**SUPPORTING INFORMATION FOR THE ARTICLE**

**The PNA-DNA Hybrid I-motif: Implications for sugar-sugar contacts in i-motif tetramerization.**

*Souvik Modi, Ajazul Hamid Wani, Yamuna Krishnan*

National Centre for Biological Sciences, TIFR, GKVK, Bellary Road, Bangalore 560 065, India.

Email: yamuna@ncbs.res.in.

**UV Spectrophotometry:**

Stock solutions of PNA and DNA were made in Millipore water and samples were prepared from a stock solution of 1 mM DNA \( d(TC_5) \) and 2 mM PNA \( p(C_5T) \) by diluting in 30 mM sodium acetate buffer, pH 4.5, to achieve the desired strand concentration. An equimolar mixture of 25 \( \mu M \) \( d(TC_5) \) and 25\( \mu M \) \( p(C_5T) \) were mixed to form the hybrid i-motif, where 25 \( \mu M \) \( d(TC_5) \) was taken to form DNA\(_4\) i-motif. The mixture was heated and maintained at 90°C for 10 min, cooled to 4°C over 3h at a rate of 0.5°C per minute and equilibrated at 4°C for 8 h. Melting experiments were monitored at 295 nm with increasing the temperature from 20°C to 90°C at a rate of 0.5°C/minute. SI-Figure 1 shows the melting profiles of DNA\(_4\) and Hybrid i-motif at the same strand concentration. We see that the hybrid i-motif shows a much more well-defined melting profile than the DNA\(_4\) i-motif and also shows a more cooperative melting transition, especially at lower concentrations.

![SI-Figure 1](image-url)
Table 1. Thermal denaturation at different strand concentration monitored at 295nm.

| Concentration (µM) | $T_{1/2}$ ($\pm$ 0.2 °C) |
|-------------------|--------------------------|
| 100               | 51.8                     |
| 50                | 51.4                     |
| 25                | 51.1                     |
| 10                | 51.05                    |
| 6                 | 51.0                     |
| 4                 | 50.8                     |
| 2                 | 50.5                     |

Thermal denaturation at different heating rate:

SI-Figure 2
Thermal denaturation of the hybrid complex at different heating rates showed a change in $T_{1/2}$ values indicating that the melt is not an equilibrium melt. This is supported by the fact that the sample showed hysteresis between the cooling and heating process. Thus the melts reflect solely the dissociation process of the hybrid i-motif, and the melting temperatures are represented by $T_{1/2}$. Thermal melting experiments on DNA$_4$ at different heating rates showed a similar $T_{1/2}$ dependence on heating rates, consistent with earlier results in the literature (Collin, D. and Gehring, K. (1998) Stability of Chimeric DNA/RNA cytosine Tetrads: Implications for i-motif formation by RNA. J. Am. Chem. Soc., 120, 4069-4072)
Stability as a function of pH:
The stability of the hybrid i-motif was investigated as a function of pH over the range pH 3 to 7. Thermal denaturation profiles of the hybrid revealed that the hybrid i-motif formed efficiently between the pH 4.2 to 5.7, above pH 5.7 and below pH 4.2, i-motif formation could not be detected by UV melting (SI-Figure 3).

Circular Dichroism confirmed the fraction of hybrid i-motif formed:

In order to clarify that the CD signal at 288 nm indeed reflects hybrid i-motif formation, we recorded CD spectra as a function of pH (Figure 4). The CD signal at 288 nm progressively decreased upon increasing or decreasing the pH from an optimal value of pH 4.5 which is in good correlation with the UV studies. The maximum signal at 288 nm is obtained at pH 4.5, near the pKₐ of N3 on cytidine revealing that this signal is due to C-C⁺ base pair formation.

Circular dichroism studies at different strand concentrations of hybrid i-motif gave an indication as to the fraction of hybrid i-motif formed. All the samples were made after mixing equimolar p(C₅T) and d(TC₅) in 30 mM sodium acetate buffer at pH 4.5. CD signal at 288 and 265 nm was converted to molar ellipticity. A plot of molar ellipticity as a function of strand concentration proved to be invariant over a concentration regime 25 µM to 200 µM (Figure 5A). This reaffirms that all the DNA strands are completely complexed in the form of hybrid i-motifs over this concentration regime. We performed thermal denaturation of the hybrid i-motif by monitoring the CD value at 288 nm as a function of temperature. SI-Figure 5B shows the
CD melting profile of the DNA₄ (100 µM strand concentration) i-motif and the hybrid i-motif (100 µM strand concentration). As the complex melts into single strands, CD-melts follow the change in chirality between complexed and single stranded states, while UV-melts follow the change in nucleobase extinction coefficients between the complexed and single stranded states. Given that two of the strands in the hybrid are achiral, it is not surprising that these two properties vary differently for hybrid and DNA₄ i-motif melting. This could explain the ~8°C and ~3°C difference in T₁/₂ values between the UV and CD measurements respectively.

**SI-Figure 5** shows dependence of CD signal with pH. Gradual decrease in signal intensity at 288 nm as well as 265 nm reveals relative pH regime of hybrid i-motif formation.

**Model of the hybrid i-motif to obtain theoretical distance estimates**

**SI-Figure 6.** Model of the hybrid i-motif constructed using PyMOL software using the NMR structure parameters of the DNA₄ i-motif (Gehring, K., Leroy, J.-L., Guéron, M. (1993) A tetrameric DNA structure with protonated cytidine–cytidine base pairs. Nature, 363, 561-565). End to end distances Rcalc, were measured incorporating the linker connecting the fluorophores.
FRET Measurements:

Distance calculation:

Energy transfer (\%E) = (1 - \frac{F_D}{F_{DA}})

R = R_o \times (1/E - 1)^{1/6}

Where

F_D = Fluorescence intensity of donor without acceptor.
F_{DA} = Fluorescence intensity of donor with acceptor.
R = Distance to be calculated.
R_o = Förster distance.

Controls:
The quantity of fluorophore-labelled strands in the control and the hybrid was verified by UV quantitation of the TMR label prior to fluorescence experiments. Control measurements employ (i) the relevant free TMR-DNA strand, as well as free TMR-PNA strand (in the absence of the respective PNA or DNA cognate partner) under identical conditions in order to correct for quenching due to non-FRET mechanisms. (ii) Post measurement, the samples were heated to melt the complex and quenched to RT and fluorescence readings of the dissociated components at RT concurred with the controls of samples prepared as in (i). r-values were less than 0.1 and hence \kappa^2 was assumed to be 2/3.

| Labeled Hybrid complex                  | Calc. Distance | Calc. Intensity | Obs. Intensity | Anisotropy^* | Obs. Distance |
|----------------------------------------|----------------|----------------|----------------|--------------|--------------|
| TMR-p(C5T)+d(TC5)                      | \sim 51 Å      | \sim 97x10^6   | \sim 108x10^6  | 0.047        | 53 ± 5 Å     |
| 3'-TMR-d(TC5)+p(C5T)                   | \sim 50.4 Å    | \sim 17x10^5   | \sim 18.3x10^5 | 0.045        | 51.7 ± 5 Å   |
| 5'-TMR-d(TC5)+p(C5T)                   | \sim 51 Å      | \sim 12.5x10^5 | \sim 12.5x10^5 | -            | 50.8 ± 5 Å   |
| 3'-TMR-d(TC5)+N-Dabcyl-p(C5T)          | \sim 20 Å      | \sim 6.9x10^5  | \sim 8.93x10^5 | 0.0365       | 21.7 ± 5 Å   |
| 5'-TMR-d(TC5)+N-Dabcyl-p(C5T)          | \sim 23 Å      | \sim 8.05x10^5 | \sim 9.3x10^5  | -            | 24.5 ± 5 Å   |

^a A model of the hybrid i-motif was constructed using PyMOL software using the NMR structure parameters of the DNA$_4$ i-motif (Gehring, K., Leroy, J.-L., Guéron, M. (1993) A tetrameric DNA structure with protonated cytidine–cytidine base pairs. *Nature*, **363**, 561-565). End to end distances $R_{calc}$ were measured incorporating the linker connecting the fluorophores. ^b Expected intensity ($F_{DA}^{calc}$), based on the calculated quenching efficiency (E), based on the calculated distance $R_{calc}$, for an experimentally obtained $F_D$. ^c Each value is the mean of four independent experiments where measurements were performed on a given complex prepared four separate times, variation in values between experiments was ± 1 Å. The accuracy of the distance is only as much as the linker lengths of the freely rotating fluorophores.

* Average value from four independent experiments were taken
Homo FRET studies were done on hybrid i-motif composed of labeled DNA with unlabelled PNA strands to determine the respective DNA strand alignment in complex (SI – Figure 7). Samples were made by mixing equimolar 5'-TMR-d(TC₅) and p(C₅T) and formed as described earlier. As a control, fluorescence of uncomplexed 5'-TMR-d(TC₅) alone at the same strand concentration was used. Around 24% quenching with respect to the control was observed when 5'-TMR-d(TC₅) was present as part of hybrid i-motif. This corresponds to a distance of ~50.8 ± 1 Å (calc. 51 Å).

SI-Figure 7

v) Kinetics of i-motif formation:

Intrigued by the superior complexation properties of PNA and DNA in the context of hybrid i-motif formation, we proceeded to investigate the kinetics of association of these component strands. We compared the association kinetics of 1:1 p(C₅T) : d(TC₅) at different strand concentrations to form the hybrid i-motif and compared it with the association of DNA strands at identical concentrations into a DNA i-motif. Association was initiated by the addition of 30 mM acetate buffer, pH 4.5 to aliquots of a 1:1 p(C₅T) and d(TC₅) in a low volume cuvette. Association was then followed at 20°C by monitoring the increase in absorbance at 295 nm due to i-motif formation from the single strands. SI-Figure 8 shows the association of 1:1 p(C₅T) : d(TC₅) at 50 µM each, and a control of 100 µM d(TC₅) alone under the same conditions. We observed the association of the hybrid is almost complete within 20 min. In contrast, association of d(TC₅) into the DNA i-motif does not even reach 5% completion in 80 min. Kinetics of hybrid i-motif formation as a function of strand concentration revealed two distinct processes occurring on two different timescales (τ₁ and τ₂) (Figure 4B). The faster reaction (τ₁ = 140 s, k₁ = 7.04× 10⁻³ s⁻¹) was found to be first-order in p(C₅T) concentration (see Table 2). A
more rigorous analysis of the association and dissociation kinetics of the hybrid i-motif is the subject of another study and will be presented elsewhere.

**Table 1.** Rates and associated rate constants of hybrid i-motif formation from 1:1 p(C₅T) and d(TC₅) at different strand concentrations.

| d(TC₅) (µM) | p(C₅T) (µM) | τ₁ (s)  | r₁ (µM s⁻¹) | k₁ (s⁻¹) | τ₂ (s)  |
|-------------|-------------|---------|-------------|----------|---------|
| 5           | 5           | 135.1   | 3.7 × 10⁻²  | 7.4 × 10⁻³ | 909.1   |
| 25          | 25          | 135     | 18.5 × 10⁻² | 7.4 × 10⁻³ | 925.9   |
| 50          | 50          | 145.3   | 34.4 × 10⁻² | 6.88 × 10⁻³ | 909     |
| 100         | 100         | 142     | 70.4 × 10⁻² | 7.04 × 10⁻³ | 775.1   |

**SI Figure 8.** Association kinetics of p(C₅T) and d(TC₅) single strands into hybrid i-motif. **(A)** Association of 1:1 p(C₅T) and d(TC₅) and d(TC₅) alone into complex form. **(B)** Formation of hybrid i-motif monitored at different strand concentrations. Absorbance at 295 nm was followed at 20°C in 30mM acetate buffer, pH 4.5;
NOESY of the Imino proton region showing 5 C-H+-C base pairs in the hybrid i-motif:

SI-Figure 9

NOESY showing focused crosspeaks of the structured PNA backbone (NH$_2$e and glycine region) in the hybrid i-motif:

SI-Figure 10
**H/D Exchange as a function of incubation time.**

Residual HDO could indeed be a possible explanation of the observation of imino proton signals at the 6 h time point. In order to check whether this was the case, we performed the following analysis of 1D spectra obtained at different time points.

For every spectrum, we chose a reference signal centered at 2.67 δppm as our internal standard, which corresponds to CH₂ groups on the PNA backbone that do not exchange. The area under the reference signal, \( R \) was obtained. For a given spectrum, the area under (a) the imino proton peaks, \( A \) (b) the Thymine imino peaks, \( B \) and (c) the region 8.2 - 8.9 δppm corresponding to the backbone amide protons and NH₂e peaks, \( C \), were obtained. \( A \), \( B \) and \( C \) were normalized with respect to \( R \) for each spectrum. The peak areas for normalized \( A \), \( B \) & \( C \) w.r.t. \( R \) in H₂O were taken as 100% unexchanged protons. Normalized peak areas at different incubation times with D₂O (\( t \)) were plotted as a function of \( t \). Normalized \( A \), \( B \) & \( C \) values at a given \( t \) were converted to percentages of the original signal in H₂O corresponding to \( t = 0 \) and 100% unexchanged protons and plotted as a function of \( t \). The corresponding graph is shown below and also included in the supporting information, pp 9.

If residual HDO is a sufficient explanation, then normalized \( A \), \( B \) & \( C \) areas should, at equilibrium, all level off to the same percentage (corresponding to the percentage of HDO in the sample) of their original signal at \( t = 0 \). However, we see that at \( t = 6 \) h \( C \) has levelled off to \( \sim 3 \) %, indicating that there is \( \sim 3 \) % HDO present in the sample. Interestingly \( A \) and \( B \) have exchanged to only \( \sim 48 \) % and \( \sim 22\% \) respectively, evidencing a slower overall rate of exchange of the cytosine imino protons. The absolute rates of exchange of these protons need not be comparable with their DNA₄ counterparts, as the duplexes here are PNA-DNA duplexes and may have different off-rates than their DNA-DNA analogues.

**SI-Figure 11.**

Percentage unexchanged protons as a function of incubation time in D₂O in the (a) imino proton regions and the non-hydrogen bonded amino and backbone amide regions.