An unprecedented knot-like G-quadruplex peripheral motif

Truong, Thi Hong Anh; Winnerdy, Fernaldo Richtia; Phan, Anh Tuân

2019

Truong, T. H. A., Winnerdy, F. R., & Phan, A. T. (2019). An unprecedented knot-like G-quadruplex peripheral motif. Angewandte Chemie International Edition, 58(39), 13834-13839. doi:10.1002/anie.201907740

https://hdl.handle.net/10356/140168

https://doi.org/10.1002/anie.201907740

COMMUNICATION

An unprecedented knot-like G-quadruplex peripheral motif

Thi Hong Anh Truong[#], Fernaldo Richtia Winnerdy[#] and Anh Tuân Phan*

Abstract: *We discovered the formation of a knot-like G-quadruplex peripheral structure by a 7-nt DNA sequence DL7: 5'-d(TGTTGGT)-3', whereby six out of its seven nucleobases participate in compact base pairing interactions. Here, solution structure of a 24-nt DNA oligonucleotide containing the DL7 sequence shows the interaction between a two-layer anti-parallel G-quadruplex core and the peripheral structure, including the construction of two sharp turns from the DNA backbone. The formation of this novel structural element highlights the intricate property of single-stranded DNA folding in presence of G-quadruplex-forming motifs. We demonstrated the compatibility of the DL7 knot-like structure with various G-quadruplexes, which could have implications in drug design and DNA engineering.*

Deoxyribonucleic acid (DNA) is a biopolymer which carries the genetic information of the living organisms.[1] Although most biological functions are regulated by the primary sequence of DNA,[2] it was established that the secondary conformations ─ the most common being the DNA double helix[3] ─ are also vital in controlling gene expression.[4] The secondary structures that might be adopted by a strand of DNA are complex, as the existence of structures such as unconventional hairpins,[5] triplexes[6] and quadruplexes[7] were discovered. Here, we report a unique secondary structure formation from a 7-nt DNA sequence d(TGTTGGT) (named *DL7*) in presence of a G-quadruplex (G4). Six out of seven nucleotides in the sequence are engaged in compact base pairing interactions, constructing a knot-like DNA structure stacking onto the G4 planar surface. The structure serves as a G4-stabilizing framework and potential drug target.

G4s are secondary structures of nucleic acids that can form in guanine-rich sequences under physiological condition.[8] The core structure of G4 is normally composed of two or more stacking G-tetrads with coordination cations at the center. Each G-tetrad comprises of four guanine bases in a square-planar arrangement with eight Hoogsteen hydrogen bonds.[9] G4-forming motifs[10] are most abundantly found in biologically-relevant sites such as telomeric[11], gene promoter[12] and replication initiation regions.[13] G4 structures highly influence some regulatory functions such as telomerase inhibition,[14] gene expression[12b, 15] and DNA replication.[16] Polymorphism of G4s[7, 17] remains a prominent topic of study, following the introductions of bulges,[18] duplex-loops,[19] triad-stacking,[20] non-canonical tetrad,[21] left-handed fold[22] and many others. In this work, the solution structure of a *DL7*-containing DNA sequence is resolved, revealing the detailed interactions between the knot-like structure and the G4 core. Based on this structure, a cut-and-paste study[23] was performed on two other G4 structures, *Pu24T* from the *c-myc* promoter[24] and an artificially designed G4 (*AQ*),[17c] whereby the *DL7* sequence was inserted as diagonal loops. Our results suggest that the *DL7* knot-like structure formation is compatible with various G4s.

NMR spectrum of a modified anti-cancer oligonucleotide reveals a unique pattern. An anti-cancer DNA aptamer containing 17 guanines *AGRO100* (also known as *AS1411*) was shown to adopt multiple G4 conformations.[25] Single G-to-T substitutions at various positions of *AGRO100* have resulted in the formation of distinct G4 structures.[22a, 26] The substitution at position 11 (*AT11*) generated a four-layer parallel G4, while another substitution at position 27 (*AT27*) resulted in a four-layer parallel left-handed G4. A G-to-T substitution in position 21 was attempted, giving the *AT21* sequence (Figure 1A). The one-dimensional (1D) NMR spectrum of *AT21* showed 12 sharp imino proton peaks including an unusual three-peak pattern at ~9-10.5 ppm (Figure 1B), suggesting a possible formation of unique base-pairing assemblies outside of G-tetrads. The CD spectrum, displaying a trough at ~265 nm and positive peaks at ~245 and ~290 nm (Figure 1C), indicated an anti-parallel G4 formation.[27]

***DL7* sequence is responsible for unprecedented base pairing interactions.** To identify the guanine and thymine bases responsible for the *AT21* structure folding, we performed ^{15}N-^{1}H

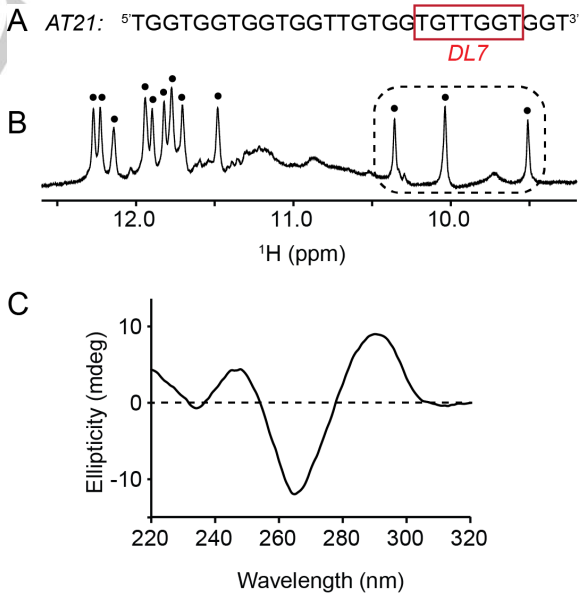


Figure 1. Spectroscopy of *AT21*. (A) DNA sequence of *AT21* with the *DL7* motif highlighted in red box. (B) 1D NMR spectrum of *AT21* showing 12 well-resolved imino proton peaks with a unique three-peak pattern shown in dashed box. (C) CD spectrum of *AT21* showing a typical profile of an anti-parallel G4 fold.

[*] Dr. T. H. A. Truong, Dr. F. R. Winnerdy, Prof. Dr. A. T. Phan
School of Physical and Mathematical Sciences
Nanyang Technological University, Singapore 637371
E-mail: phantuan@ntu.edu.sg

Prof. Dr. A. T. Phan
NTU Institute of Structural Biology
Nanyang Technological University, Singapore 636921

[#] These authors contributed equally to this work.

Supporting information for this article is given via a link at the end of the document.((Please delete this text if not appropriate))

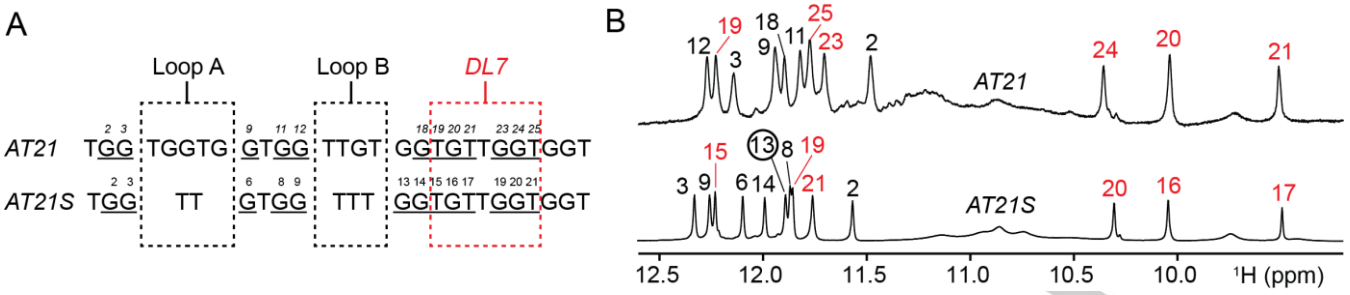


Figure 2. *AT21* derivative *AT21S*. (A) The DNA sequence of *AT21* and *AT21S*. Loop A, loop B and *DL7* are shown in dashed boxes. The underlined bases indicate the observation of their imino protons. (B) 1D NMR spectra for *AT21* and *AT21S*. Twelve and thirteen imino proton peaks are observed for *AT21* and *AT21S* respectively. Resonance assignments are shown with residue numbers; imino protons from *DL7* are indicated in red.

HMQC experiments using multiple site-specific 2% ^{15}N-labeled samples[28] and obtained full assignments of the twelve observed imino protons (Figure S1 in the Supporting Information). It revealed the participation of G2-G3, G9, G11-G12, G18-T21 and G23-T25 in the base pairing interactions, which included six out of seven nucleobases of *DL7* (Figure 2A, top). Interestingly, all three upshifted peaks at ~9-10.5 ppm were identified to be imino proton peaks of the *DL7* sequence (Figure 2B, top), suggesting major contribution of *DL7* towards the unique non-G-tetrad structure. The two segments in the middle of the sequence, T4-G8 and T13-T16, did not show observable imino protons indicative of base pairing, suggesting their possible roles as loops (Figure 2A).

Mutations of loops preserve the overall structure. The presence of unstructured long loops is known to destabilize G4 structures,[29] therefore we attempted several loop modifications to get a more stable G4 structure without changing its folding pattern. The loops were designated as loop A (T4-G8) with five nucleotides and loop B (T13-T16) with four nucleotides. We performed various loop modifications on *AT21* and recorded 1D NMR and CD spectra for each modified sequence (Figure S2, Table S1). Most mutated sequences showed very similar CD spectra and well-resolved NMR spectra preserving the unusual three-peak pattern. For further structural study, we chose one sequence where loop A was converted to TT and loop B to TTT, designated as *AT21S* (Figure 2A, bottom), due to its well-resolved NMR spectrum. The melting temperature (T_m) of *AT21S* was observed to be ~39°C, approximately 6°C higher than its predecessor *AT21* (Figure S3).

Resonance assignments of *AT21S*. The 1D NMR spectrum of *AT21S* showed thirteen imino proton peaks, which corresponded to ten guanines (H1) and three thymines (H3) as identified by the 2D ^{15}N-^{1}H HMQC spectrum (Figure S4). The imino proton assignments of *AT21S* were obtained by comparing NOESY spectra of *AT21* and *AT21S* (Figure S5A, B). Note the different numbering of nucleotides due to the loop shortening (Figure 2A). Guanine aromatic proton (H8) assignments were achieved by a long-range through-bond coupling (via C5) HMBC experiment.[28] Every aromatic proton of base-pairing guanines was assigned (Figure S5C). The additional imino proton observed for *AT21S*, but not *AT21*, was identified to correspond to G13 (Figure S6A). It is likely that the G17 counterpart in *AT21* was broadened beyond detection due to motions of the original loops A and B. The complete imino proton assignments of *AT21S* are presented (Figure 2B).

***AT21S* forms two-layer G4 accompanied by a compact peripheral structure.** The H1-H8 region of the NOESY spectrum obtained on a sample dissolved in H_2O revealed two cyclic patterns: G2•G14•G8•G19 and G3•G6•G9•G13 (Figure 3A), indicating the formation of a two-layer G4 structure. We inferred from the data that the strand orientation of G13-G14 was opposite to those of G2-G3 and G8-G9, while G6 and G19 formed a discontinuous G-tract. In the NOESY spectrum obtained on a sample dissolved in D_2O, strong intra-residue H8-H1' cross-peak were observed for four guanines: G2, G8, G13 and G19, indicating their *syn* glycosidic bond conformations. We also found four rectangular patterns from the intra-residue and inter-residue H1'-H8 cross peaks originated from the following pairs: G2-G3, G8-G9, G13-G14 and G19-G6 (Figure S6B). These rectangular patterns signified *syn-anti* glycosidic bond conformation steps, consistent with the anti-parallel fold.[28] Overall, the NMR signal of *AT21S* indicated a chair-type anti-parallel G4 core with one discontinuous G-tract (Figure 3B).

The unusual peripheral structure was analyzed by examining the upshifted imino proton region of the NOESY spectrum. The observed imino-imino and imino-methyl proton cross-peaks indicated the formation of novel structural elements as illustrated, where the T15•T17•G20 triad and G16•T21 base pair assembled in a compact fashion (Figure 3C, D). These structural elements were constructed entirely by the *DL7* sequence d(TGTTGGT), with the three upshifted peaks at ~9-10.5 ppm determined to be imino protons of guanines and thymines which were hydrogen-bonded with thymine O2 or O4 atoms. Based on all the NMR data, we found that the peripheral structure is stacked onto the G4 core.

To probe the stabilizing role of the *DL7* motif, we substituted *DL7* by a three-thymine sequence capable of forming a diagonal loop,[30] giving rise to the *AT21S-3T* sequence (Table S2). NMR and UV-melting experiments showed that the structure of the mutated sequence *AT21S-3T* (T_m<25°C) was less stable than that of *AT21S* (Figure S7). These results indicated the stabilizing nature of the peripheral structure of *DL7*.

Structural analysis reveals a knot-like backbone progression. The structure of *AT21S* was calculated using constraints generated from the NMR data (Table 1A). One hundred structures were built by simulated annealing and refinement processes, whereby the ten lowest-energy structures were selected (Figure 4A, Table 1B). A representative structure is displayed in cartoon form (Figure 4B), showing the stacking

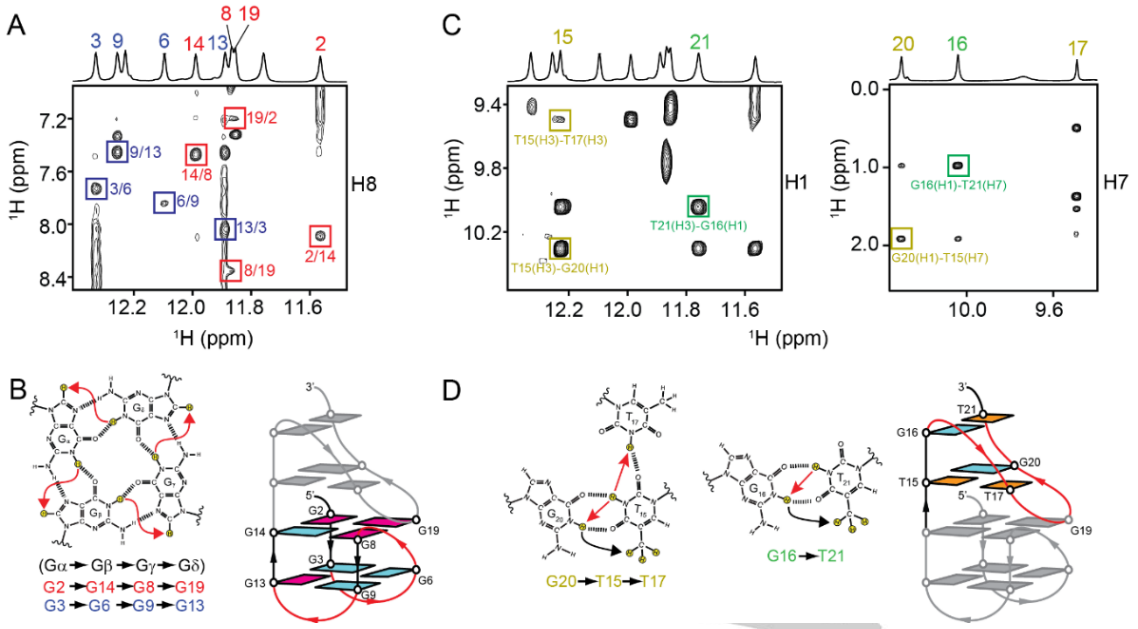


Figure 3. 2D NOESY (mixing time, 200 *ms*; temperature, 25°C) analysis of *AT21S*. (A) H1-H8 region of the NOESY spectrum of *AT21S*. Indicated in boxes are the H1-H8 NOE cyclic-connectivity patterns indicating the formation of G-tetrads. (B) The schematic of NOE patterns observed in G-tetrads and the proposed G4 core of *AT21S*. The arrows indicate DNA backbone orientation; cyan and magenta represent *anti*- and *syn*-guanine respectively; greyed area shows the peripheral structure. (C) Different regions of the NOESY spectrum of *AT21S*. Indicated in boxes are some key NOE cross-peaks used to determine the base interactions. (D) The schematics of NOE peaks seen in the spectra as well as proposed fold of the peripheral structure of *AT21S*. The arrows indicate DNA backbone orientation; cyan represents *anti*-guanine; orange represents thymine; greyed area shows the G4 core structure.

interaction between the G4 core and the peripheral structure. *Syn-anti* steps are observed for three G-tracts (G2-G3, G8-G9 and G13-G14), as well as one discontinuous G-tract (G19-G6). There are two edgewise loops (T4-T5 and T10-T12), one V-shaped loop[31] (T7) and one diagonal loop (T15-T18) in the structure. The guanine residue G19, which is a part of both the *DL7* motif and one G-tetrad, serves as a bridge between the peripheral structure and the G4 core.

The stacking interfaces between different base pair, triad, and tetrad layers are presented in detail (Figure S8). Both the T•T•G layer formed in the *DL7* motif and the T•T•T layer formed between the edgewise loops cap the two G-tetrads and protect their imino protons from exchange with solvent, consistent with the observation of these protons 1 hour after dissolving the sample in D_2O (Figure S9). The overall backbone progression of *AT21S*, specifically in *DL7* region, resembles the shape of a knot with two back-to-back U-turns in a seven-nucleotide span (Figure 4C).

Compatibility of *DL7* with various G-quadruplexes. The resolved *AT21S* structure revealed the importance of the *DL7* sequence in forming a G4-stabilizing knot-like structure. To confirm that the *DL7* sequence is sufficient for that function, we performed 1D NMR experiments on truncated sequences (Figure S10). The results indicated that the 5'-end thymine (T1) and the 3'-end segment G22-T24 are unnecessary in maintaining the fold. In a broader context, the *DL7* motif represents a particular case, where a diagonal loop adopts a well-defined structure. The significance of this knot-like structure was gauged by testing its compatibility with other G4 structures. For this purpose, we chose a G-rich sequence from the *c-myc* promoter *Pu24T* (Table S3),[24] which formed a three-layer parallel G4 structure containing a snapback triad-forming diagonal loop. We executed a cut-and-paste study by incorporating the *DL7* sequence to the diagonal loop of *Pu24T*, designated as *Pu24T-mod* (Table S3, Figure S11A). The 1D NMR spectrum of *Pu24T-mod* displayed sharp imino proton peaks corresponding to the G-tetrad core, as well as three ancillary imino proton peaks (Figure 5B) at 9-10.5 ppm. Highly similar spectral patterns of *Pu24T-mod* with *AT21S* in this area strongly suggest the formation of a similar knot-like structure by *DL7* in *Pu24T-mod* (Figures 5A and S11B), albeit less robust as shown by the broadened ancillary peaks. A similar cut-and-paste study was also performed on an engineered G4,[17c] *AQ-mod* (Table S3, Figure S12A). Three up-field shifted imino proton peaks at 9-10.5 ppm were observed (Figures 5C and S12B). The

Table 1. NMR restraints and statistics of the computed structures of *AT21S*.

A. NMR Restraints		
Distance restraints	H_2O	D_2O
Intra-residue	4	275
Inter-residue	82	115
Other restraints		
Hydrogen bond	42	
Dihedral angle	10	
Planarity	4	
B. Structure Statistics		
NOE violations		
Number (>0.2 Å)	0.3 ± 0.5	
Deviations from the ideal geometry		
Bond lengths (Å)	0.003 ± 0.000	
Bond angles (°)	0.695 ± 0.006	
Impropers (°)	0.366 ± 0.006	
Pairwise heavy atom RMSD value (Å)		
G-tetrad core	0.726 ± 0.183	
All heavy atoms	1.170 ± 0.248	

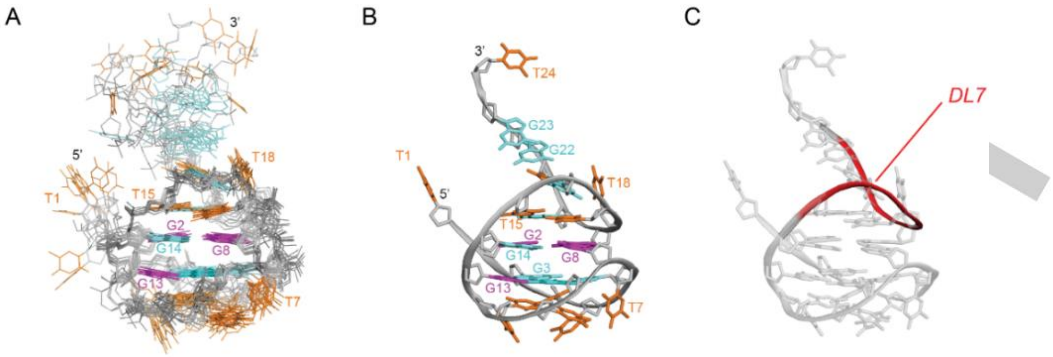


Figure 4. Solution structure of *AT21S*. (A) Superimposed ten lowest-energy structures of *AT21S*. (B) Cartoon representation of the lowest-energy structure of *AT21S*. Cyan, magenta and orange indicates *anti*-guanine, *syn*-guanine and thymine respectively. (C) Cartoon views of *AT21S* highlighting *DL7* backbone in red, showing two sharp turns constructing a knot-like progression.

comparison between these two sequences (*Pu24T-mod* and *AQ-mod*) and our reference sequence (*AT21S*) is summarized in Figure 5. We performed unambiguous spectral assignments using site-specific ^{15}N enrichment (Figure S13) and confirmed the identities of the up-field shifted imino proton peaks in each sequence to the respective expected position in the *DL7* segment (Figure 5). 2D NOESY of *Pu24T-mod* and *AQ-mod* also revealed the correlations between imino protons of the *DL7* residues, supporting the formation of a similar *DL7* knot-like structure (Figure S14). The CD profiles of *Pu24T-mod* and *AQ-mod* were consistent with the proposed parallel and (3+1) hybrid G4 folding topologies (Figure S15), which were designed to incorporate the knot-like *DL7* structure. In the perspective of G4 engineering, the *DL7* sequence can be incorporated into a G4-forming sequence to produce a compact and self-stabilizing structure.

Possible biological relevance of the *DL7* motif. Given the short nature of the *DL7* sequence (7-nt), we expect a high number of occurrences of *DL7* and similar motifs in the human genome. A bioinformatics search on the reference strands of the human genome resulted in ~270,000 exact matches, exceeding the expected number of random occurrences by ~1.5 times (see Supporting Information). The formation of the *DL7* knot-like structure could modulate the recognition and stability of adjacent G4s, affecting their role in biological processes involving DNA, such as transcription, replication and repair. A similar knot-like structure could also form and modulate the RNA structure and function. RNA pseudoknot is known to play a role in RNA splicing and ribosomal frame-shifting.[32]

In summary, the *DL7* sequence was established to form a compact structure as a G-quadruplex peripheral. This knot-like structure is comprised of unique base pairing interactions from six out of seven nucleobases and accommodated two sharp turns on the DNA backbone. Its G4-stabilizing property was examined from the resolved *AT21S* structure, where the T•G base pair and T•T•G triad of *DL7* were observed to stack onto the G-tetrad core. We successfully incorporated the *DL7* motif into other G4 structures. The *DL7* motif is an example of an intricate folding of a short single-stranded DNA in presence of a G4 motif. The compatibility and stabilizing effects of *DL7* could prove useful in G4 sequence engineering.

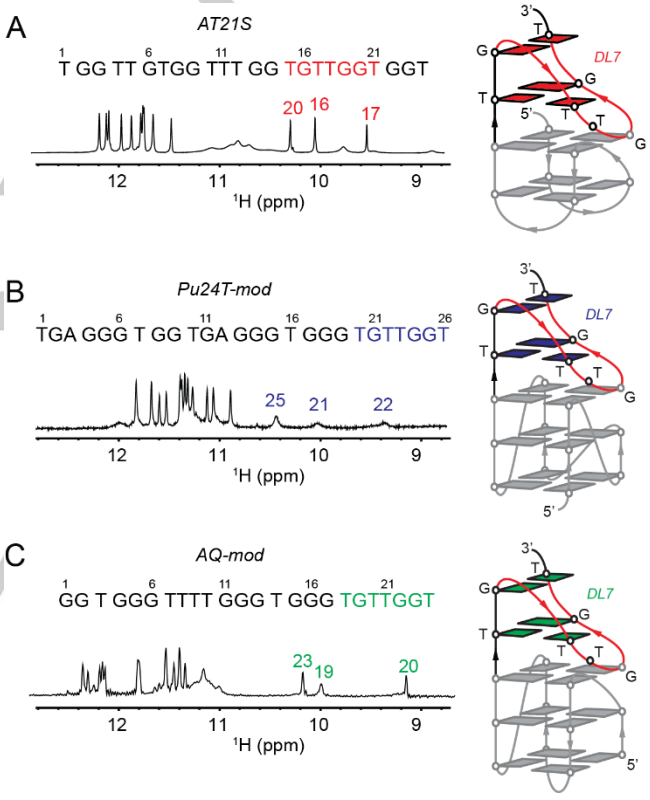


Figure 5. Compatibility of *DL7* with various G-quadruplexes. The sequences, imino proton NMR spectra and proposed folds of (A) *AT21S*, (B) *Pu24T-mod* and (C) *AQ-mod*. Indicated in the sequences with different colours are the *DL7* segments. The assigned three-peak pattern is shown in all spectra. The schematics show three folding topologies of anti-parallel, parallel, and (3+1) hybrid respectively.

Acknowledgements

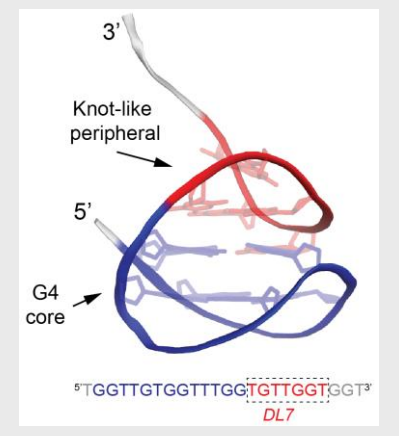
This research was funded by Singapore National Research Foundation Investigatorship (NRF-NRFI2017-09) and Ministry of Education Academic Research Fund Tier 2 (MOE2015-T2-1-092) and grants from Nanyang Technological University to A.T.P. The authors acknowledge the use of NMR facilities at NTU Institute of Structural Biology.

Keywords: G-quadruplex • DNA • Knot • NMR structure • Compact fold

[1] A. D. Hershey, M. Chase, *J Gen Physiol* **1952**, *36*, 39-56.
[2] F. Crick, *Nature* **1970**, *227*, 561-563.
[3] J. C. Wang, *Proc Natl Acad Sci U S A* **1979**, *76*, 200-203.
[4] a) D. Rhodes, H. J. Lipps, *Nucleic Acids Res* **2015**, *43*, 8627-8637; b) N. Maizels, *EMBO Rep* **2015**, *16*, 910-922.
[5] a) A. Kovanda, M. Zalar, P. Sket, J. Plavec, B. Rogelj, *Sci Rep* **2015**, *5*, 17944; b) M. Gajarsky, M. L. Zivkovic, P. Stadlbauer, B. Pagano, R. Fiala, J. Amato, L. Tomaska, J. Sponer, J. Plavec, L. Trantirek, *J Am Chem Soc* **2017**, *139*, 3591-3594.
[6] a) J. C. Hanvey, M. Shimizu, R. D. Wells, *Proc Natl Acad Sci U S A* **1988**, *85*, 6292-6296; b) V. Limongelli, S. De Tito, L. Cerofolini, M. Fragai, B. Pagano, R. Trotta, S. Cosconati, L. Marinelli, E. Novellino, I. Bertini, A. Randazzo, C. Luchinat, M. Parrinello, *Angew Chem Int Ed Engl* **2013**, *52*, 2269-2273.
[7] a) S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res* **2006**, *34*, 5402-5415; b) J. T. Davis, *Angew Chem Int Ed Engl* **2004**, *43*, 668-698; c) D. J. Patel, A. T. Phan, V. Kuryavyi, *Nucleic Acids Res* **2007**, *35*, 7429-7455; d) J. L. Mergny, D. Sen, *Chem Rev* **2019**, *119*, 6290-6325.
[8] D. Sen, W. Gilbert, *Nature* **1988**, *334*, 364-366.
[9] M. Gellert, M. N. Lipsett, D. R. Davies, *Proc Natl Acad Sci U S A* **1962**, *48*, 2013-2018.
[10] a) J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res* **2005**, *33*, 2908-2916; b) A. K. Todd, M. Johnston, S. Neidle, *Nucleic Acids Res* **2005**, *33*, 2901-2907.
[11] a) J. R. Williamson, M. K. Raghuraman, T. R. Cech, *Cell* **1989**, *59*, 871-880; b) W. I. Sundquist, A. Klug, *Nature* **1989**, *342*, 825-829.
[12] a) T. Simonsson, P. Pecinka, M. Kubista, *Nucleic Acids Res* **1998**, *26*, 1167-1172; b) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc Natl Acad Sci U S A* **2002**, *99*, 11593-11598.
[13] a) C. Cayrou, P. Coulombe, A. Vigneron, S. Stanojcic, O. Ganier, I. Peiffer, E. Rivals, A. Puy, S. Laurent-Chabalier, R. Desprat, M. Mechali, *Genome Res* **2011**, *21*, 1438-1449; b) E. Besnard, A. Babled, L. Lapasset, O. Milhavet, H. Parrinello, C. Dantec, J. M. Marin, J. M. Lemaitre, *Nat Struct Mol Biol* **2012**, *19*, 837-844.
[14] A. M. Zahler, J. R. Williamson, T. R. Cech, D. M. Prescott, *Nature* **1991**, *350*, 718-720.
[15] L. A. Cahoon, H. S. Seifert, *Science* **2009**, *325*, 764-767.
[16] a) J. Lopes, A. Piazza, R. Bermejo, B. Kriegsman, A. Colosio, M. P. Teulade-Fichou, M. Foiani, A. Nicolas, *EMBO J* **2011**, *30*, 4033-4046; b) K. Paeschke, J. A. Capra, V. A. Zakian, *Cell* **2011**, *145*, 678-691.
[17] a) J. Dai, M. Carver, D. Yang, *Biochimie* **2008**, *90*, 1172-1183; b) A. T. Phan, *FEBS J* **2010**, *277*, 1107-1117; c) S. A. Dvorkin, A. I. Karsisiotis, M. Webba da Silva, *Sci Adv* **2018**, *4*, eaat3007.
[18] a) V. T. Mukundan, A. T. Phan, *J Am Chem Soc* **2013**, *135*, 5017-5028; b) B. De Nicola, C. J. Lech, B. Heddi, S. Regmi, I. Frasson, R. Perrone, S. N. Richter, A. T. Phan, *Nucleic Acids Res* **2016**, *44*, 6442-6451; c) A. Kerkour, J. Marquevielle, S. Ivashchenko, L. A. Yatsunyk, J. L. Mergny, G. F. Salgado, *J Biol Chem* **2017**, *292*, 8082-8091; d) Y. Liu, W. Lan, C. Wang, C. Cao, *J Biol Chem* **2018**, *293*, 8947-8955; e) A. Sengar, J. J. Vandana, V. S. Chambers, M. Di Antonio, F. R. Winnerdy, S. Balasubramanian, A. T. Phan, *Nucleic Acids Res* **2019**, *47*, 1564-1572.
[19] a) K. W. Lim, A. T. Phan, *Angew Chem Int Ed Engl* **2013**, *52*, 8566-8569; b) K. W. Lim, T. Q. Nguyen, A. T. Phan, *J Am Chem Soc* **2014**, *136*, 17969-17973; c) H. J. Kang, Y. Cui, H. Yin, A. Scheid, W. P. Hendricks, J. Schmidt, A. Sekulic, D. Kong, J. M. Trent, V. Gokhale, H. Mao, L. H. Hurley, *J Am Chem Soc* **2016**; d) B. Onel, M. Carver, G. Wu, D. Timonina, S. Kalarn, M. Larriva, D. Yang, *J Am Chem Soc* **2016**, *138*, 2563-2570; e) M. L. Greco, A. Kotar, R. Rigo, C. Cristofari, J. Plavec, C. Sissi, *Nucleic Acids Res* **2017**, *45*, 10132-10142; f) E. Butovskaya, B. Heddi, B. Bakalar, S. N. Richter, A. T. Phan, *J Am Chem Soc* **2018**, *140*, 13654-13662.
[20] a) A. Kettani, S. Bouaziz, W. Wang, R. A. Jones, D. J. Patel, *Nat Struct Biol* **1997**, *4*, 382-389; b) K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J Am Chem Soc* **2006**, *128*, 9963-9970; c) J. Dai, C. Punchihewa, A. Ambrus, D. Chen, R. A. Jones, D. Yang, *Nucleic Acids Res* **2007**, *35*, 2440-2450; d) A. Matsugami, Y. Xu, Y. Noguchi, H. Sugiyama, M. Katahira, *FEBS J* **2007**, *274*, 3545-3556; e) J. Wirmer-Bartoschek, L. E. Bendel, H. R. A. Jonker, J. T. Grun, F. Papi, C. Bazzicalupi, L. Messori, P. Gratteri, H. Schwalbe, *Angew Chem Int Ed Engl* **2017**, *56*, 7102-7106.
[21] a) A. Kettani, S. Bouaziz, A. Gorin, H. Zhao, R. A. Jones, D. J. Patel, *J Mol Biol* **1998**, *282*, 619-636; b) C. Caceres, G. Wright, C. Gouyette, G. Parkinson, J. A. Subirana, *Nucleic Acids Res* **2004**, *32*, 1097-1102; c) V. Kocman, J. Plavec, *Nat Commun* **2017**, *8*, 15355.
[22] a) W. J. Chung, B. Heddi, E. Schmitt, K. W. Lim, Y. Mechulam, A. T. Phan, *Proc Natl Acad Sci U S A* **2015**, *112*, 2729-2733; b) B. Fu, J. Huang, Y. Chen, Y. Wang, T. Xue, G. Xu, S. Wang, X. Zhou, *Chem Commun* **2016**, *52*, 10052-10055; c) C. Zhao, H. Song, P. Scott, A. Zhao, H. Tateishi-Karimata, N. Sugimoto, J. Ren, X. Qu, *Angew Chem Int Ed Engl* **2018**, *57*, 15723-15727; d) B. Bakalar, B. Heddi, E. Schmitt, Y. Mechulam, A. T. Phan, *Angew Chem Int Ed Engl* **2019**, *58*, 2331-2335.
[23] L. Hu, K. W. Lim, S. Bouaziz, A. T. Phan, *J Am Chem Soc* **2009**, *131*, 16824-16831.
[24] A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat Chem Biol* **2005**, *1*, 167-173.
[25] M. M. Dailey, M. C. Miller, P. J. Bates, A. N. Lane, J. O. Trent, *Nucleic Acids Res* **2010**, *38*, 4877-4888.
[26] N. Q. Do, W. J. Chung, T. H. A. Truong, B. Heddi, A. T. Phan, *Nucleic Acids Res* **2017**, *45*, 7487-7493.
[27] a) J. Kypr, I. Kejnovska, D. Renciuk, M. Vorlickova, *Nucleic Acids Res* **2009**, *37*, 1713-1725; b) A. I. Karsisiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew Chem Int Ed Engl* **2011**, *50*, 10645-10648; c) R. Del Villar-Guerra, J. O. Trent, J. B. Chaires, *Angew Chem Int Ed Engl* **2018**, *57*, 7171-7175.
[28] M. Adrian, B. Heddi, A. T. Phan, *Methods* **2012**, *57*, 11-24.
[29] A. Guedin, J. Gros, P. Alberti, J. L. Mergny, *Nucleic Acids Res* **2010**, *38*, 7858-7868.
[30] P. Hazel, J. Huppert, S. Balasubramanian, S. Neidle, *J Am Chem Soc* **2004**, *126*, 16405-16415.
[31] N. Zhang, A. Gorin, A. Majumdar, A. Kettani, N. Chernichenko, E. Skripkin, D. J. Patel, *J Mol Biol* **2001**, *311*, 1063-1079.
[32] a) O. Namy, S. J. Moran, D. I. Stuart, R. J. Gilbert, I. Brierley, *Nature* **2006**, *441*, 244-247; b) C. W. Pleij, K. Rietveld, L. Bosch, *Nucleic Acids Res* **1985**, *13*, 1717-1731.

COMMUNICATION

A knot-like DNA secondary structure was formed by 7-nt sequence *DL7*, d(TGTTGGT) next to a G-quadruplex, whereby six out of its seven nucleobases participate in compact base pairing interactions.



*Thi Hong Anh Truong, Fernaldo Richtia Winnerdy, Anh Tuân Phan**

Page No. – Page No.

An unprecedented knot-like G-quadruplex peripheral motif