Article

Formation of an Interlocked Quadruplex Dimer by d(GGGT)

Yamuna Krishnan-Ghosh, Dongsheng Liu, and Shankar Balasubramanian

J. Am. Chem. Soc., **2004**, 126 (35), 11009-11016• DOI: 10.1021/ja049259y • Publication Date (Web): 17 August 2004

J|A|C|S
ARTICLES
Published on Web 08/17/2004

Formation of an Interlocked Quadruplex Dimer by d(GGGT)

Yamuna Krishnan-Ghosh, Dongsheng Liu, and Shankar Balasubramanian*

Contribution from the University Chemical Laboratories, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

Received February 11, 2004; E-mail: sb10031@cam.ac.uk

Abstract: A tetranucleotide sequence d(GGGT) has been shown to self-assemble into an interlocking quadruplex dimer. UV-melting studies indicated the existence of two species that each showed distinct quadruplex melting transitions, a low-T_m species, $\mathbf{Q_l}$, and a high-T_m species, $\mathbf{Q_h}$. Conditions were controlled to favor the formation of either $\mathbf{Q_l}$ or $\mathbf{Q_h}$. $\mathbf{Q_l}$ and $\mathbf{Q_h}$ each showed circular dichroism spectra characteristic of parallel quadruplexes. Negative ion nano-electrospray ionization mass spectrometry confirmed that $\mathbf{Q_l}$ was a tetrameric complex, $d(GGGT)_4$, and $\mathbf{Q_h}$ was an octameric complex, $d(GGGT)_8$. High-resolution ^{1}H NMR spectroscopy evidenced that $d(GGGT)_4$ was a C_4-symmetric parallel tetramolecular quadruplex. The ^{1}H NMR spectrum of $d(GGGT)_8$ was consistent with a structure formed by the dimerization of a parallel, "slipped" tetramolecular quadruplex that has its diagonal strands staggered by one base. This "slippage" results in two guanine bases at the 5′ end of the quadruplex being presented diagonally that are not involved in tetrads. Two such "slipped" quadruplexes dimerize via these free G-bases at the 5′ ends by forming an extra G-tetrad. Each "slipped" quadruplex contributes two guanine bases to this extra G-tetrad. The formation of a novel GTGT tetrad is also observed at both the 3′ ends of the interlocked quadruplex dimer.

Introduction

G-quadruplexes are four-stranded structures formed from DNA sequences that contain stretches of contiguous guanine bases, usually in the presence of alkali metal cations.[1] G-quadruplex DNA has been in focus recently because of its putative existence and biological function at telomeres, which affects replication,[2] and in promoter sites influencing transcription.[3] The guanine bases in G-quadruplexes form G-tetrads where four G bases hydrogen bond via the Hoogsteen sites in a cyclic fashion.[4] Quadruplexes can adopt several morphologies that include parallel or antiparallel configurations and may be intermolecular[5] or intramolecular.[6] Parallel quadruplex-forming sequences that have a stretch of consecutive guanines at their 3′ or 5′ ends have been shown to form extended quadruplexes by associating into higher order structures.[7] However, it is not known how such higher order structures are arranged at the molecular level. This article addresses the molecular arrangement of strands in higher order structures formed from the tetranucleotide sequence d(GGGT) (**G3T**).

Experimental Section

Materials. Oligonucleotides were obtained from Invitrogen (UK). All chemicals used were obtained from Sigma and were of analytical grade. Milli-Q water was used for all experiments.

UV Spectrophotometry. All UV measurements were carried out on a Varian Cary 1E UV/visible spectrophotometer. Samples were prepared by diluting the oligonucleotide from a 1 mM stock solution into 100 mM K phosphate buffer, pH 7.4, to give the appropriate strand concentration. For samples comprising only low-T_m species, 50 μM G3T in phosphate buffer was heated to 90 °C for 10 min, followed by slow cooling over 3 h to room temperature and equilibration for 24 h at 4 °C. Samples comprising only high-T_m species were prepared by heating 1 mM G3T in phosphate buffer at 55 °C for 24 h. This was followed by cooling to 30 °C and reheating to 55 °C three times at a rate of 0.5 °C/min. The sample was then cooled to 5 °C over 2 h and equilibrated for 24 h at 5 °C. UV-melting studies were carried out from 5 °C to 95 °C at a heating rate of 0.5 °C/min.

CD Spectroscopy. CD spectra were recorded on a JASCO J-810 circular dichroism spectrometer. A 200 μM concentration of oligonucleotide in 100 mM K^+ phosphate buffer, pH 7.4, was heated to 90 °C for 10 min, slowly cooled to room temperature over 3 h, and equilibrated at 4 °C for 24 h. Spectra were recorded between 350 and 220 nm and have been presented as the average of five successive runs. All spectra are subtracted from a baseline corresponding to buffer alone.

Mass Spectrometry. Nanospray ionization mass spectrometry was carried out on a Q-TOF-1 mass spectrometer (Micromass, Manchester, UK) at a source temperature of 35 °C with a cone voltage of 30 V and an analyzer pressure of 1×10^{-5} Torr. The low-T_m species ($\mathbf{Q_l}$) was prepared from 100 μM G3T in 100 mM K^+ phosphate buffer, pH 7.4,

(1) Gellert, M.; Lipsett, M. N.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1962**, *48*, 2013−2018.
(2) Hurley, L. H. *Nat. Rev. Cancer* **2002**, *2*, 188−200. Neidle, S.; Parkinson, G. *Nat. Rev. Drug. Discov.* **2002**, *1*, 383−393.
(3) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593−11598.
(4) (a) Williamson, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 703−730. (b) Guschlbauer W.; Chantot, J.-F.; Thiele, D. *J. Biomol Struct. Dyn.* **1990**, *8*, 491−511.
(5) (a) Wang, Y.; Patel, D. J. *Biochemistry* **1992**, *35*, 8112−8119. (b) Haider, S.; Parkinson, G. N.; Neidle, S. *J. Mol. Biol.* **2002**, *320*, 189−200.
(6) Smith, F. W.; Feigon, J. *Nature* **1992**, *356*, 164−168. Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876−880.
(7) Sen, D.; Gilbert, W. *Biochemistry*, **1992**, *31*, 65−70. Poon, K.; Macgregor, R. B. *Biopolymers* **1998**, *45*, 427−434. Chen, F. M. *Biophys. J.* **1997**, *73*, 348−356. Guo, Q.; Lu, M.; Kallenbach, N. R. *Biochemistry* **1993**, *32*, 3596−3603. Marsh T. C.; Henderson, E. *Biochemistry* **1994**, *33*, 10718−10724.

by heating to 90 °C for 10 min, followed by slow cooling over 3 h to room temperature and equilibration for 24 h at 4 °C. A 20-fold dilution was carried out just prior to injection in order to visualize $\mathbf{Q_l}$. For the high-T_m species ($\mathbf{Q_h}$), 400 μM G3T in 100 mM phosphate buffer, pH 7.4, was heated at 55 °C for 24 h and slowly cooled to room temperature. This was followed by cooling to 30 °C and reheating to 55 °C at a rate of 0.5 °C/min. This was repeated three times. The sample was then cooled to 5 °C over 2 h and equilibrated for 24 h at 5 °C.This sample was diluted 40-fold just prior to injection to investigate $\mathbf{Q_h}$ by nano-ESI-MS.

^{1}H NMR Spectroscopy. The low-T_m sample ($\mathbf{Q_l}$) was prepared from 100 μM solution of G3T by annealing from 90 °C to 4 °C over 4 h, followed by equilibration for 12 h at 4 °C. The high-T_m sample ($\mathbf{Q_h}$) was prepared from a 5 mM solution of **G3T** by heating the sample for 24 h at 55 °C in 100 mM phosphate buffer, pH 6.8. This was followed by cooling to 30 °C, reheating to 55 °C at a rate of 0.5 °C/min, and repeating three times. ^{1}H NMR spectra were recorded on a Bruker 500 AMX spectrometer at 12 °C. D_2O exchange studies were carried out by lyophilizing the above samples and resuspending in D_2O.

Results and Discussion

Design Considerations. The oligonucleotide d(GGGT), **G3T**, incorporated a stretch of guanines at the 5′ terminus which is a prerequisite for the formation of higher order quadruplex structures. The sequence was chosen to be a tetranucleotide in order to rule out the structural possibilities associated with intramolecular folding of longer oligonucleotides. A T residue was included at the 3′ end in order to prevent higher order structure formation at the 3′ end, and also to desymmetrize the sequence to facilitate structural characterization by NMR spectroscopy.

G3T Forms Two Kinds of Quadruplexes, $\mathbf{Q_l}$ and $\mathbf{Q_h}$, under Different Conditions. DNA duplexes show a positively sloped sigmoidal curve when the UV absorbance at 260 nm (A_{260}) is plotted as a function of temperature (T). On the other hand, quadruplex DNA melting is characterized by an *inverse* sigmoidal thermal melting curve observed at 295 nm.[8] Intermolecular quadruplexes are generally formed at strand concentrations on the order of 10^{-4} M, at which the signal at 295 nm is saturating. In such cases, quadruplex melting is usually followed at 305 nm.[8] The ability of **G3T** to self-assemble into quadruplexes at different strand concentrations was investigated by UV-melting experiments monitored at 305 nm. A given concentration of **G3T** was self-assembled in 100 mM potassium phosphate buffer, pH 7.4, by heating to 90 °C for 10 min, followed by cooling to 4 °C at a rate of 0.5 °C/min and equilibration for 24 h at 4 °C.[9] Figure 1A shows the first derivative of the melting curves of **G3T** at different concentrations, where each transition in the melting curve appears as a peak in the corresponding first derivative plot. The melting profile of 200 μM **G3T** in 100 mM KH_2PO_4 buffer (pH 7.4) showed two distinct melting transitions: a lower melting species, which we have designated $\mathbf{Q_l}$, with a T_m of 38.5 ± 1 °C and a high melting species, which we have designated $\mathbf{Q_h}$, with a T_m of 79 ± 1 °C (Figure 1A). Conditions for self-assembly could be adjusted such that either $\mathbf{Q_h}$ or $\mathbf{Q_l}$ could be predominantly formed from **G3T**. When the strand concentration was increased to 400 μM and equilibrated as described in the Experimental Section, the UV-melting profile showed a preponderance of $\mathbf{Q_h}$ (Figure 1A). UV-melting at 200 μM and 100 μM **G3T** revealed that decreasing the concentration of **G3T** favored the formation of $\mathbf{Q_l}$ at the expense of $\mathbf{Q_h}$. A comparison of the peak areas of $\mathbf{Q_l}$ relative to $\mathbf{Q_h}$ for each melting profile indicates that $\mathbf{Q_h}$ is much more sensitive than $\mathbf{Q_l}$ to decreased strand concentration. This suggests that $\mathbf{Q_h}$ has a higher molecularity than the $\mathbf{Q_l}$.[10] To assess the importance of the sequence for the formation of the higher T_m species, $\mathbf{Q_h}$, the 5′ end of **G3T** was "blocked" with a T residue to give d(TGGGT). A 200 μM concentration of d(TGGGT) in 100 mM phosphate buffer was allowed to self-assemble and its UV-melting profile recorded. In this case, only a single melting transition with a T_m of 52 ± 1 °C was observed, indicative of a single self-assembled species (Figure 1A). This suggests that the terminal G nucleobase in **G3T** is required for one of the two species formed.

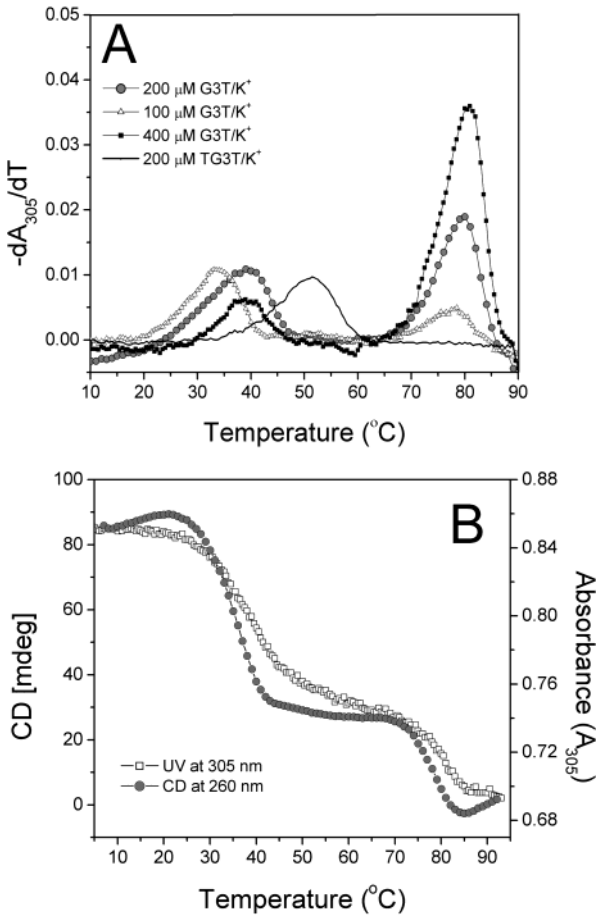


Figure 1. (A) Plots of the first derivative ($-dA_{305}/dT$) versus temperature (T) of temperature-dependent UV absorbance at 305 nm (A_{305}) of **G3T** and **TG3T** at various concentrations. (B) Comparison of the melting profile of 200 μM **G3T** in 100 mM K^+ phosphate buffer, pH 7.4, by UV spectrophotometry at 305 nm (A_{305}) and circular dichroism (mdeg) at 260 nm.

$\mathbf{Q_l}$ and $\mathbf{Q_h}$ Are Parallel Quadruplexes. To further characterize $\mathbf{Q_l}$ and $\mathbf{Q_h}$, circular dichroism (CD) spectra were recorded at various temperatures. A 200 μM concentration of **G3T** in 100 mM phosphate buffer, pH 7.4, was investigated by both CD and UV as a function of temperature. The UV-melting profile of 200 μM **G3T** at 305 nm (Figure 1B) showed two

(8) Mergny, J.-L.; Phan, A.-T.; Lacroix, L. *FEBS Lett.* **1998**, *435*, 74−78.
(9) Circular dichroism has been shown to be supportive, albeit not confirmatory, evidence of quadruplex structure. See: Dapić, V.; Abdomerović, V.; Marrington, R.; Peberdy, J.; Rodger, A.; Trent, J. O.; Bates, P. J. *Nucleic Acids Res.* **2003**, *31*, 2097−2107.
(10) $d[(\mathbf{G3T})_n]/dT = k[(\mathbf{G3T})]^n$, where n is the molecularity and k is the rate constant of formation of the complex $(\mathbf{G3T})_n$. The higher the value of n, the greater the effect of concentration on complex formation.

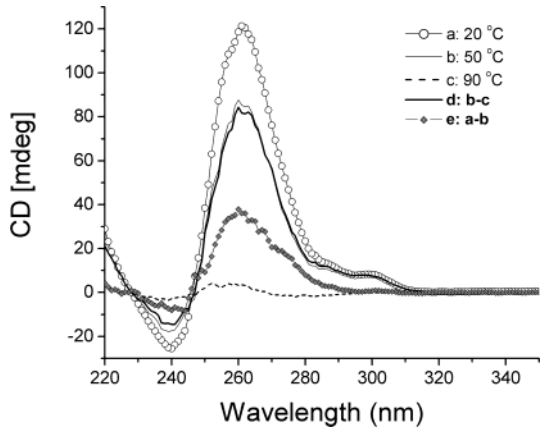


Figure 2. Circular dichroism of 200 μM **G3T** in 100 mM K^+phosphate buffer, pH 7.4, of (a) low-T_m species (**Q_l**) and high-T_m species (**Q_h**) at 20 °C; (b) **Q_h** and single-stranded **G3T** at 50 °C; (c) single-stranded **G3T** at 90 °C; (d) **Q_l**; and (e) **Q_h**.

melting transitions due to **Q_l** at 38.5 ± 1 °C and **Q_h** at 79 ± 1 °C. The CD trace of this sample at 20 °C showed a positive maximum at 260 nm (Figure 2). This was therefore the preferred wavelength at which temperature-induced changes in CD were monitored. Temperature-dependent CD on this sample revealed two melting transitions at 37 ± 1 °C and 77.5 ± 1 °C (Figure 1B), comparable to the T_m's observed in the UV-melting profile at 200 μM G3T. Moreover, the relative contributions of high-T_m species and low-T_m species to the overall transition are similar to those observed by UV. We therefore propose that the high- and low-T_m species observed by CD are the same high- and low-T_m species as observed in the UV melt. Figure 2 shows the CD trace of 200 μM **G3T** at 20, 50, and 90 °C. At 20 °C, **G3T** strands are self-assembled, giving a mixture of two complexes, **Q_l** and **Q_h**. The CD trace, Figure 2a, at this temperature is a result of both these species. At 50 °C, **Q_l** has melted into single strands, and therefore the solution is now a mixture of **Q_h** and single strands of **G3T**. The CD trace, Figure 2b, at 50 °C thus has contributions from **Q_h** and single-stranded DNA from melted **Q_l**. At 90 °C, **Q_h** also melts into single strands, and the trace at this temperature, Figure 2c, results from only single strands of **G3T**. Accordingly, trace c in Figure 2 is consistent with single-stranded DNA, showing negligible CD compared to traces a and b. That trace b is indistinguishable from its difference spectrum[11] (b − c), represented by trace d in Figure 2, suggests the single-stranded DNA component makes a negligible contribution to the CD spectrum. Thus, trace b may be considered to result only from **Q_h**. This shows a positive maximum at 260 nm (Figure 2) which is characteristic of parallel quadruplexes.[12] Hence the difference spectrum[11] of trace (a − b), designated trace e, represents the CD spectrum of **Q_l**. Trace e also shows a positive maximum at 260 nm indicative of it, also, being a parallel quadruplex.[12]

Q_l Is Tetrameric, $(G3T)_4$, while Q_h Is Octameric, $(G3T)_8$. To establish the molecularity of quadruplexes **Q_l** and **Q_h**, they

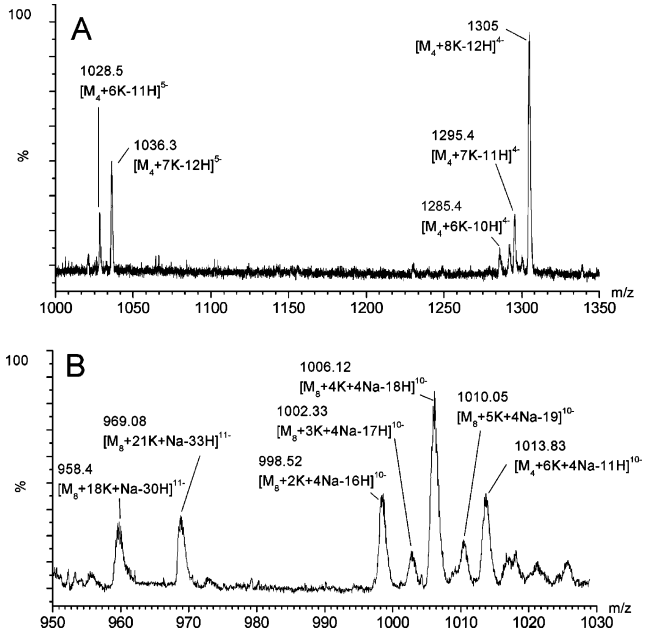


Figure 3. Partial nano-ESI-MS spectrum showing multiply charged states of (A) the low-T_m species, **Q_l**, and (B) at an expanded scale the high-T_m species, **Q_h**.

were subjected to negative ion nano-electrospray ionization mass spectrometry (nano-ESI-MS) under conditions that preserve non-covalent interactions. ESI-MS has been used previously to characterize intermolecular DNA quadruplexes.[13] Analysis of a solution of 100 μM **G3T** in 100 mM K^+ phosphate buffer, pH 7.4, shown to consist mainly of **Q_l** by UV analysis (Figure 1A), showed broad peaks centered at *m*/*z* 1285.4, 1295.4, and 1305 (Figure 3A). The spacing between these peaks ($\Delta m/z$) corresponded to 9.8 ± 0.2 units, indicating that these peaks were due to a 4^- charged species corresponding to $[M_4 + 6K - 10H]^{4-}$, $[M_4 + 7K - 11H]^{4-}$, and $[M_4 + 8K - 12H]^{4-}$, respectively, where M is the molecular weight of **G3T** (1228.5 Da). Broad peaks centered at *m*/*z* 1028 and 1036, corresponding to $[M_4 + 6K - 11H]^{5-}$ and $[M_4 + 7K - 12H]^{5-}$, respectively, were also observed (Figure 3A), which confirmed that **Q_l** is a tetramolecular quadruplex, **$(G3T)_4$**.

Analysis of a solution of 400 μM **G3T** in 100 mM K^+ phosphate buffer, shown to consist mainly of **Q_h** by UV analysis (Figure 1A), gave broad peaks centered at *m*/*z* 998.5, 1002.3, 1006.12, 1010, and 1013.8 (Figure 3B). The spacing between these peaks ($\Delta m/z$) corresponded to 3.8 ± 0.2 units, indicating that these peaks were due to a 10^- charged species. The observed peaks correspond to $[M_8 + 2K + 4Na - 16H]^{10-}$, $[M_8 + 3K + 4Na - 17H]^{10-}$, $[M_8 + 4K + 4Na - 18H]^{10-}$, and $[M_8 + 5K + 4Na - 19H]^{10-}$. Similar peaks centered at *m*/*z* 958.4 and 969.0, corresponding to $[M_8 + 18K + Na - 30H]^{11-}$ and $[M_8 + 21K + Na - 33H]^{11-}$, were also observed.[14] This confirms that **Q_h** is a quadruplex with eight **G3T** strands, i.e., **$(G3T)_8$**.

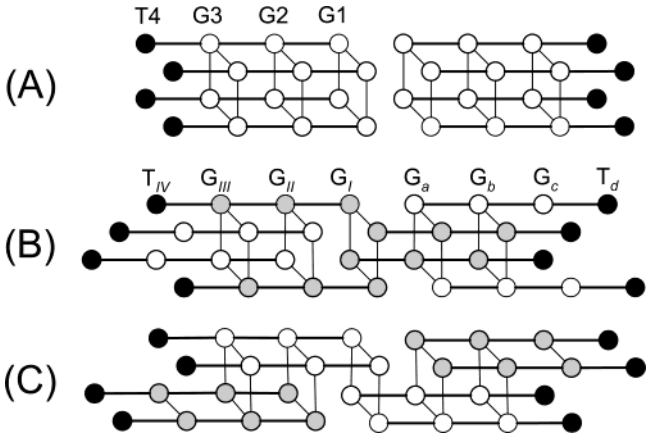
(11) These difference spectra do not take into account differential amounts of single-stranded (ss) DNA at each temperature. Since the CD of both high- and low-T_m species is so much larger than that of ss-DNA, it is unlikely that the different amounts of ss-DNA will affect the gross appearance of the trace. A direct comparison can be made of the trace at 50 °C and the difference spectrum of the traces at 50 °C and 90 °C.

(12) Hardin, C. C.; Perry, A. G.; White, K. *Biopolymers* **2001**, *56*, 147−194.

(13) Goodlett, D. R.; Camp, D. G.; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, *22*, 181−183.

(14) Na^+/K^+ exchange in quadruplexes is fairly fast ($\sim 10^{-3}$s) and occurs without much disruption of the overall quadruplex structure. See: (a) Hud, N. V.; Smith, F. W.; Anet, F. A. L.; Feigon, J. *Biochemistry* **1996**, *35*, 15383−15390. (b) Deng, H.; Braunlin, W. H. *J. Mol. Biol.* **1996**, *255*, 476−483. Sodiated species are likely to be formed by exchange of K^+ ions in the sample with Na^+ ions from the instrument.

Chart 1. Possible Arrangements of Strands in $(G3T)_8$: "Blunt-End" Stacked Model with D_{4h}-Type Symmetry (A) and "Interlocking" Models with C_{2v}-Type Symmetry (B) and C_{2h}-Type Symmetry (C)[a]



[a] Thymines are represented by black circles and guanines on slipped strands by gray circles; potassium ions have been omitted for clarity.

Possible Structural Models of $(G3T)_8$. We considered the likely structures that could be adopted by eight **G3T** strands while preserving quadruplex character. The octamer **$(G3T)_8$** could be formed by a blunt-end stacking of two tetramers **$(G3T)_4$** (Chart 1A) to give a semi-continuous quadruplex with D_4-type symmetry. Blunt-end stacking of terminal G-tetrads that are part of different quadruplexes has been observed in the crystalline quadruplexes.[15] Alternatively, if one or two strands are staggered by one or more guanine bases, this could result in an interlocked but extended quadruplex formed by the dimerization of two staggered or "slipped" quadruplexes.[16] Dimerization of two complementary slipped quadruplexes at the 5′ ends would result in a continuous interlocking structure. Chart 1B,C shows interlocking structures arising from slippage of diagonal and adjacent strands, respectively.

^{1}H NMR Spectroscopy on $(G3T)_4$ and $(G3T)_8$. The structures of **$(G3T)_4$** and **$(G3T)_8$** were investigated in detail by high-resolution ^{1}H NMR spectroscopy.

(a) $(G3T)_4$ Is a Parallel Quadruplex. ^{1}H NMR of quadruplexes requires strand concentrations in the millimolar regime in order to clearly observe the imino proton resonances. At these concentrations, **G3T** forms the octamer, **$(G3T)_8$**, rather than the tetramer, **$(G3T)_4$**. Measurements carried out on samples of **G3T** that were predominantly in the form of **$(G3T)_4$** could only be done at a maximum strand concentration of 100 μM. Under these conditions, the hydrogen-bonded guanine imino proton resonances could not be observed. The 1D ^{1}H NMR spectrum of 100 μM **G3T** in 100 mM phosphate buffer (pH 7) in the region 7−8 δ ppm shows the H8 protons on the guanines and H6 protons on the thymine (Figure 4A). The spectrum shows a set of minor peaks, indicated by #, corresponding to a minor population of octamer and a set of major peaks corresponding to the tetramer, **$(G3T)_4$**. The H6 of the thymine could be easily identified by its strong NOE with the $-CH_3$ protons at 1.65 δ ppm (not shown). The peaks in the region 7−8 δ ppm may be assigned by following the observed NOEs between the base

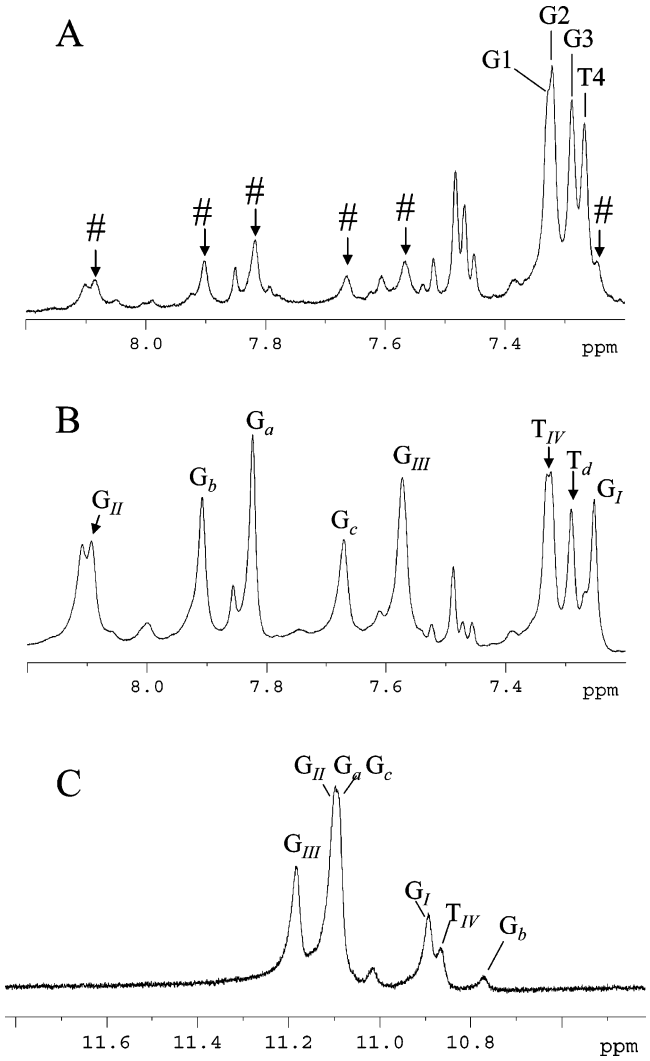


Figure 4. Portion of the 1D ^{1}H NMR spectrum in water at neutral pH, 12 °C, showing aromatic protons of (A) 100 μM **G3T** [guanine H6 assignments correspond to the guanine labels in Chart 1A; peaks indicated by # correspond to a minor population of **$(G3T)_8$**]; (B) 5 mM **G3T**, 100 mM K^+ phosphate buffer [guanine H6 assignments correspond to the guanine labels shown in Chart 1B]; and (C) imino proton (NH) region of 5 mM **G3T**.

protons (H8 on G and H6 on T) and their own and flanking sugar H1′ protons in **$(G3T)_4$** (Supporting Information). The magnitude of these NOEs at short mixing times (50 ms, data not shown) suggests that the nucleotides adopt anti glycosidic orientations along the length of the quadruplex.[5a] From the 1D spectrum we see that there are only three H8 protons, corresponding to G1, G2, and G3, and one H6, corresponding to T4. This implies that all the strands in **$(G3T)_4$** are equivalent. The only way strands from an asymmetric sequence can associate in a tetramolecular complex that maintains strand equivalence is a parallel-stranded structure with a 4-fold rotational symmetry. Thus **$(G3T)_4$** is a parallel, tetramolecular quadruplex. This is in line with previous studies on short (5−7 bases), quadruplex-forming oligonucleotides where all the gycosidic bonds were found to be in the anti conformation.[5a,17] Without exception, these sequences formed parallel quadruplexes with 4-fold rotational symmetry.

(15) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876−880. Laughlan, G.; Murchie A. I.; Norman, D. G.; Moore, M. H.; Moody, P. C.; Lilley, D. M.; Luisi, B. *Science* **1994**, *265*, 520−524.

(16) Spackova, N.; Berger, I.; Sponer J. *J. Am. Chem. Soc.* **1999**, *121*, 5519−5534.

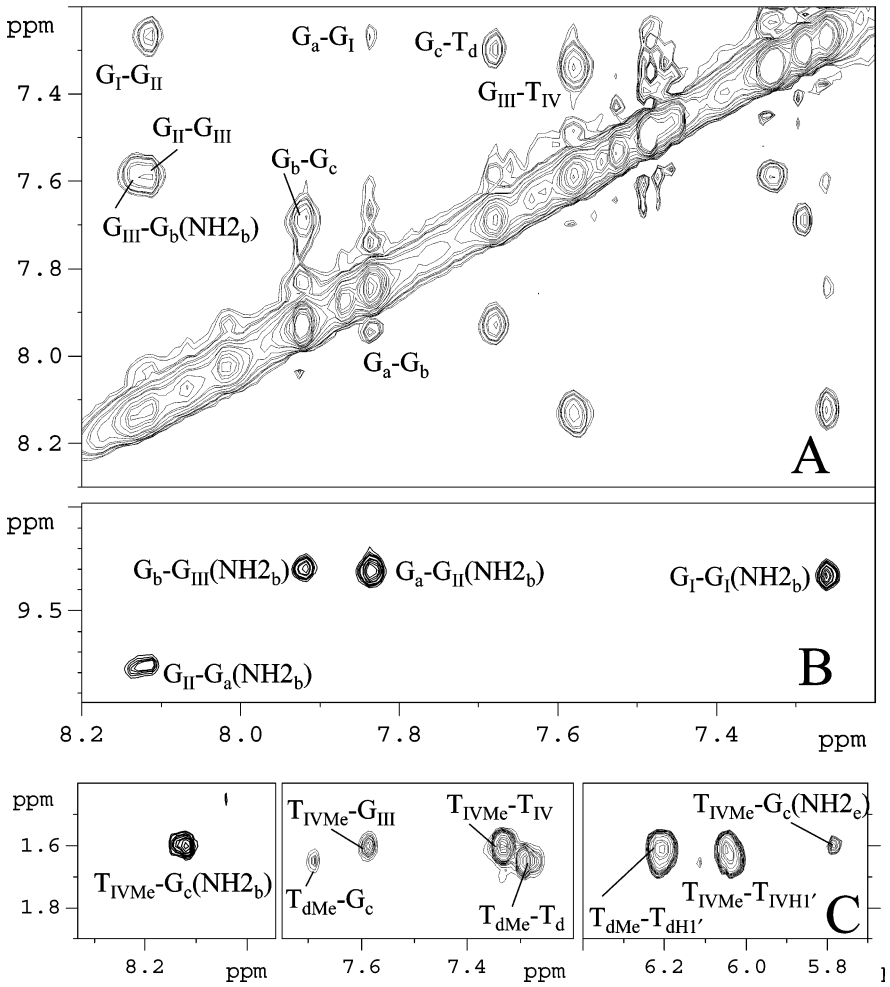


Figure 5. Portion of the NOESY spectrum, pH 6.7, 12 °C, 5 mM **G3T**, 10% D_2O/H_2O, showing (A) sequential H8−H8 NOEs, (B) H8−NH_{2b} NOEs, and (C) T-CH_3 NOEs in **$(G3T)_8$**.

(b) $(G3T)_8$ Is an Extended Quadruplex with Interlocking Strands. Spectral measurements on samples of **G3T** that were predominantly in the form of **$(G3T)_8$** were done at strand concentrations of 5 mM.

(i) $(G3T)_8$ Is Composed of Two Kinds of G3T Strands. The 1D ^{1}H NMR spectrum of **$(G3T)_8$** (Figure 4B) showed two distinct T H6's (T_d = 7.34 and T_{IV} = 7.29 δ ppm) that were identified by their strong NOEs with the two CH_3 signals at 1.5−1.7 δ ppm (Figure 5C, Q,Q′). Sequential NOE connectivities in the region 7−8.5 δ ppm (Figure 5A) established that T_{IV} H6 was serially connected to three H8 protons designated G_{III}, G_{II}, and G_I, respectively. NOE connectivities also revealed that T_d H6 was serially connected as well to three H8 protons designated G_c, G_b, and G_a, respectively. Thus, the 1D spectrum showed two thymine H6 protons and six guanine H8 protons (Figure 4B), while the NOE data (Figure 5A) established that each T H6 was serially connected to three G H8 protons. This implies that there are two kinds of **G3T** strands in the **$(G3T)_8$** quadruplex. These data rule out the blunt-end stacked quadruplex model, Chart 1A, in which all the **G3T** strands in this structure would be equivalent.

Importantly, while tracing the chain connectivities T_{IV}-G_I and T_d-G_a in Figure 5A, we observed a cross-peak D between G_I H8 and G_a H8, indicating that the two nonequivalent strands that make up **$(G3T)_8$** are arranged with their 5′ ends adjacent to each other.

(ii) All the Glycosidic Bonds in $(G3T)_8$ Are Anti. NOEs in the region shown in Figure 6 correspond to contacts between the guanine H8 protons and (i) their own sugar H1′ protons as well as (ii) their 5′-flanking sugar H1′ protons. The NOEs between the base H8 and its own sugar H1′ are marked by the residue position. The dotted lines show the connectivities for the strand G_I-T_{IV}, while the solid lines indicate the corresponding NOEs for the strand G_a-T_d. The separation between a nucleobase proton (G H8 or T H6) and its own H1′ proton is 2.5 Å for a syn glycosidic conformation and 3.7 Å for the anti orientation. NOESY contour plots recorded at short mixing times (50 ms, data not shown) revealed that the cross-peaks in Figure 6 labeled G_x, where x = a−d, I−IV, all showed weaker intensities than the NOEs between the T H6 and T CH_3 protons (This inter-proton separation is fixed at 2.9 Å). This is a well-established method of confirming anti glycosidic bond conformations in nucleotides.[5a] Thus, in **$(G3T)_8$**, the glycosidic bonds in all of the eight distinct nucleotides are in the anti conformation.

(iii) Diagonal Strands in $(G3T)_8$ Are Staggered. NOESY spectra taken in 10% D_2O/H_2O revealed the existence of five pairs of well-resolved amino protons separated by differences in chemical shifts ranging between 2 and 3.5 δ ppm. The

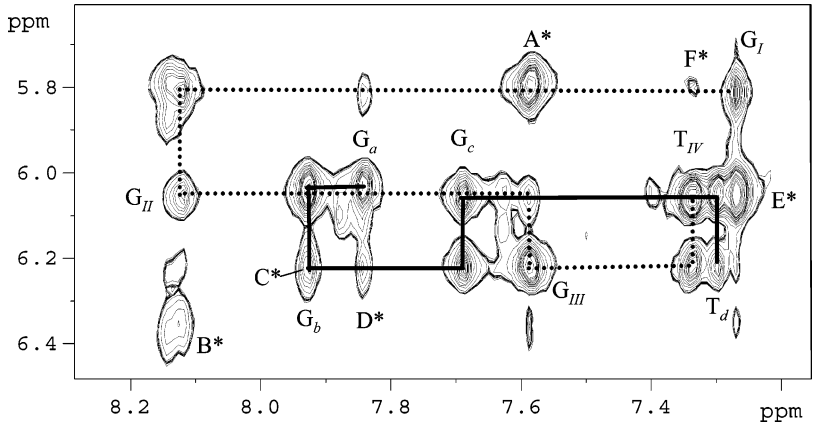


Figure 6. Portion of the NOESY spectrum, pH 6.7, 5 mM **G3T**, 12 °C, 10% D_2O/H_2O, showing sequential H8(H6)−sugar H1′ and H8−NH_{2e} NOEs: A*, $G_{III}(H8)-G_b(NH_{2e})$; B*, $G_{II}(H8)-G_a(NH_{2e})$; C*, $G_b(H8)-G_{III}(NH_{2e})$; D*, $G_a(H8)-G_{II}(NH_{2e})$; E*, $G_I(H8)-G_I(NH_{2e})$; and F*, $T_{IV}(H6)-G_c(NH_{2e})$.

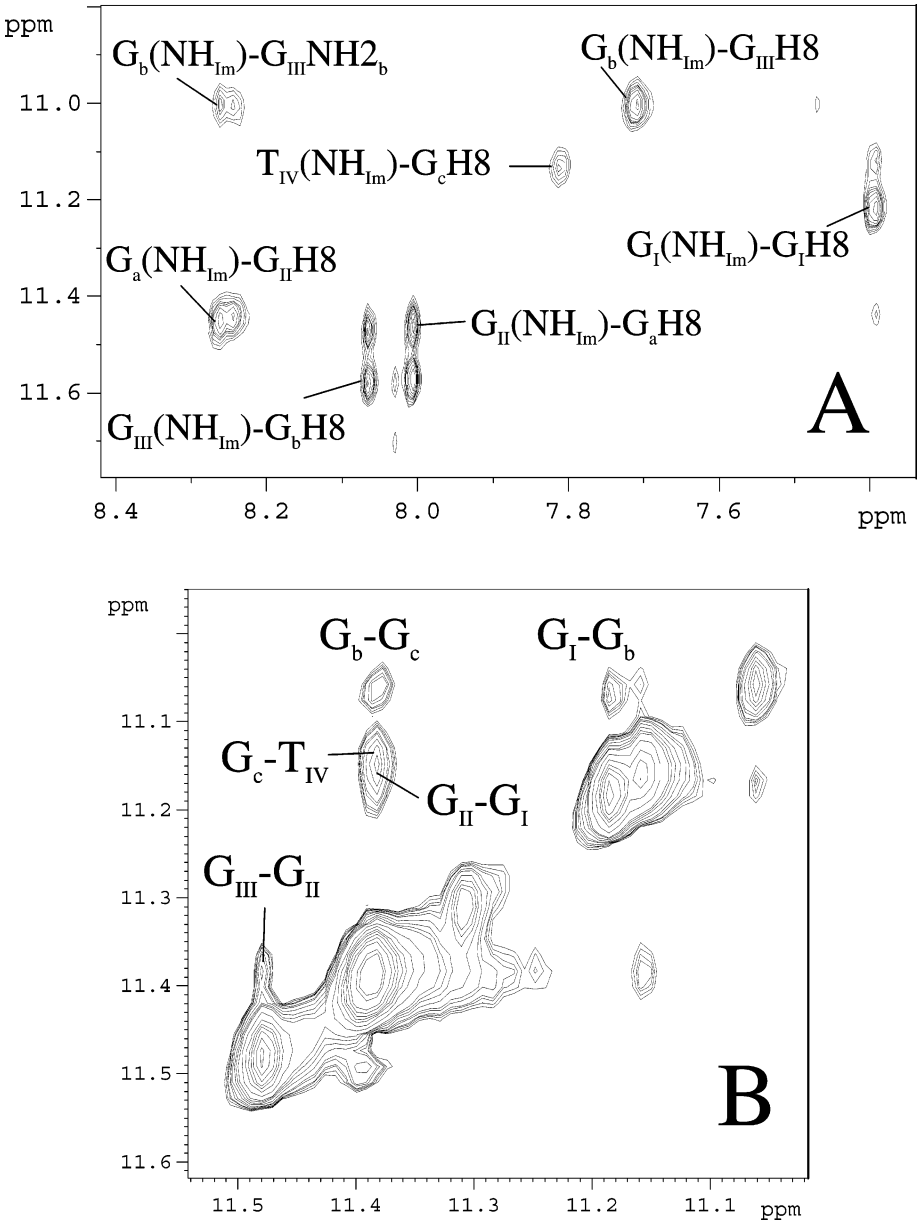
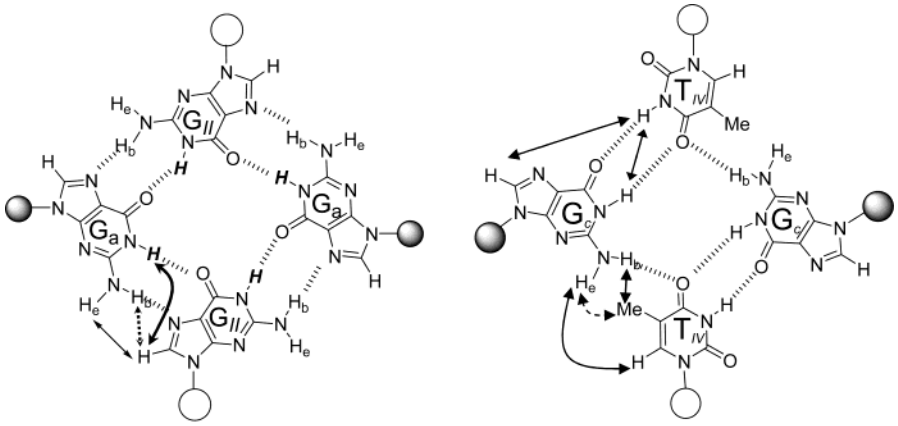


Figure 7. Portion of the NOESY spectrum, pH 6.7, 5 mM **G3T**, 12 °C, 10% D_2O/H_2O, showing (A) imino proton (NH_{Im}) contacts in the region 7.5−8.4 δ ppm and (B) $NH_{Im}-NH_{Im}$ cross-peaks in the region 11−12 δ ppm.

downfield-shifted amino protons resonating between 8 and 10 δ ppm must be hydrogen-bonded (NH_{2b}), while the upfield-shifted amino protons between 5.5 and 6.5 δ ppm must be non-hydrogen-bonded (NH_{2e}). Barring G_c, every H8 proton showed



(A) (B)

Figure 8. (A) Representative G-quartet in **(G3T)$_8$** with C_{2v}-type symmetry showing the expected H8 NOEs. (B) Proposed structure of the GTGT tetrad consistent with the experimentally observed NOEs indicated by arrows.

one set of cross-peaks (A*−E*) to a hydrogen-bonded amino proton (NH_{2b}) (Figure 5A,B) and one set of cross-peaks (A*−E*) with an exposed amino proton (NH_{2e}) (Figure 6). The observation of two kinds of guanine amino protons confirms Hoogsteen base pairing, and the existence of five such pairs confirms the presence of five G-tetrads.

The observation of a single set of G(H8)−NH_{2b} NOEs and G(H8)−NH_{2e} NOEs per guanine H8 is consistent with a C_{2v}-type symmetrical structure where diagonal strands are "slipped". The symmetrical nature of a representative tetrad in such a structure is shown in Figure 8A). The absence of two sets of H8−NH_{2b} NOEs and two sets of H8−NH_{2e} NOEs per guanine H8 rules out model C in Chart 1, where adjacent strands are slipped.

(iv) (G3T)$_8$ Is Stabilized by Peripheral GTGT Tetrads. $T_{IV}(CH_3)$ showed additional NOEs with an amino proton at 8.15 δ ppm (Figure 5C, P*) and another at 5.8 δ ppm (Figure 5C, R*). These amino protons can only belong to G_c, as all the others have already been accounted for. The difference in chemical shifts is about 2.4 δ ppm, suggesting that G_c is also present in a Hoogsteen type of arrangement where one amino proton is hydrogen-bonded (NH_{2b}) and the other is exposed to the solvent (NH_{2e}). The only nucleobases that could take part in a Hoogsteen-type arrangement with G_c are T_{IV} and T_d. Because of its anti conformation, T_d can be ruled out. Furthermore, the NOE between T_{IV}(H6) and $G_c(NH_{2e})$ (Figure 6, F*) confirms that T_{IV} and G_c are in the same plane, where one amino proton on G_c is hydrogen-bonded, possibly to the carbonyl group on T_{IV} (Figure 8B). G-T base pairing has been recently observed in the crystal structure of ss-DNA bound to *Schizosaccharomyces pombe* Pot1p, where the unusual G-T base pairs compact the DNA strand.[18,19]

(v) (G3T)$_8$ Is Hoogsteen Base Paired with Five G-Tetrads. Peaks in the region 10−12 δ ppm characteristic of hydrogen-bonded imino protons (NH_{Im}) on the guanines (Figure 4C) were observed, showing that the guanines in **(G3T)$_8$** are indeed present as tetrads. Considerable overlapping of imino protons was observed. However, these were resolved in the 2D spectra (Figure 7). Imino protons corresponding to G_a and G_b could be assigned from NOEs with the distinct G_{II} (not shown) and G_{III} NH_{2b} protons (Figure 7A). The symmetric G_I tetrad was identified by the NOE cross-peak with its own H8 (Figure 7A). Interestingly, G_c(H8) shows an NOE with the amidic proton on T_{IV}(NH) (Figure 7, peak E), which has shifted downfield to 11.1 δ ppm. This suggests that the T_{IV}-(NH) is hydrogen-bonded, possibly to G_c(O6). This is reinforced by a relatively strong NOE of T_{IV}(NH) with $G_c(NH_{Im})$ (Figure 7B, L), confirming the presence of a GTGT tetrad. $G_{II}(NH_{Im})$ and $G_{III}(NH_{Im})$ were assigned from sequential NOEs (Figure 7B).

Thus **(G3T)$_8$** has five G-tetrads and two GTGT tetrads. This eliminates C_2-symmetric structures involving "slippage" by more than one nucleobase as such structures would lead to an interlocked structure with five tetrads (including the two GTGT tetrads) or less. Blunt-end stacking by two discrete tetramolecular quadruplexes would yield a dimer having six G-tetrads. However, the C_{2v}-type interlocking structure (Chart 1B) has a total of seven tetrads, including the two extra GTGT tetrads at the 3′ ends, favoring this particular mode of interlocking over any other. This is a possible explanation for the preference of **G3T** to form the C_{2v}-type interlocked quadruplex over any other possible interlocked structure.

Conclusions

Sequences that possess a single stretch of guanines at either their 5′ or 3′ end have been shown to form higher order quadruplex structures that could consist of either discrete quadruplexes stacking end-to-end or interlocking strands.[7] While the possibility of an interlocking quadruplex motif has been postulated,[6,16] this paper describes unequivocal evidence of such a motif in these higher order structures. One component of the

(17) (a) Cheong, C.; Moore, P. B. *Biochemistry* **1992**, *31*, 8406. (b) Hardin, C. C.; Corregan, M. J.; Lieberman, D. V.; Brown, B. A. *Biochemistry* **1997**, *36*, 15428. (c) Aboulela, F.; Murchie, A. I. H.; Norman, D. G.; Lilley, D. M. J. *J. Mol. Biol.* **1994**, *243*, 458.

(18) Lei, M.; Podell, E. R.; Baumann, P.; Cech, T. R. *Nature* **2003**, *426*, 198−203.

(19) Other novel tetrads include: Zhang, N.; Gorin, A.; Majumdar, A.; Kettani, A.; Chernichenko, N.; Skripkin, E.; Patel, D. J. *J. Mol. Biol.* **2001**, *312*, 1073−1088. Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2001**, *105*, 10366−10371. Pan, B. C.; Xiong, Y.; Shi, K.; Deng, J. P.; Sundaralingam, M. *Structure* **2003**, *11*, 815−823. Meyer, M.; Schneider, C.; Brandl, M.; Suhnel, J. *J. Phys. Chem. A.* **2001**, *105*, 11560−11573.

driving force of oligomerization of quadruplexes is probably hydrophobic stabilization. Two discrete tetramolecular quadruplexes would have more aromatic quartet faces exposed to solvent than one dimerized quadruplex. By assembling into a higher order structure, the system is able to exclude a greater number of solvent molecules. The formation of the interlocked structure B in particular, over other possible structures, is probably also due to the system maximizing the number of stacked tetrads. Interlocking structures slipped by more than one base are not optimal from the point of view of the number of tetrads generated, and are therefore disfavored.

Acknowledgment. We thank the Interdisciplinary Research Collaboration in Nanotechnology for funding and Professor Ben Luisi for valuable discussions. Y.K.-G. thanks the Royal Commission for the Exhibition of 1851 for a Research Fellowship.

Supporting Information Available: NOE assignments in **(G3T)$_4$** from H8/H6−H1′ NOEs. This material is available free of charge via the Internet at http://pubs.acs.org.

JA049259Y