Supporting Information

PNA forms an I-motif.

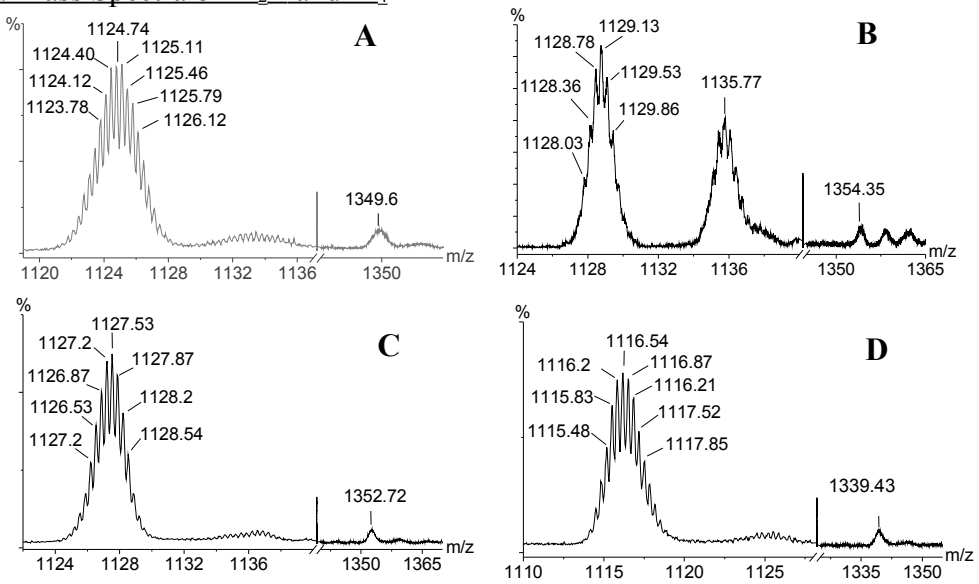
Yamuna Krishnan-Ghosh, Elaine Stephens and Shankar Balasubramanian*.

University Chemical Laboratories, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

Synthesis and Characterization of p(C_5T).

All chemicals were obtained from Sigma-Aldrich unless otherwise stated. Fmoc protected PNA monomers and coupling reagents for solid phase PNA synthesis were obtained from Applied Biosystems (Cheshire, UK). Resins were obtained from Novabiochem. Solvents and other chemicals used were of analytical grade. Milli-Q water was used for all aqueous chemistry. Nano-electrospray ionization mass spectra (Nano-ESI-MS) were recorded on a Q-TOF-1 mass spectrometer (Micromass, UK). The PNA oligomer was synthesized on tentagel resin with a Wang linker using standard protocols employing Fmoc chemistry.((i)Christensen,L., Fitzpatrick,R., Gildea,B., Petersen,K.H., Hansen,H.F., Koch,T., Egholm,M., Buchardt,O., Nielsen,P.E., Coull,J. and Berg,R.H. *J. Peptide Sci*. 1995, **3**, 175-183. (ii) Koch,T., Hansen,H.F., Anderson,P., Larsen,T., Batz,H.G., Otteson,K. and Orum,H. *J. Peptide Res*. 1997, **49**, 80-88.) Deprotection and cleavage from the resin was achieved simultaneously with 95/2.5/2.5 TFA/ water/TIS. The solvent was evaporated under reduced pressure and preciptated using dry ether. The PNA was purified by RP-HPLC. **p(C_5T)**: retention time = 11.8 min. ESI-MS **p(C_5T)** monoisotopic $[M+H]^+$ 1669.65 (calc 1669.66), $[M+2H]^{2+}$ 835.3 (calc 835.33), $[M+3H]^{3+}$ 557.20 (calc 557.22). **p(C_5T)** was stored as a 0.3 M aqueous solution at 4ºC.

Expanded Mass Spectra of M_2^{3+} and M_4^{5+}



Figures SI- A-D show expanded plots of the peaks corresponding to M_2^{3+} and M_4^{5+} for samples (**2**), (**3**), (**4**) and (**5**) respectively. The separation between the dimer peaks averages 0.33 Da indicating that this peak is due to a triply charged species. MS/MS on the peaks in the regime m/z 1325-1400 confirmed that this peak was due to a quintuply charged tetramer (See Figure SI-A).

MS-MS on the Tetramer, M_4^{5+}:

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+}. MS-MS on the peak at m/z 1336 at energies of 30 eV yielded peaks corresponding to M_2^{3+} and

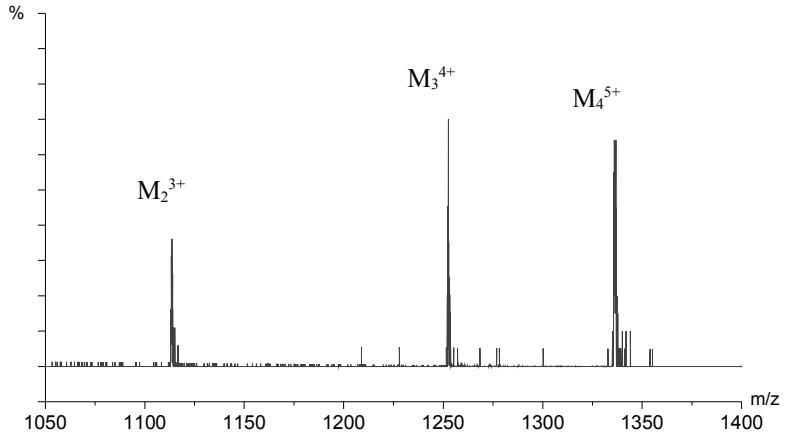


Figure SI-E. MS-MS on the tetramer species M_4^{5+} at 30 eV.

M_3^{4+} (Figure SI-E) in addition to M^+. However at higher energies, (50-60 eV) only M^+ and M^{2+} were observed (not shown). The peaks corresponding to M_2^{3+}, M_3^{4+} were always observed simultaneously. This suggests that the removal of one strand from the tetramer (to give M_3^{4+}) just as easily releases the second strand (leaving behind M_2^{3+}) with the simultaneous appearance of monomeric **p(C$_5$T)**. This is consistent with an i-motif type arrangement, where removal of one of the four strands leaves behind a dimer with a loosely associated third strand. Therefore for clarity, in the main manuscript we have chosen to present the analysis on M_2^{3+} rather than M_4^{5+} to elucidate hydrogen bonding in the component subunit.

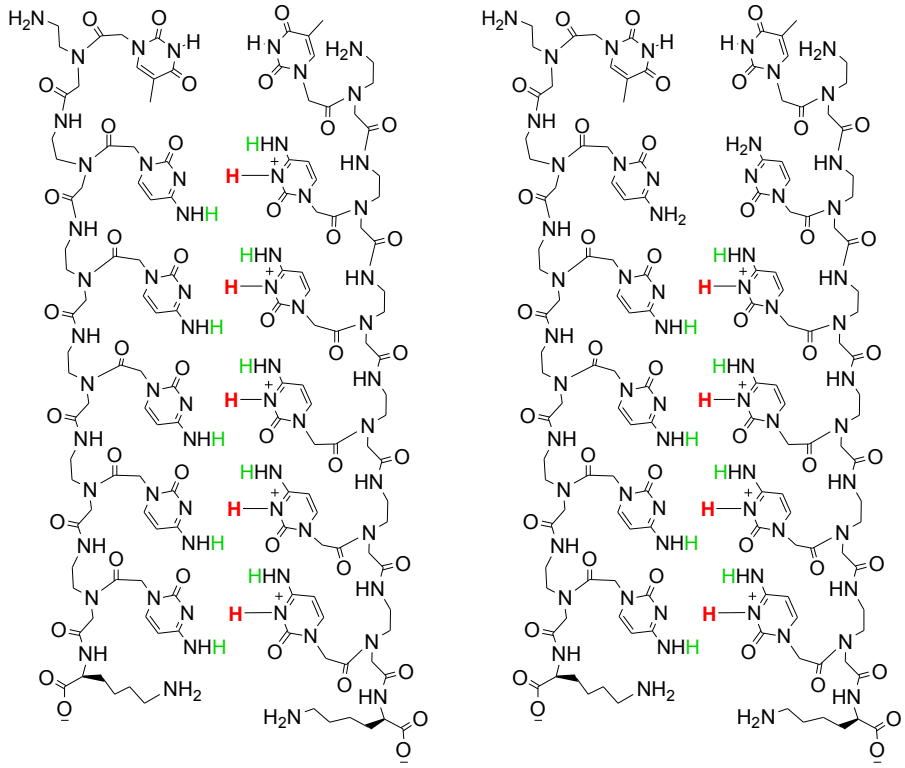
M_4^{5+} forms an I-motif – H/D exchange analysis

The putative structure of M_4^{5+}, the unsymmetrically charged tetramer is shown below. From the peak intensities, this tetramer is a relatively minor population. It consists of two dimeric subunits (**F**) and (**G**). (**F**) is a dimer with a structure identical to M_2^{3+} (Chart 1, Main M/S) with 5 protons and 5 C-C$^+$ base pairs. (**G**) is a dimer (M_2^{2+}) with 4 buried protons and hence, 4 C-C$^+$ base pairs. Thus, M_4^{5+}, the unsymmetrically charged tetramer represents a small population of an i-motif where one C-C$^+$ pair out of a total ten has frayed. In fact, H/D exchange revealed that M_4^{5+} had ~9 protons that were impossible to exchange (See SI-Table 1 (**3**)-(**4**)).

SI-Table 1. Observed m/z and associated molecular weight (MW) as determined by Nano-ESI-MS before and after H/D exchange.

Sample	Observed m/z (M_4^{5+})	Computed MW of M_4^{5+}	Expected[a] m/z of M_4^{5+}	Expected[a] MW of M_4^{5+}	Sites exchanged in M_4^{5+}
(**1**)	1336.18	6680.9 ± 1.06	1335.93	6679.65	0
(**2**)	1349.6	6748.0 ± 1.56	1349.21	6746.05	66
(**3**)	1354.35	6772.2± 1.13	1354.64	6773.20	94
(**4**)	1352.72	6763.6 ± 1.52	1352.83	6764.15	85
(**5**)	1339.43	6697.15 ± 1.02	1339.55	6697.75	66[c]

[a] Molecular weight calculated assuming the number of exchanges listed in the right-most column; [b]Number of deuteriums that have undergone back-exchange (D/H) in H_2O.



F **G**

SI- Figure F&G: M_4^{5+} is composed of two duplexes **F** and **G**, with 5 and 4 C-C$^+$ base pairs respectively. The green hydrogens are involved in neutral H-bonds and exchange upon heating; The red hydrogens are involved in buried, charged H-bonds and were not exchanged even upon heating.