Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry. This journal is © The Royal Society of Chemistry 2015

Supporting Information for

#### Interdependence of Pyrene Interactions and Tetramolecular G4-DNA Assembly

by Osman Doluca,<sup>*a,c*</sup> Jamie M. Withers,<sup>*a*</sup> Trevor S. Loo,<sup>*a*</sup> Patrick J. B. Edwards,<sup>*a*</sup> Carlos Gonzalez,<sup>*b*</sup> Vyacheslav V. Filichev<sup>*a*</sup>

### **Oligonucleotide Synthesis**

Oligonucleotides (ONs) were synthesised with an Mer-Maid 4 automated DNA synthesiser from BioAutomation Corporation using 4,5-dicyanoimidazole (DCI) as an activator. Oxidation and deprotection times were set to 40 s and the activation time at 60 s for 1.0 µmol synthesis scale. For coupling of modified nucleotides, automated DNA synthesis was paused after the deblocking step. TINA phosphoramidite (10 mg per coupling) was added into each 1 µmol column and 750 µL of an activator (0.25 M DCI in dry acetonitrile) was directly injected onto the column under argon by the automated DNA synthesiser. Coupling time was extended for hand coupling (up to 5 min). All oligonucleotides were synthesised in DMT-on mode. TINA-conjugated ONs were purified using reverse-phase HPLC instrument (Waters<sup>™</sup> 600) using a C-18 column from Alltech with 250 mm length and 10 mm internal diameter. The residue was dissolved in H<sub>2</sub>O (1 mL). Buffer A [0.05 M TEAA in H<sub>2</sub>O (pH = 7.0)] and buffer B (75 % acetonitrile, 25 % H<sub>2</sub>O). Flow 2.5 mL min<sup>-1</sup>. Gradients: 2 min 100 % buffer A, linear gradient to 100 % buffer B in 48 min, linear gradient to 100 % buffer A in 2 min, 100 % buffer A for 10 min. The corresponding UV-active fractions were freeze-dried. After purification, DMT-on-ONs were treated with 80 % aqueous AcOH (100 µL) for 20 min at +4 °C to remove 5'-O-DMT group. NaOAc (3 M, 50 µL) followed by EtOH (1 mL) were added and vortexed. Solutions were cooled to -18 °C for two hours in order to precipitate the ONs. Samples were centrifuged for 20 min at 13000 rpm and supernatant was removed. Pellets were washed with EtOH (2  $\times$  500 µL), centrifuged and the supernatant was removed. Then, the rest of the solvent was evaporated at 50 °C for 15 min. 100 µL H<sub>2</sub>O was added to dissolve the ONs. Purities of ONs were confirmed by denaturing gel electrophoresis using 20 % polyacrylamide gel (0.75 mm thickness, 19:1 acrylamide/bisacrylamide ratio) and found to be more than 95 % pure. Gels were prepared in 1× TBE buffer (100 mM Tris, 90 mM boric acid, and 10 mM EDTA) under denaturing conditions (7 M urea). ONs were loaded onto gels after preincubation at 90 °C for 10 min. Molecular weight of ONs was confirmed by mass-spectroscopy analysis (Table S1).

## **Polyacrylamide Gel Electrophoresis**

For non-denaturing 20 % PAGE, gels were prepared in HEPES buffer (50 mM) in corresponding salt concentrations (110 mM NaCl or KCl, the same buffer was used as a running buffer) with 0.75

mm thickness, 19:1 acrylamide/bisacrylamide ratio. ONs were prepared at 100  $\mu$ M strand concentration, incubated in 50 mM HEPES using the same salt concentrations and heated up to 90 °C for between 10 and 30 min before cooling down and incubating at 4 °C overnight. All gel electrophoresis were performed at room temperature. After the electrophoresis, gels were stained with 5 % Stains-All® in 50 % water/formamide for 5–10 min and then destained in H<sub>2</sub>O until complete washing of the dye from the gel background occurred.

## **UV-Vis Spectroscopy**

UV-Vis spectroscopy was performed using Cary 100Bio UV-Vis spectrometer using quartz cuvettes with 1 cm pathlength and a 2 × 6 multicell block with a Peltier temperature controller. Extinction coefficient ( $\epsilon$ ) of TINA is 22000 [L/(mol×cm)].

# 1.1.1 UV-Vis Thermal Difference Spectra (UV-Vis TDS)

Thermal difference spectra were determined for the DNA complexes (with 10  $\mu$ M strand concentration) by subtraction of UV-Vis spectra obtained in Li cacodylate buffer (10 mM) and NaCl or KCl (110 mM) at 20 °C after 2-7 days on incubation of samples at 4°C from UV-Vis spectra obtained after 30 min incubation at 90 °C.

## 1.1.2 Determination of Melting and Annealing Temperatures

The melting temperatures ( $T_m$  [°C]) were determined for G-quadruplexes [prepared at 10 µM strand concentration in Li cacodylate buffer (10 mM) and NaCl or KCl (110 mM)] as the maxima of the first derivative plots of the melting curves obtained by measuring absorbance at 373 nm against increasing temperatures (0.18, 0.5 or 1 °C/min). The change in absorbance was also recorded at 295 nm (see Figure S3).

## Association Rate Constant (kon) Measurements

Association rate constants ( $k_{on}$ ) for G-quadruplex formations were measured using CD and UV-Vis spectroscopy as previously described (Mergny, De Cian et al. 2005). Stock solution of ONs was preincubated at 90 °C for 30 min and then transferred into Li cacodylate buffer (10 mM) and NaCl (110 mM) solution at 20 °C. Stock solution of dTG<sub>6</sub>T was treated with LiOH at 90 °C for 15 min to ensure full denaturation of the complex followed by HCl neutralisation. As soon as the ONs were diluted in a buffer, changes in absorbance at 275, 280, 295 and 373 nm or changes in CD signal at

265 (G3X, TG<sub>6</sub>T) or 295 nm (GXG) were then recorded every 5 min for 12.5 h. The data were used to plot the change in absorbance versus time (s) in Kleida Graph software. A curve was fit to the data using the following equation (Mergny, De Cian et al. 2005):

$$Abs_{t} = Abs_{f} + (Abs_{i} - Abs_{f}) \times (1 + (n-1) \times k_{on} \times \mathbb{C}^{n-1} \times t)^{\frac{1}{1-n}}$$

 $Abs_i$  refers to absorbance of the sample when all strands exist in unfolded, single-stranded form.  $Abs_f$  represent the absorbance of the sample when equilibrium is reached between folded and unfolded forms of the ONs. C represents the strand concentration, *n* represents the order of the reaction, which is 4 for tetramolecular G-quadruplexes, and *t* represents the time elapsed in seconds. Experimental data and  $k_{on}$  values are presented in Figure S5 and Tables 1 and S4.

## **Circular Dichroism (CD) Spectroscopy**

CD spectra were recorded using a Chirascan CD spectrophotometer (150 W Xe arc) from Applied Photophysics with a Quantum Northwest TC125 temperature controller. CD spectra (average of at least 3 scans) were recorded between 200 and 500 nm with 1 nm intervals, 120 nm/min scan rate and 1 cm pathlength followed by substraction of a background spectrum (buffer only). CD spectra of G-quadruplexes were recorded at 10  $\mu$ M strand concentration in Li cacodylate (10 mM) supplemented with NaCl or KCl (110 mM) at pH 7.2 or in 150 mM ammonium acetate buffer, pH 7.0 at 20 °C.

#### Fluorescence Spectroscopy

Fluorescence spectroscopy studies were performed using a FluoroMax-4 Spectrofluorometer from HORIBA Scientific. Fluorescence emission spectra were recorded for all TINA-containing ONs between 380 and 600 nm when excited at 373 nm (excitation and emission slits were 2 nm). Fluorescence excitation spectra were recorded between 250 and 450 nm while monitoring emission at 480 nm (excitation and emission slits were 2 nm). Samples were prepared as described for UV-Vis spectroscopy and adjusted to an identical optical density at an excitation wavelength of 373 nm. For convenience the values in the y-axis (fluorescence intensity) were divided by 10<sup>4</sup>.

#### ESI-mass Spectrometry

TINA G4-DNA samples were annealed at 400-500  $\mu$ M concentration in a Na-containing buffer (75-100  $\mu$ L), 10 M aq. ammonium acetate (35  $\mu$ L) was added and DNA were precipitated upon addition of ethanol (1 mL) followed by incubation for 2 hr at -20 °C. Samples were centrifuged, solvent was decanted and pillets after washing with ethanol (2 × 500  $\mu$ L) were dried *in vacuo*. Pillets were dissolved in 150 mM ammonium acetate buffer (250  $\mu$ L, pH 7.0). Samples were left at +4 °C for 5 days. Each sample was diluted to 10  $\mu$ M concentration in 150 mM ammonium acetate buffer and just before an injection samples was further diluted with 15% MeCN. Samples were also analysed using 15 % MeOH as a co-solvent but it did not result in an improved electrospray signal. Each diluted sample (3  $\mu$ L) was injected into an Agilent 1200 analytical liquid chromatography system (Agilent Technologies, Hanover, Germany) delivering an isocratic gradient containing 50% (v/v) acetonitrile and 50% (v/v) water at a flow rate of 200  $\mu$ L min<sup>-1</sup> in bypass mode. The LC eluent was directly infused into a duel electrospray ionization (ESI) source coupled to an Agilent 6520 quadrupole time of flight (Q-TOF) mass spectrometer (Agilent Technologies, Hanover, Germany). Mass drifts were constantly corrected with reference standards in real time. Ions generated by ESI operated at a nitrogen flow of 5.0 L min<sup>-1</sup>, gas temperature of 290 °C and a neubiliser spray pressure at 30 psi were analysed in negative ion mode with a capillary voltage of 3,500 V, a fragmentator voltage of 175 V and a skimmer voltage of 65 V.

Total ion count (TIC) was recorded in profile mode over the m/z range of 100-3,000 and analysed using Agilent MassHunter Workstation Qualitative Analysis software version B.6.0633.10 (Agilent Technologies, Hanover, Germany).

#### NMR Spectroscopy of Oligonucleotides

NMR spectroscopy of G-quadruplexes formed by **GXG** and **TXG** were performed on 700 MHz Brüker instrument and processed with Topspin software. ONs (100, 500 or 1000  $\mu$ M) were incubated overnight in 10 % D<sub>2</sub>O, 10 mM Na<sup>+</sup> phosphate buffer supplemented with or without NaCl (100 mM) or KCl (10 mM), pH 7.0 at 20 °C prior to NMR spectroscopy. Trimethylsilyl propionate (TSP, 25  $\mu$ M) was used as internal standard in <sup>1</sup>H NMR ( $\delta$  = 0.003 ppm).

IP-COSY (1), TOCSY and NOESY experiments were recorded in  $H_2O/D_2O$ . The NOESY spectra were acquired with mixing times of 100 and 250 ms, and the TOCSY spectra were recorded with standard MLEV-17 spin-lock sequence, and 60-ms mixing time. Water suppression was achieved by including either a WATERGATE (2) or excitation-sculpting (3) module in the pulse sequence prior to acquisition. NOESY spectra in  $D_2O$  were acquired with 150 ms and 250 ms mixing times. Two-dimensional experiments in  $D_2O$  were carried out at temperatures ranging from 5 °C to 25 °C, whereas spectra in  $H_2O$  were recorded at 5 °C to reduce the exchange with water. The spectral analysis program Sparky (4) was used for semiautomatic assignment of the NOESY cross-peaks and quantitative evaluation of the NOE intensities.

**NMR-based molecular modeling.** The computer package Sybyl was used to build TINA structure and to perform the adequate insertion in the DNA sequence. Structural model were calculated with the SANDER module of the molecular dynamics package AMBER (5). Parametrization of the AMBER force field for TINA was carried out following standard methods in ANTECHAMBER program. Charges were calculated using RESP/6-31G(d). In a first step, a model

of a dimeric quadruplex formed by the association of two hairpins was built. Atomic coordinates for initial modeling were taken from of the solution structure of the Bombyx Mori telomeric Gquadruplex (1AFF) determined by Kettani, et al (6). This structure was chosen because it is stabilized by two G-tetrads and its content of syn and anti guanines is consistent with the experimental information obtained for our TINA conjugates. Only the guanine residues were considered. Two TINA blocks were used to connect contiguous antiparallel G-tracts, giving rise to two hairpins oriented in a head-to-head way. Thymine residues were added at the 5'- and 3'-ends of each of the resulting hairpins. This model was energy minimized and submitted to a short molecular dynamics run in vacuo to eliminate bas steric contacts. The relative orientation of the two TINA moieties was constrained during these calculations. As a second step, the resulting dimeric structures were used to build a tetramer consisting of two dimeric quadruplexes with the program Sybyl. The relative orientation of the two quadruplexes was selected manually trying to maximize TINA-TINA contacts. The resulting structures were submitted to a molecular dynamics run including explicit solvent, periodic boundary conditions and the Particle-Mesh-Ewald method to evaluate long-range electrostatic interactions (7). Thus, the structures obtained in the previous step were placed in the center of a water-box with around 4000 water molecules and 24 sodium counterions to obtain electroneutral systems. We used the parmbsc0 revision of the parm99 force field (8) including suitable parameters for the TINA nucleotide calculated for this work. The TIP3P model was used to describe water molecules (9). The protocol for the constrained molecular dynamics refinement in solution consisted of an equilibration period of 160 ps using a standard equilibration process (10), followed by 10 independent 500 ps runs. Averaged structures were obtained by averaging the last 250 ps of individual trajectories and further relaxation of the structure.

Analysis of the representative structures as well as the MD trajectories was carried out with the programs MOLMOL (11), the analysis tools of AMBER, and other "in house" programs.

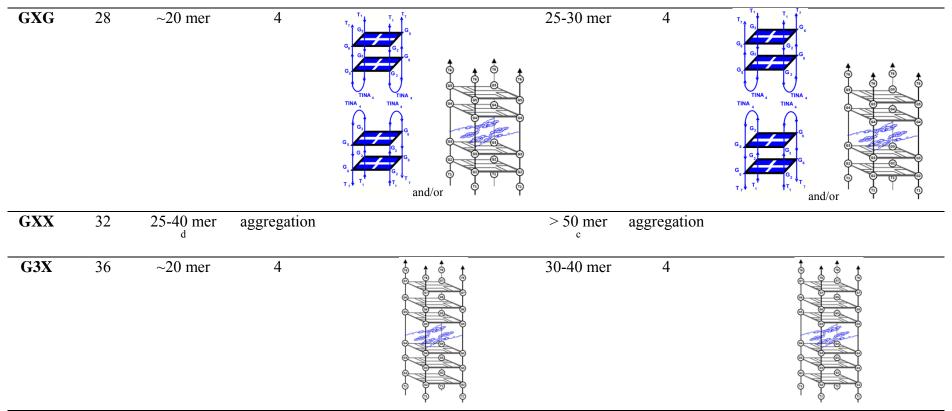
Oligonucleotide	Calculated m/z, Da	Observed m/z, Da
XTG	2331.7	2330.6
TXG	2331.7	2330.6
GXG	2331.7	2329.6
GXX	2800.1	2801.3
G3X	2990.1	2989.6

 Table S1. Results of mass spectroscopy analysis of ONs synthesised.

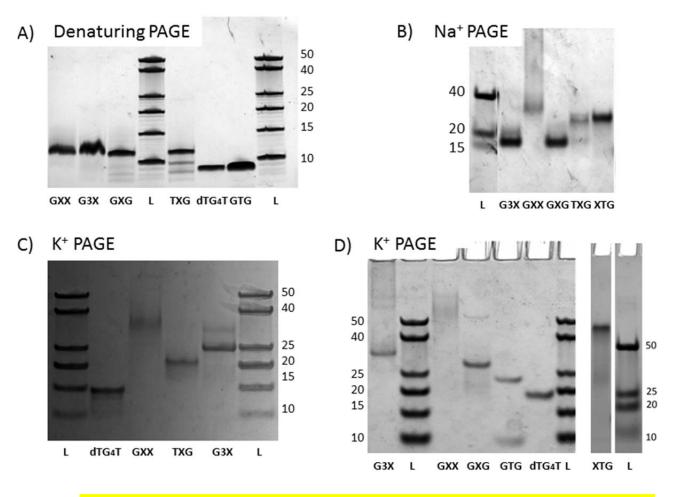
[a] Mass spectrometry analysis was performed using a M@LDI micromass instrument (from Waters) in the positive-ion mode using 2-aminobenzoic acid or 6-aza-2-thiothymine as matrices and dibasic ammonium citrate or imidazole as co-matrices. ONs were desalted using C<sub>18</sub>ZipTips (Millipore) prior to loading on the MALDI plate.

**Table S2.** Retardation of the G-quadruplexes formed by TINA-TG<sub>4</sub>T sequences in native PAGE (20 %) and molecularity in the presence of 110 mM NaCl or KCl in 50 mM HEPES buffer at pH 7.2, room temperature.

	N <sup>a</sup>		1	NaCl			KCl
	$(N=n\times 4)$	mobility <sup>b</sup>	molecularity	Proposed assembly <sup>f</sup>	mobility <sup>b</sup>	molecularity	Proposed assembly
TG4T	24	10-15 mer	4		15-20 mer	4	
XTG	28	25-30 mer	4		> 50 mer	aggregation	and at high DNA concentrations
TXG	28	25-30 mer	4		25 mer <sup>c</sup>	4	
GTG	28	10-15 mer	ssDNA <sup>e</sup>		~10 mer 20-25 mer	ssDNA 4	

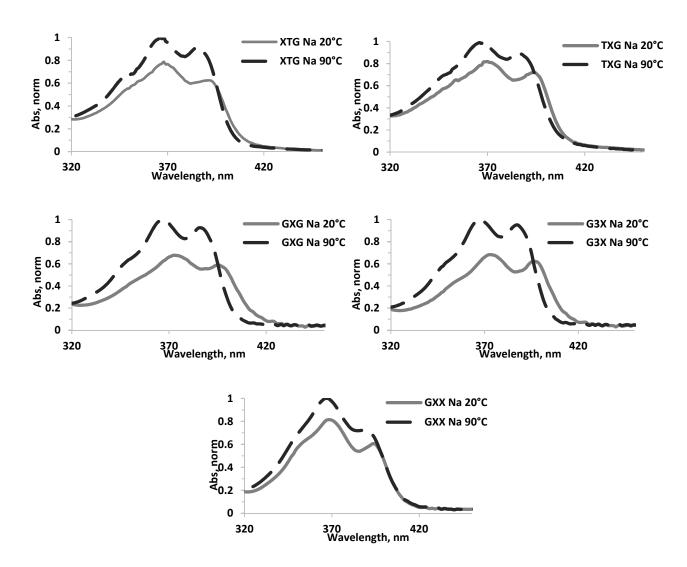


a) N refers to number of nucleotides/modifications within the tetramolecular G-quadruplex formed by corresponding sequence. b) Level of mobility in comparison to the oligothymidylate ladder (see Figure S1 for discussion). c) a slower migrating band between  $dT_{25}$  and  $dT_{40}$  is also observed on the gel. d) Observed as a smear on the gel. e) ssDNA is single-stranded DNA. f) the images of proposed assemblies are schematic.

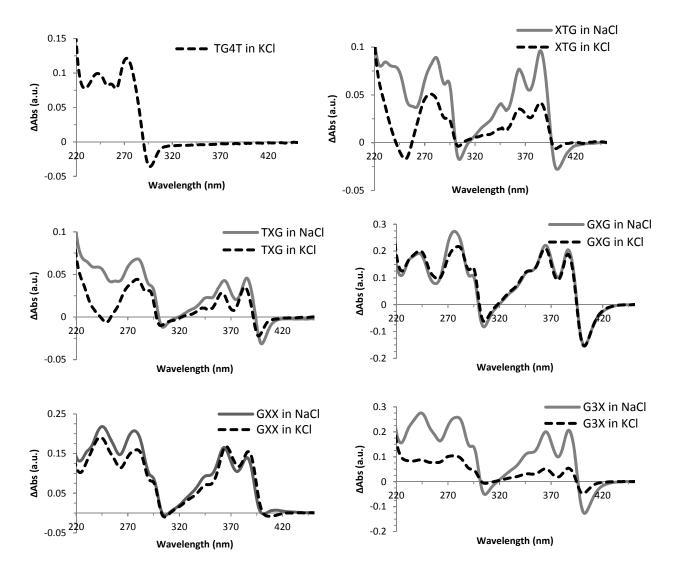


**Figure S1.** A) Denaturing PAGE (20 %, 7M urea) analysis of oligonucleotides (100  $\mu$ M, samples were heated at 95 °C for 30 min in 7M urea prior loading); B-D) Representative images of native PAGE (20 %) of oligonucleotides (100  $\mu$ M) in the presence of 110 mM NaCl (B) and 110 mM KCl (C, D) in 50 mM HEPES buffer at pH 7.2, room temperature. Ladder (L) contains 10-, 15-, 20-, 25-, 40- or 50-mer oligothymidylates as indicated.

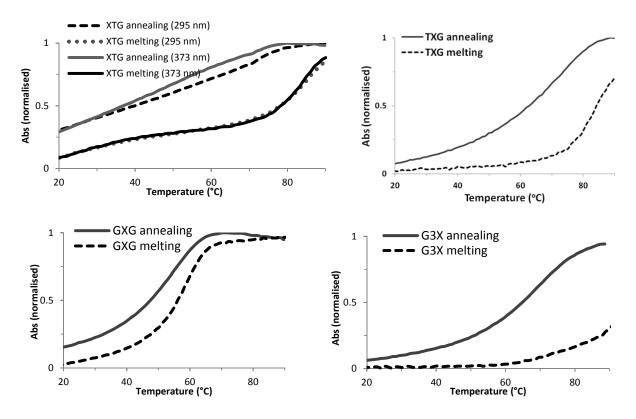
TINA-modified  $TG_{4(6)}T$  sequence have similar mobility in the denaturing PAGE (Fig. S1A), just above  $dT_{10}$  component of the ladder and are slightly retarded in comparison with controls  $dTG_4T$ and **GTG**. In native PAGE, all complexes migrate slower in K<sup>+</sup> than in Na<sup>+</sup> relative to the ladder components. **G3X** and **GXG** sequences form complexes with similar mobility in Na<sup>+</sup> and K<sup>+</sup>. The presence of TINA and an additional phosphate results in slower mobility of **G3X** and **GXG** in comparison with (**TG**<sub>4</sub>**T**)<sub>4</sub> and (**GTG**)<sub>4</sub> in K<sup>+</sup> (Figure S1D). Of note, **GTG** only partially forms G4-DNA at 100 µM oligonucleotide concentration in K<sup>+</sup> (two bands are observed). **TXG** and **XTG** sequences form a hydrophobic pocket at the 5'-end of corresponding G4-DNAs which lead to slower mobility in the native gel (Na<sup>+</sup>) in comparison to **GXG** and **G3X** assemblies (Fig. S1B). In K<sup>+</sup>, **TXG** complex migrates similar to **G3X** (25-30 mer, Fig. S1C), whereas **XTG** has significantly slower mobility (>50 mer, Fig. S1D) indicating that **XTG** forms a dimeric G4-DNA (8-strands) assisted by TINA-TINA aggregation (Table S2). Two transitions detected in melting profile (Figure S4) suggest that dimeric aggregate of (**XTG**)<sub>4</sub> dissociates at 25-30 °C followed by dissociation of G4-DNA (> 90 °C). It also means that 5'-dT prevents formation of aggregates in **TXG** both Na<sup>+</sup> and K<sup>+</sup>. **GXX** sequence form a slow migrating smear in both Na<sup>+</sup> and K<sup>+</sup> which suggest aggregation.



**Figure S2.** Normalised UV-vis spectra of TINA-TG<sub>4</sub>T constructs in the TINA absorbance region at 10  $\mu$ M strand concentration before (90 °C) and after annealing (20 °C) in the presence of NaCl in 10 mM Li cacodylate buffer at pH 7.2.



**Figure S3.** UV-Vis Thermal difference spectra of G-quadruplex samples at 10  $\mu$ M strand concentration after annealing in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2.

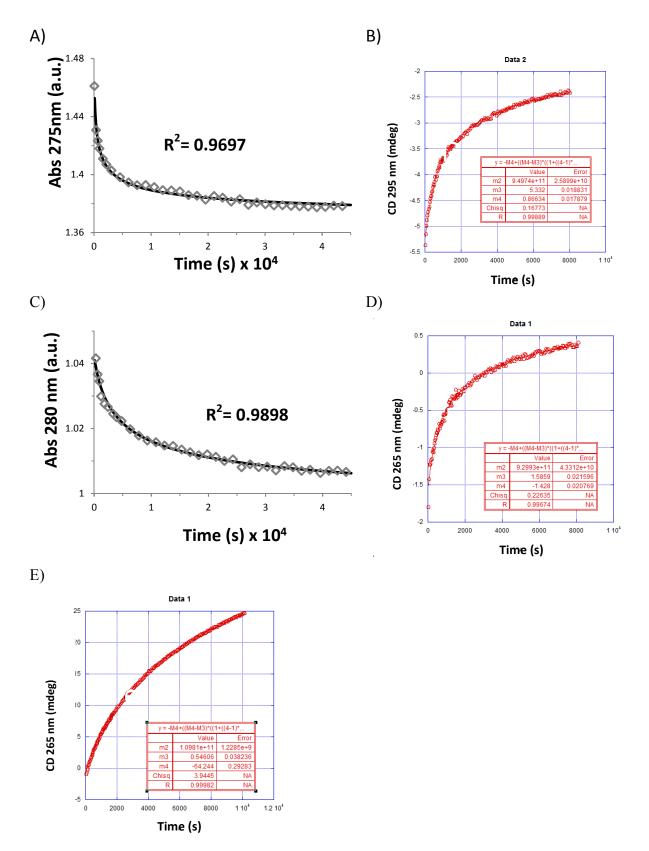


**Figure S4.** Melting and annealing profiles of modified G-quadruplex assemblies after incubation in 110 mM NaCl, 10 mM Li cacodylate buffer at pH 7.2. Oligonucleotide concentration was 10  $\mu$ M. The profiles are based on absorbance data recorded at 373 nm apart from **XTG** where profiles are reported at 295 and 373 nm (0.18 °C/min temperature ramp).

**Table S3.**  $T_{\frac{1}{2}}$  values of **GXG**, **GXX** and **G3X** obtained during melting and annealing, using 0.5 or 1.0 °C/mintemperature ramp in the presence of 110 KCl in 10 mM Li cacodylate buffer at pH 7.2.

	0.5 °C/min	1.0 °C/min	
	melting/annealing (°C)	melting/annealing (°C)	
GXG	59.5 / 47.5	62.5 / 44.5	
GXX	41.0 / 36.5	42.0 / 38.5	
G3X	>90 <sup>a</sup> / 41.5	>90 <sup>a</sup> / 41.5	

<sup>a</sup> melting process was complete only after 30 min incubation at 90 °C.



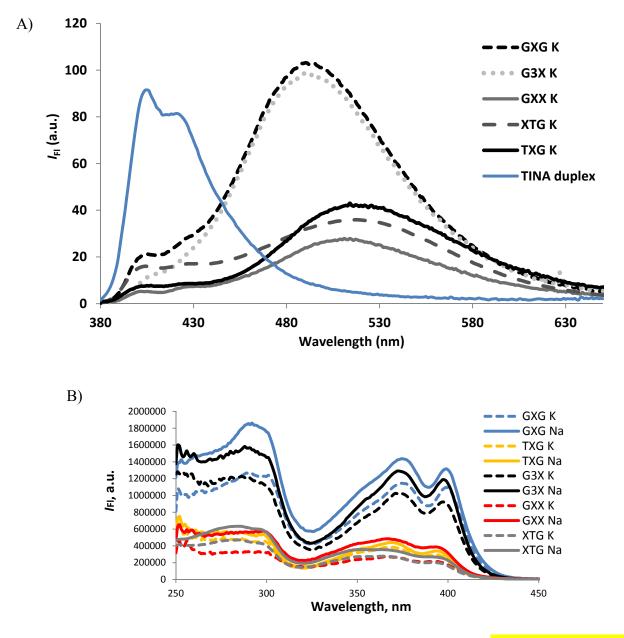
**Figure S5.** Examples of isothermal renaturation processes recorded by UV-Vis and CD spectroscopy of **GXG** (20  $\mu$ M strand concentration in A, 10  $\mu$ M in B), **G3X** (10  $\mu$ M in C, 10  $\mu$ M in D) and **TG<sub>6</sub>T** (20  $\mu$ M in E) in the presence of 110 mM NaCl in 10 mM Li cacodylate buffer at pH 7.2, 20 °C. Experimental data – dotted lines, fitted curves – solid lines. The order (n) of the reaction

in separate strands was initially floating and the data could be fitted with n = 3.97, we defined n = 4 for determination of  $k_{on}$  Table 1.

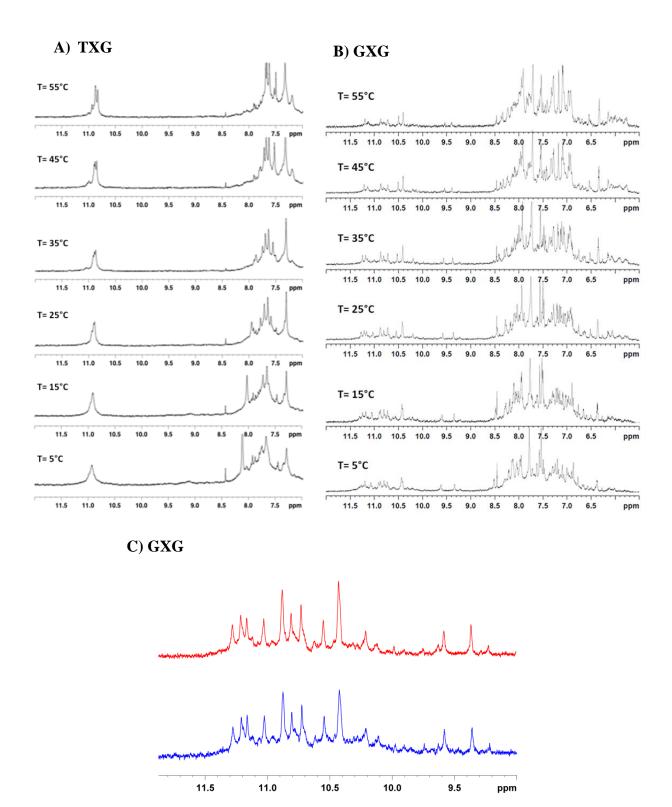
Abbreviation	Sequence	$k_{\rm on}$ / UV-Vis	$k_{\rm on}$ / CD
GXG	dTGG <b>X</b> GGT	$2.73 \times 10^{11}$	$7.30 \times 10^{11}$
G3X	dTGGGXGGGT	$2.28 \times 10^{11}$	$9.38 \times 10^{11}$
TG <sub>6</sub> T	dTGGGGGGT	$\sim 1.0 \times 10^{10}$	$8.90\times10^{10}$

**Table S4.**  $k_{on}$  values determined using UV-Vis and CD.<sup>*a*</sup>

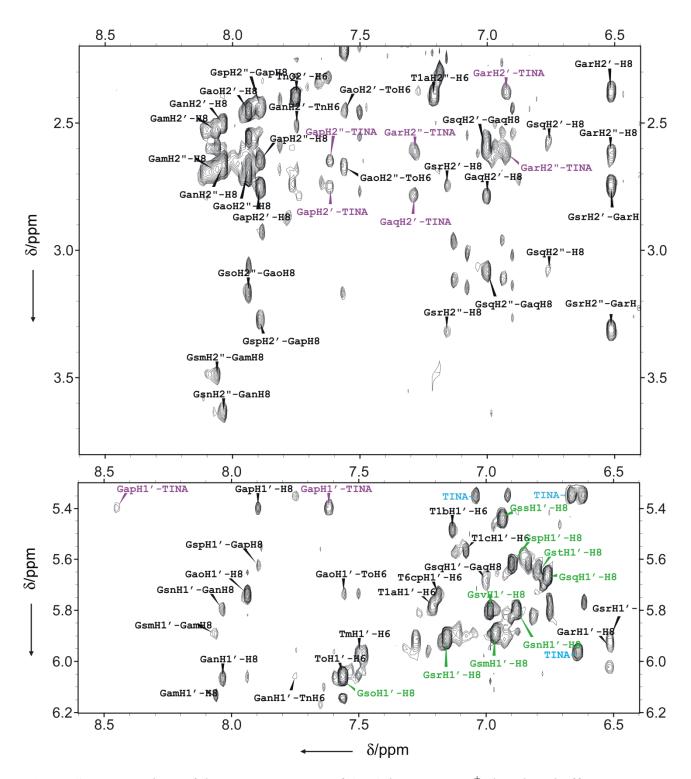
<sup>a</sup>  $k_{on}$  values (± 30%) of selected assemblies were experimentally obtained by monitoring ellipticities at 265 (**G3X**, TG<sub>6</sub>T) and 295 nm (**GXG**) using CD spectrometer or by monitoring absorbance at 275 (**GXG**) and 280 nm (**G3X**) using UV-Vis spectrometer in 10 mM Li cacodylate buffer and 110 mM NaCl at 20 °C, pH 7.2.  $k_{on}$  value for TG<sub>6</sub>T (UV-Vis) at 20 °C was estimated from the literature.



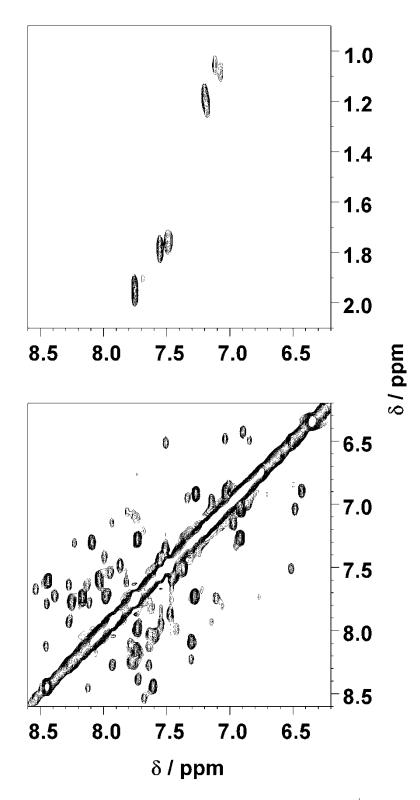
**Figure S6.** A) Fluorescence emission spectra of TINA-TG<sub>4(6)</sub>T assemblies and TINA duplex (5'dAGCTTGXCTTGAG/5'-dCTCAAGXCAAGCT) at 10  $\mu$ M strand concentration in the presence of 110 mM KCl in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.  $\lambda_{ex}$  = 373 nm. Poor overlap of TINA moieties in the middle of the duplex does not result in excimer formation and only monomeric fluorescence is observed (12). B) Fluorescence excitation spectra at 10  $\mu$ M strand concentration in the presence of 110 mM NaCl or 110 mM KCl in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.  $\lambda_{em}$  = 495 nm.



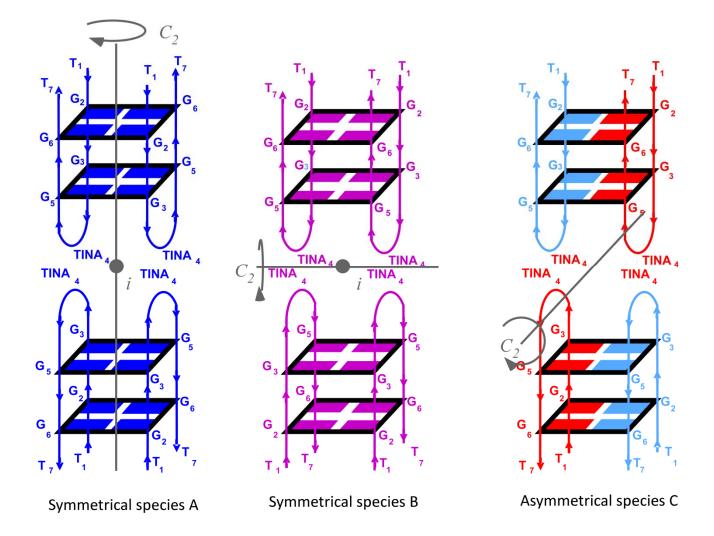
**Figure S7.** NMR spectra of **TXG** (A) and **GXG** (B) showing imino and aromatic regions at 500  $\mu$ M strand concentration in 10 mM Na<sup>+</sup> phosphate buffer supplemented with 10 mM KCl (A) or 100 mM NaCl (B), 10% D<sub>2</sub>O, pH 7.0. (C) NMR spectra of **GXG** showing imino region at 500 (red) and 100  $\mu$ M (blue) strand concentrations in 10 mM Na<sup>+</sup> phosphate buffer and 100 mM NaCl, 10% D<sub>2</sub>O, pH 7.0.



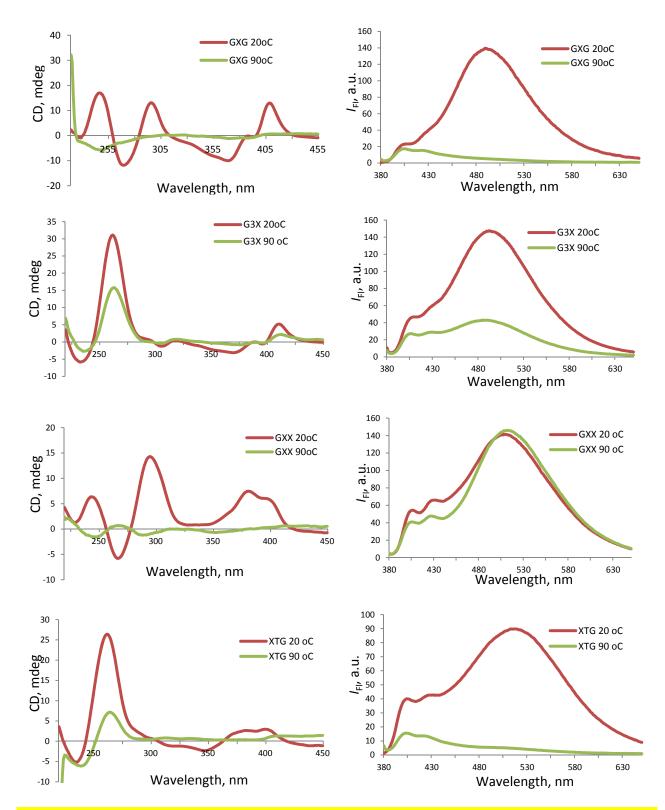
**Figure S8.** Two regions of the NOESY spectra of **GXG** in 10 mM Na<sup>+</sup> phosphate buffer, 10% D<sub>2</sub>O, pH 7.0. NOE cross-peaks are labelled according to partial assignments. Ga and Gs stand for *anti*-and *syn*-guanines, respectively. Strong intra-residual H8-H1' cross-peaks, characteristic of syn-Gs, are labelled in green (bottom panel). Three tracts of Gs $\rightarrow$ Ga $\rightarrow$ T were found and are labelled as m, n, and o. Three Gs $\rightarrow$ Ga $\rightarrow$ TINA protons are labelled as p, q, and r. The three remaining Gs, not showing sequential cross-peaks, are labelled as s, t, and u. DNA-TINA cross-peaks are shown in magenta, and TINA-TINA cross-peaks in cyan. TINA protons could be identified in the TOCSY spectra (see Fig S9), but unambiguous assignment of specific protons was not possible.



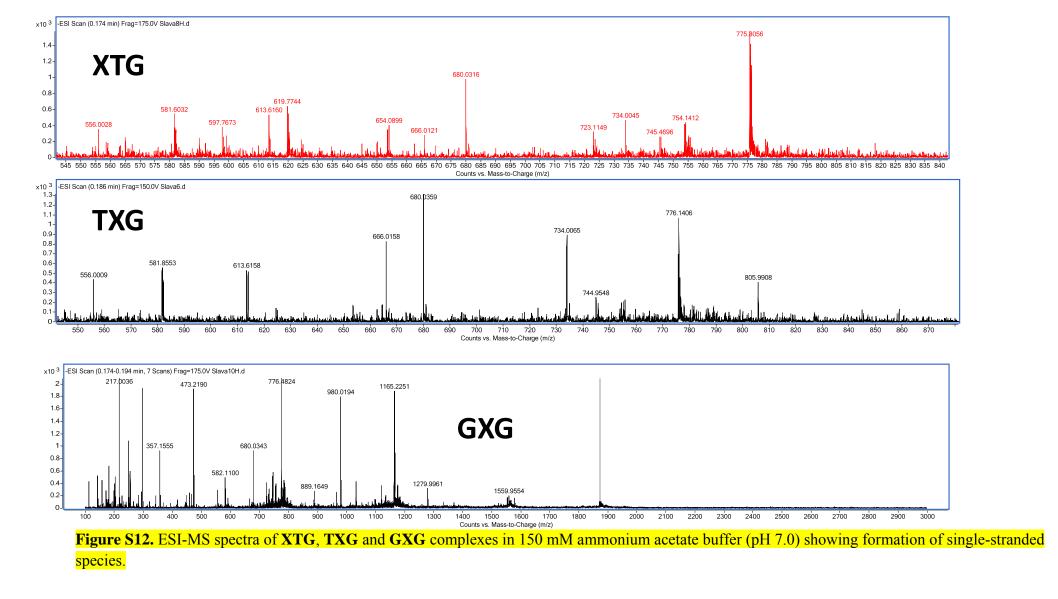
**Figure S9.** Two regions of the IP-COSY spectrum of **GXG** in Na<sup>+</sup> phosphate buffer in 90%  $H_2O/D_2O$ . Top: Methyl-H6 regions, showing at least six thymines. Bottom: Multiple cross-peaks between TINA protons. The number of signals indicates the presence of several TINA species.



**Figure S10.** Possible configurations of dimers formed by TINA mediated interaction of two antiparallel quadruplex subunits. Species A and B have both a  $C_2$  symmetry axis and a centre of inversion. The four hairpins of each structure are magnetically equivalent, but those of A are not equivalent to those in B. In structure C the hairpins are no longer all equivalent. Each subunit is formed from a pair of magnetically non-equivalent hairpins which are related *between* the subunits by a  $C_2$  axis. Excluding coincidental overlap, these structures could account for six of the thymine Methyl-H6 correlations in **Figure S9**. A further species with no symmetry and generating additional NMR signals could be formed by the dimerization of subunits from species A and B/C.



**Figure S11.** CD (left) and fluorescence emission (right) spectra of TINA-TG<sub>4(6)</sub>T at 20 and 90 °C at 10  $\mu$ M strand concentration in the presence of 110 mM NaCl in 10 mM Li cacodylate buffer at pH 7.2,  $\lambda_{ex} = 373$  nm. It should be noted that excimer is still observed for **GXX** at 90 °C due to the presence of two TINA moieties in the sequence. In other constructs excimer is significantly reduced (**G3X**) or disappears after incubation at 90 °C for 10 min.



#### References

- . [1] Xia, Y., Legge, G., Jun, K. Y., Qi, Y., Lee, H., & Gao, X. *Magn Reson Chem* 2005, 43(5), 372-379.
- . [2] Piotto, M.; Saudek, V.; Sklenar, V. J. Biomol. NMR 1992, 2, 661
- [3] Hwang, T. L., & Shaka, A. J. 1995, *Journal of Magnetic Resonance, Series A, 112*, 275-279
- . [4] Goddard, D. T.; Kneller, G.; Program Sparky 3. ed.; University of California: San Francisco
- [5] D. A. P. David A Case, James W Caldwell, Thomas E Cheatham III, Junmei Wang, Wilson S Ross, CL Simmerling, TA Darden, Kenneth M Merz, Robert V Stanton, AL Cheng, James J Vincent, Mike Crowley, Vickie Tsui, Holger Gohlke, RJ Radmer, Yong Duan, Jed Pitera, Irina Massova, George L Seibel, U Chandra Singh, PK Weiner, Peter A Kollman, University of California, San Francisco, 2007.
- . [6] Kettani, et al. Nat Struct Biol. 1997, 4, 382-389
- . [7] T. Darden, D. York, L. Pedersen, J. Chem. Phys. 1993, 98, 10089-10092.
- [8] a) T. E. Cheatham, P. Cieplak, P. A. Kollman, *J. Biomol. Struct. Dyn.* 1999, 16, 845-862; b) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* 1996, 118, 2309-2309.
- . [9] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* 1983, 79, 926-935.
- . [10] G. C. Shields, C. A. Laughton, M. Orozco, J. Am. Chem. Soc. 1997, 119, 7463-7469.
- . [11] Koradi, R.; Billeter, M.; Wüthrich, K. J. Mol. Graphics 1996, 14, 29
- . [12] V. V. Filichev, I. V. Astakhova, A. D. Malakhov, V. A. Korshun, E. B. Pedersen, *Chem.—Eur. J.* 2008, *14*, 9968-9980.