid based nanoscale assemblies. We thank Prof. Carol V. Robinson for valuable comments, the BBSRC, the Royal Commission for the Exhibition of 1851 and the IRC in Nanotechnology for funding.‡

Notes and references

‡ Note added in proof: After submission of this paper we became aware of a related and complementary study using PNA TC_8 (ref. 20).

1 E. O. Akrinimiski, C. Sander and P. O. P. Ts'O, *Biochemistry*, 1963, **2**, 340–344; K. A. Hartman, Jr and A. Rich, *J. Am. Chem. Soc.*, 1965, **87**, 2033–2039; B. Inman, *J. Mol. Biol.*, 1964, **9**, 624.

2 R. E. Marsh, R. Bierstedt and E. L. Eichhorn, *Acta Crystallogr.*, 1962, **15**, 310–316.

3 L. Cai, S. Raghavan, R. Ratliff, R. Mayzis and A. Rich, *Nucleic Acids Res.*, 1998, **26**, 4696–4705; K. Gehring, J.-L. Leroy and M. Guéron, *Nature*, 1993, **363**, 561.

4 D. Liu and S. Balasubramanian, *Angew. Chem., Int. Ed.*, 2003, **42**, 5734–5736.

5 P. E. Nielsen, M. Egholm, R. M. Berg and O. Buchardt, *Science*, 1991, **254**, 1497.

6 (*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*) P. Wittung, P. E. Nielsen, O. Buchardt, M. Egholm and B. Norden, *Nature*, 1994, **368**, 561–563.

7 (*a*) E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew. Chem., Int. Ed.*, 1998, **37**, 2796–2823; (*b*) H. Knudsen and P. E. Nielsen, *Nucleic Acids Res.*, 1996, **24**, 494–500; (*c*) P. Wittung, P. E. Nielsen and B. Norden, *Biochemistry*, 1997, **36**, 7973–7979; (*d*) P. Wittung, P. E. Nielsen and B. Norden, *J. Am. Chem. Soc.*, 1997, **119**, 3189.

8 (*a*) Y. Krishnan-Ghosh, E. Stephens and S. Balasubramanian, *J. Am. Chem. Soc.*, 2004, **126**, 5944–5945; (*b*) B. Datta, C. Schmitt and B. A. Armitage, *J. Am. Chem. Soc.*, 2003, **125**, 4111–4118.

9 U. Diederichsen, *Angew. Chem., Int. Ed.*, 1998, **37**, 2273.

10 (*a*) S. Sforza, T. Tedeschi, R. Corradini, A. Dossena and R. Marchelli, *Chem. Commun.*, 2003, 1102–1103; (*b*) M. C. Griffith, L. M. Risen, M. J. Greig, E. A. Lesnik, K. G. Sprankle, R. H. Griffey, J. S. Kiely and S. M. Freier, *J. Am. Chem. Soc.*, 1995, **117**, 831.

11 We cannot rule out the contribution of the duplex to peaks in the M_2^{3+} regime. However, the observation of a large peak centred at *m*/*z* 1135.77, Fig. 1(**3**), for the fully deuterated species ($\Delta m/z$ = 6.64) indicates the existence of a substantial population of a +6 charged, mono-potassiated species, namely $[M_4 - H^+ + K^+]^{6+}$, implying that the non-potassiated tetramer is a major component of the M_2^{3+} peak.

12 Analysis of M_4^{5+} peaks also confirmed i-motif formation (see ESI†). The non-identical subunits, M_2^{3+} and M_2^{2+}, differed only in the number of C–C^+ base pairs.

13 Other combinations of buried charges and carboxylates possible for M_2^{3+} were ruled out *via* H/D exchange (ESI†). The data suggests that amines at the termini of each PNA strand are neutral in the observed species. A possible explanation could be depression in the pK_a values of the corresponding protonated amines caused by the multiple positive charges of the C–C^+ base pairs. We have previously observed reduced levels of protonation of terminal amines in PNA G-quadruplexes by mass spectrometry (ref. 8*a*).

14 S. A. Lorenz, E. P. Maziarz and T. D. Woods, *J. Am. Soc. Mass Spectrom.*, 2001, **12**, 795–804.

15 (*a*) C. S. Maier, M. I. Schimerlik and M. L. Deinzer, *Biochemistry*, 1999, **38**, 1136–1143; (*b*) F. Wang and X. J. Tang, *Biochemistry*, 1996, **35**, 4069–4078; (*c*) A. Miranker, C. V. Robinson, S. E. Radford, R. T. Aplin and C. M. Dobson, *Science*, 1993, **262**, 896; (*d*) M. Vairamani and M. L. Gross, *J. Am. Chem. Soc.*, 2003, **125**, 42–43.

16 (*a*) For an N–N distance of 270 pm, N–H N^+ has a bond energy of 12.9 kcal mol^{-1}. See H. Bock, T. Vaupel, C. Näther, K. Ruppert and Z. Havlas, *Angew. Chem., Int. Ed. Engl.*, 1992, **31**, 299–301; (*b*) G. M. Blackburn and M. J. Gait, *Nucleic Acids in Chemistry and Biology*, 2nd edn, 1996, Oxford University Press, ch. 2, p20.

17 A. T. Phan and J.-L. Mergny, *Nucleic Acids Res.*, 2002, **30**, 4618–4625.

18 pH dependent studies at 15 °C on unbuffered solutions of d(TC_5) at millimolar concentrations by NMR evidenced i-motif formation in pH range 2.2–6.8: J.-L. Leroy, K. Gehring, A. Kettani and M. Guéron, *Biochemistry*, 1993, **32**, 6019–31.

19 G. M. Blackburn and M. J. Gait, *Nucleic Acids in Chemistry and Biology*, Oxford University Press, Oxford, 2nd edn, 1996, ch. 2, p18.

20 N. K. Sharma and K. N. Ganesh, *Chem. Commun.*, 2005, 4330–4332.

