Electronic Supplementary Material (ESI) for Chemical Communications.

A novel minimal motif for left-handed G-quadruplex formation

Poulomi Das,[a] Fernaldo Richtia Winnerdy,[a,b] Arijit Maity,[a] Yves Mechulam,[c] and Anh Tuân Phan*[a,b]

[a] School of Physical and Mathematical Sciences, Nanyang Technological University, 637371, Singapore

[b] NTU Institute of Structural Biology, Nanyang Technological University, Singapore 636921. Singapore

[c] Laboratoire de Biologie Structurale de la Cellule (BIOC), Ecole Polytechnique, CNRS-UMR7654, Institut Polytechnique de Paris, Palaiseau 91128, France

*Correspondence should be addressed to A.T.P. (phantuan@ntu.edu.sg).

Supplementary Information

Experimental Procedures

DNA sample preparation
DNA samples were purchased from Integrated DNA Technologies. The samples were dissolved in a buffer solution containing 20 mM KPi (pH 7), 70 mM KCl, 10% D_2O, 20 μM DSS to reach a concentration of ~100 μM. They were further annealed by heating them at 95 ^{0}C for 5 minutes in a water bath and letting them cool down slowly to room temperature.

Nuclear Magnetic Resonance (NMR) spectroscopy
NMR experiments were performed on a 600 MHz Bruker AVANCE II spectrometer at 25 ^{0}C, unless otherwise stated. DNA concentration was 100-200 μM. All spectra were processed and analysed using TopSpin 4.0.6.

Circular Dichroism (CD) spectroscopy
JASCO-815 spectropolarimeter was used to measure the CD spectra of DNA samples. The previously annealed DNA samples were diluted to a final concentration of ~4-5 μM in the same buffer solution containing 20 mM KPi (pH 7), 70 mM KCl, 10% D_2O, 20 μM DSS. 500 μl of the diluted DNA sample was transferred to a 1-cm path-length cuvette and CD spectrum was recorded at 25 ^{0}C with 5 or 10 accumulations and with baseline subtraction of the signal contributed by the buffer. Each spectrum was further corrected to be zero at the wavelength of 320 nm.

Crystallization and X-ray structure determination
1 mM solution of *2xMotif2* folded in 10 mM potassium cacodylate buffer and 100 mM potassium chloride (pH 6.5) were used for crystallization assays with the help of sitting-drop methods. The initial screening of crystal formation was done using Natrix 2 screen from Hampton Research using two (DNA:screening buffer) ratios of 1:1 and 1:2. Crystals were obtained in the condition containing 0.08 M Strontium Chloride hexahydrate, 0.04 M Sodium cacodylate trihydrate pH 6.0, 35% v/v (+/-)-2-Methyl-2,4-pentanediol (MPD) and 0.012 M Spermine tetrahydrochloride in the 1:2 ratio drop. This condition was further optimized manually, and rectangular plate shaped crystals were obtained in the similar condition but with 50% v/v (+/-)-2-Methyl-2,4-pentanediol (MPD). These crystals were fished and directly flash-frozen in liquid nitrogen. Diffraction data were collected at the MX1 Beamline of the Australian Synchrotron equipped with an Eiger detector. The diffraction data were processed with autoproc (including STARANISO) from Global Phasing Ltd.[1] Although the data collected had resolution up to 1.69 Å, the completeness of the data at resolution above 2.0 Å was considerably low. Hence, the structure was solved initially using molecular replacement in Phenix with data up to 2.0 Å resolution which had reasonable completeness. The model was further built manually in COOT[2] and was subjected to several cycles of refinement in Phenix[3] with resolution up to 2.0 Å. Then, inclusion of all the data up to 1.69 Å resolution clearly improved the map quality (Figure S4). Further refinements after adding waters were done using pdb redo.[4] The last steps of refinement were done in phenix.refine[5] with the TLS parameters and the final map was generated using the data up to the highest resolution 1.69 Å**.**

Gel electrophoresis
DNA samples were annealed in a buffer solution containing 20 mM KPi (pH 7.0), 70 mM KCl, 10% D_2O, 20 μM DSS at a concentration of ~100 μM as mentioned previously. Sucrose was added to the samples to reach 8 % final concentration. The samples were loaded into 10-20 % native gels (37.5:1 acrylamide/bis-acrylamide) containing 1x TBE supplement. The gels were run in 1x TBE, supplemented with 1 mM KCl at 25 ^{0}C, 80 V, for 95 minutes. DNA bands were visualized by UV shadowing in AlphaImager HP.

UV melting
UV melting experiments were conducted on a JASCO V-650 spectrophotometer. 3-5 μM of samples were taken in a cuvette of a pathlength of 1 cm and used for measuring the melting data. The samples were heated from 25 to 90 ^{0}C and subsequently cooled from 90 to 25 ^{0}C at a rate of 0.1 ^{0}C/min. The UV absorption of the samples were measured at 295 nm and 320 nm at every 0.5 ^{0}C as the samples were heated and cooled. The data collected at 295 nm were subtracted from the 320 nm for background correction and further normalized. The melting temperatures (T_m) were determined from the normalized melting curves where the DNA folded fraction was 50%.

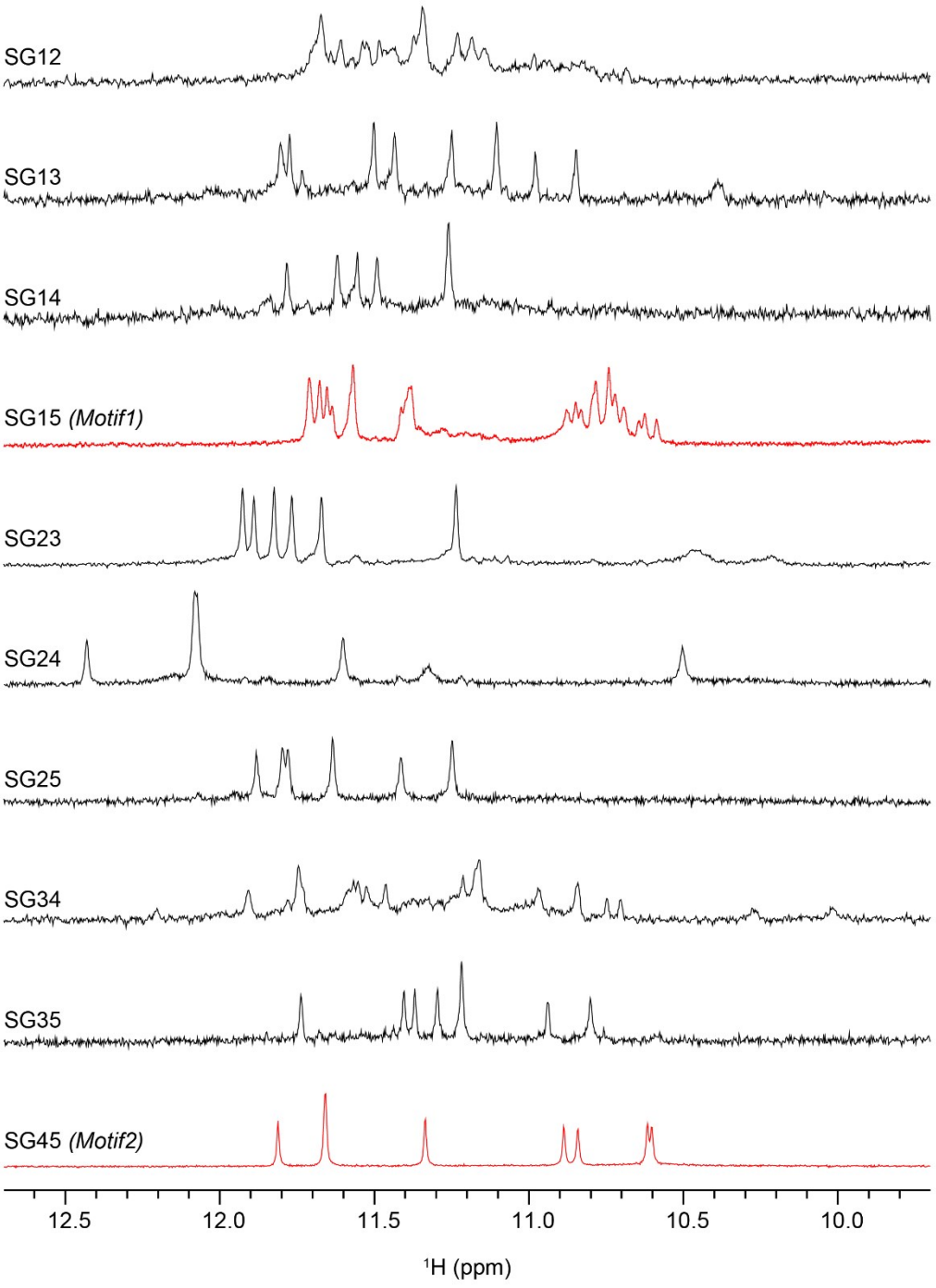

Figure S1: 1D imino proton NMR spectra of all the designed 12-nt DNA sequences to test the possible formation of a left-handed G-quadruplex.

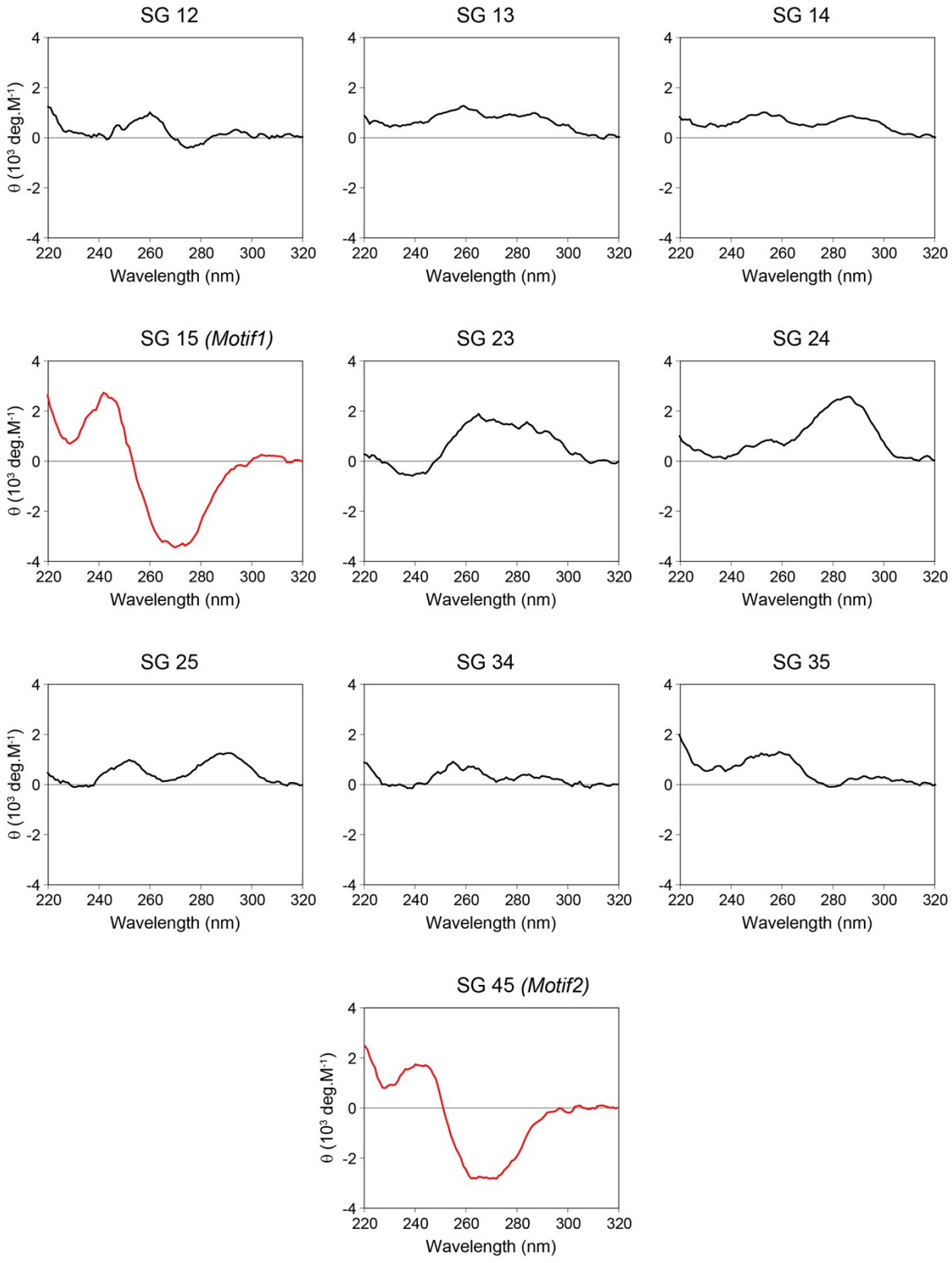

Figure S2: CD spectra of all the designed 12-nt DNA sequences to test the possible formation of a left-handed G-quadruplex.

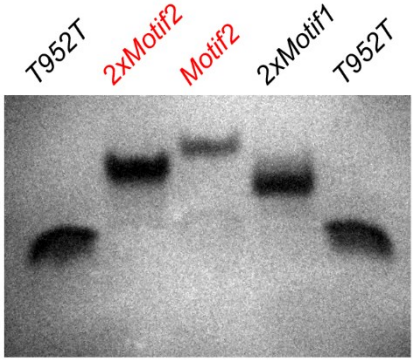

Figure S3: Gel electrophoresis or electrophoretic mobility shift assay (EMSA) of the DNA sequences *2xMotif2*, *Motif2* and *2xMotif1*. *T95-2T*, a three-layered right-handed DNA G-quadruplex, was used as a reference.

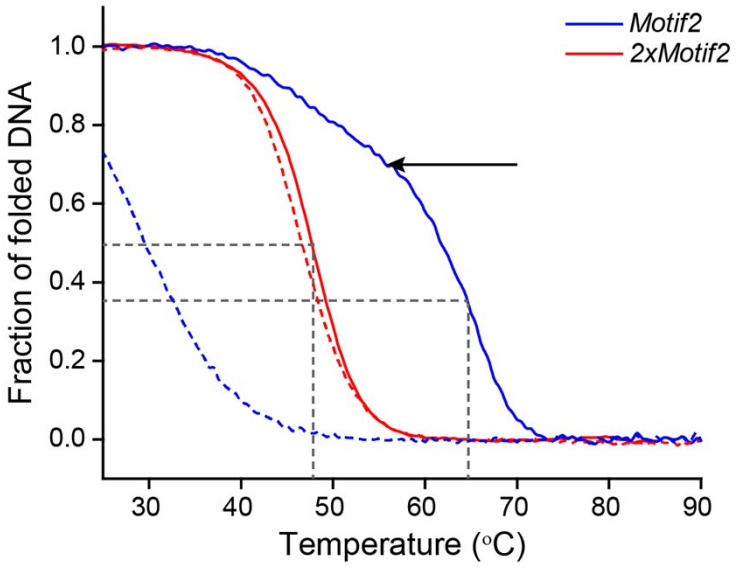

Figure S4: UV melting and cooling profiles of the sequences *Motif2* and *2xMotif2*. Solid and dashed lines indicate heating and cooling profiles, respectively. *Motif2* displays two-step melting profile and the arrow indicates the approximate starting of the second step melting of *Motif2*. The dashed grey lines indicate the melting temperatures (T_m) which are obtained from the heating curves for comparison.

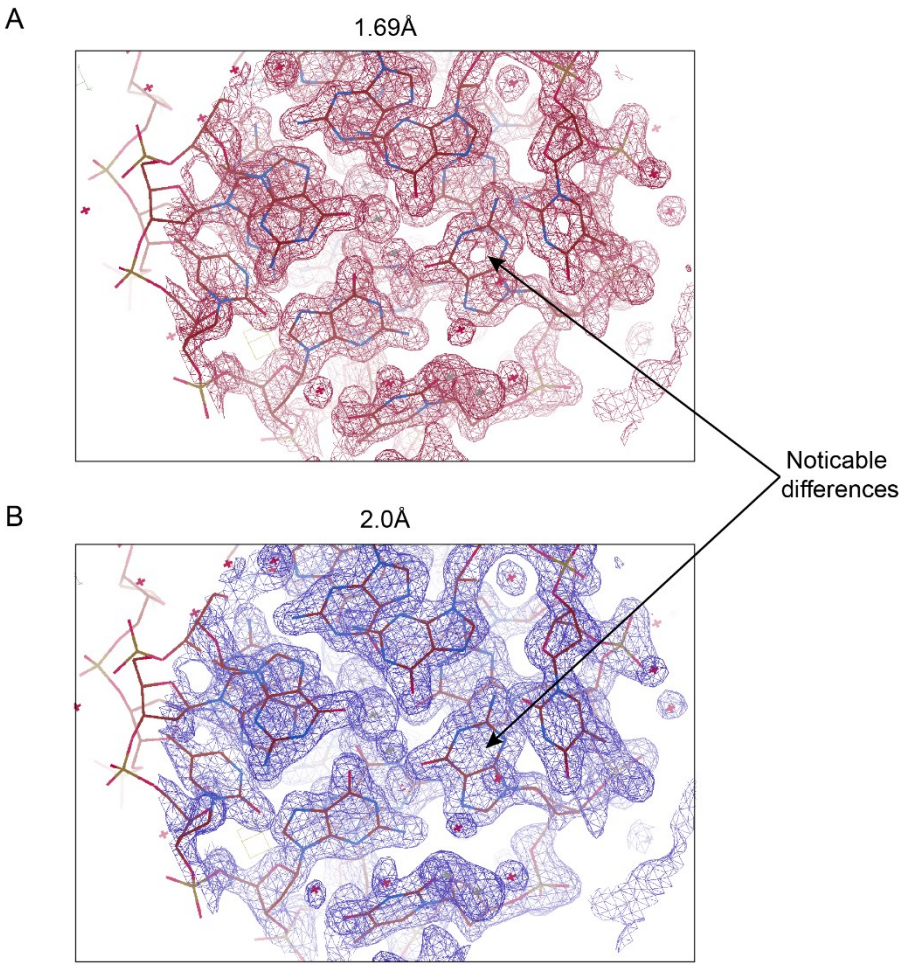

Figure S5: Comparison of electron density map of a part of the *2xMotif2* crystal structure calculated at (A) 1.69 Å and (B) 2.0 Å, showing the improvement of map quality with the inclusion of all the data up to 1.69 Å resolution.

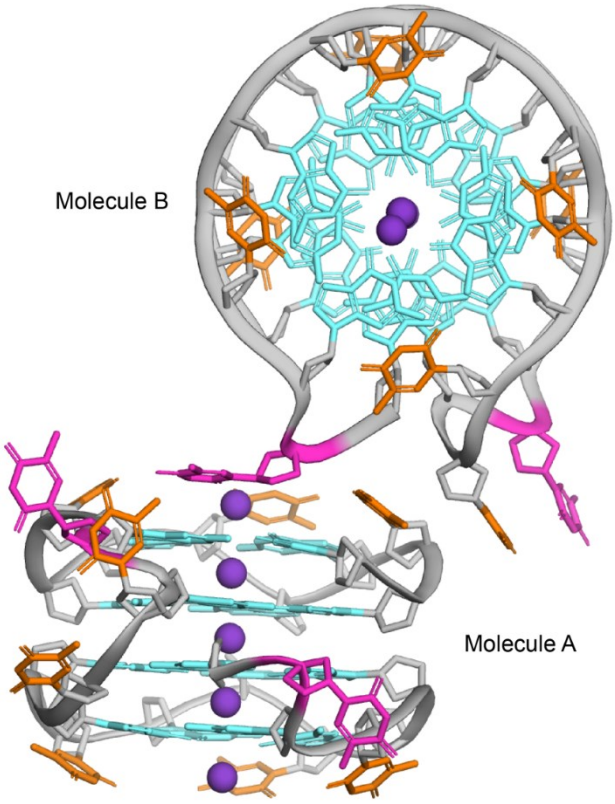

Figure S6: Two crystal structures of *2xMotif2* named as Molecule A and Molecule B found in a unit cell connected with non-crystallographic symmetry (NCS). The following colour code is used: nucleotide backbones in grey, G-tetrad guanines in cyan, capping and linker thymines in orange, bulge thymines in magenta, and potassium ions in purple.

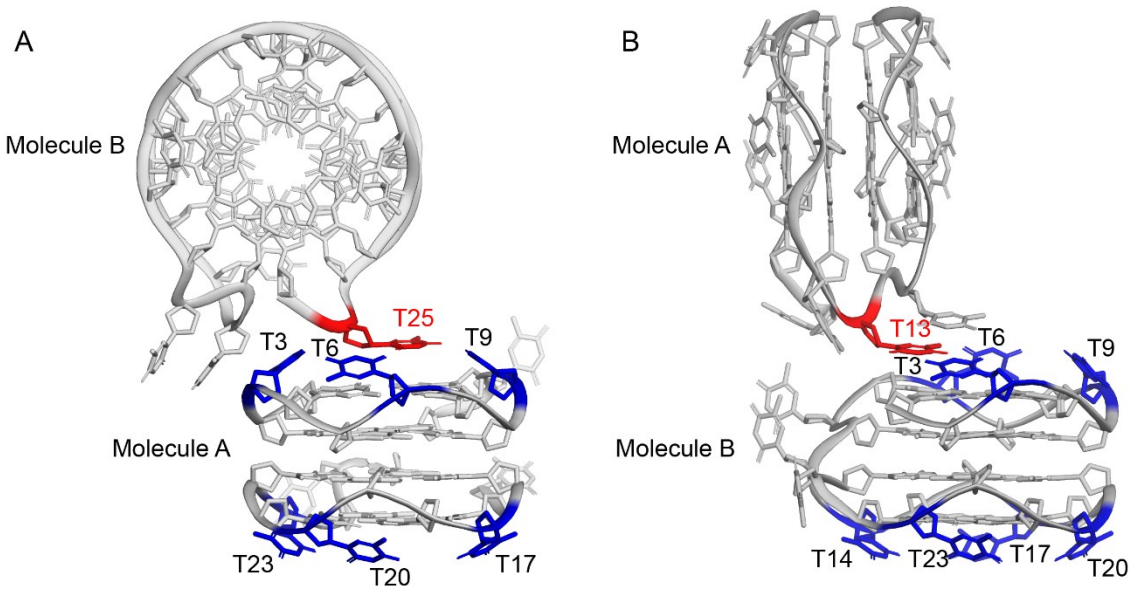

Figure S7: (A) T-cappings (in blue) of Molecule A are highlighted with an additional capping by T25 (in red) of Molecule B on the upper tetrad due to crystal packing. (B) T13 of Molecule A (in red) acts as an extra capping on the upper tetrad of Molecule B apart from the rest of the T-cappings from Molecule B (in blue).

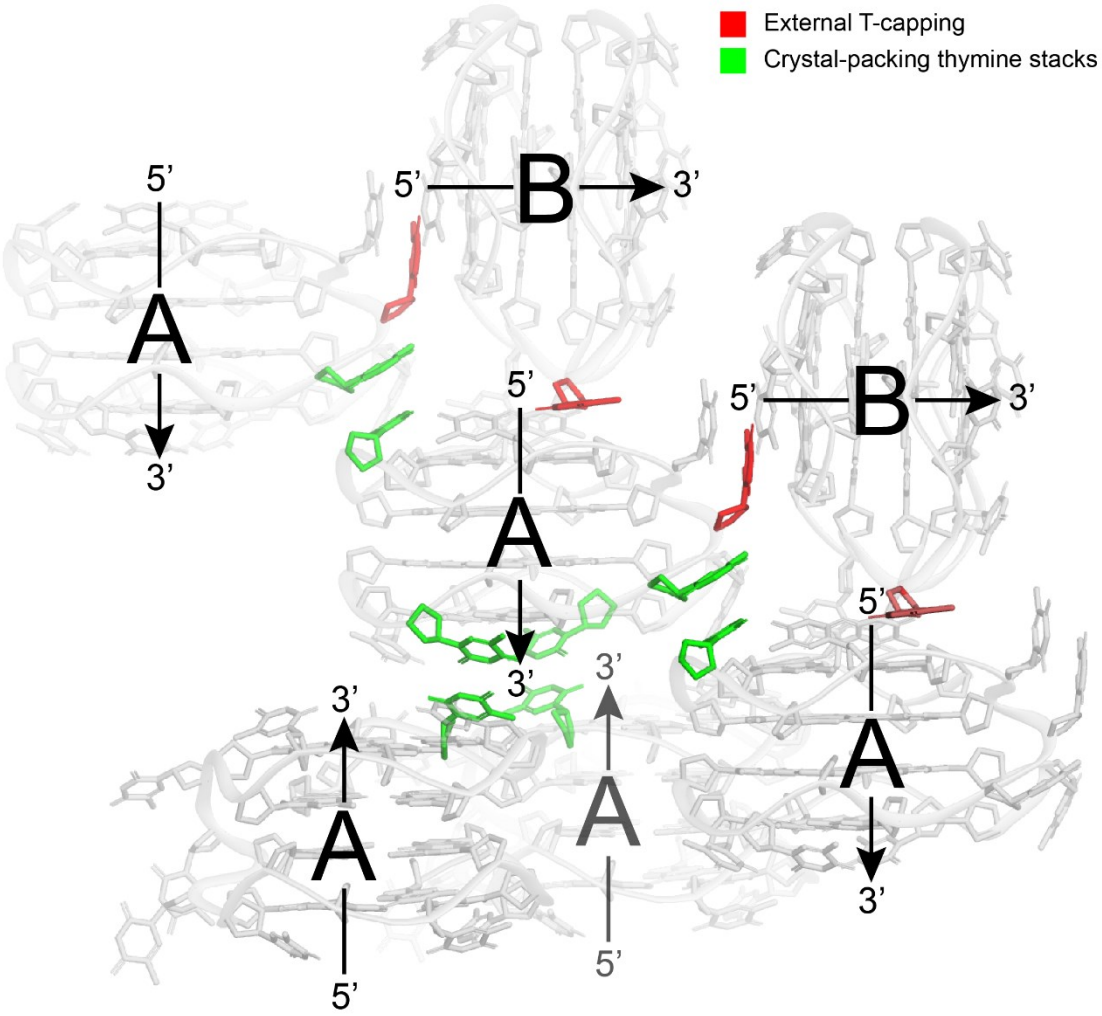

Figure S8: Crystal packing of the *2xMotif2* structure. The stacking of thymine residues on each other are highlighted in green, while the stacking of thymine residues on the upper tetrads acting as T-cappings are highlighted in red. Molecules A and B are indicated, while the arrows showed the direction from the 5’-block to the 3’-block of each molecule.

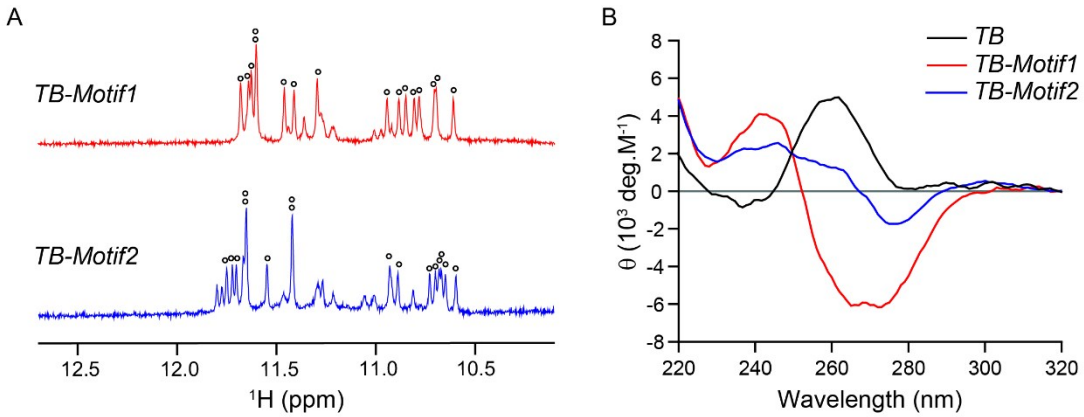

Figure S9: (A) Comparison of 1D ^{1}H NMR spectra of *TB* attached to *Motif1* and *Motif2* respectively, showing the formation of left-handed G4 species predominantly in *TB-Motif1* and formation of similar left-handed species in *TB-Motif2* with additional minor species. The peaks representing the left-handed G4 conformation are marked with open circles. The peaks for the left-handed G4 of *TB-Motif2* are marked based on the homology with *TB-Motif1*. (B) CD spectra of *TB*[6], *TB-Motif1* and *TB-Motif2* displaying the conversion of a right-handed G4 conformation of *TB* to left-handed G4 conformation when it is attached to *Motif1* and *Motif2*. *Motif2* has lower ability of LHG4 structural conversion as compared to *Motif1*.

Table S1. Sequences for investigation of LHG4 structural conversion ability of *Motif1* and *Motif2*. The split-guanines are marked in bold.

Name	Sequence (5'-3')
TB	GG TGG TGG TGG
TB-Motif1 (ZG4)	T GG TGG TGG TGG TT **G**TG GTG GTG GT**G** TT
TB-Motif2	GG TGG TGG TGG TT GG TGG TGG T**G**T**G**

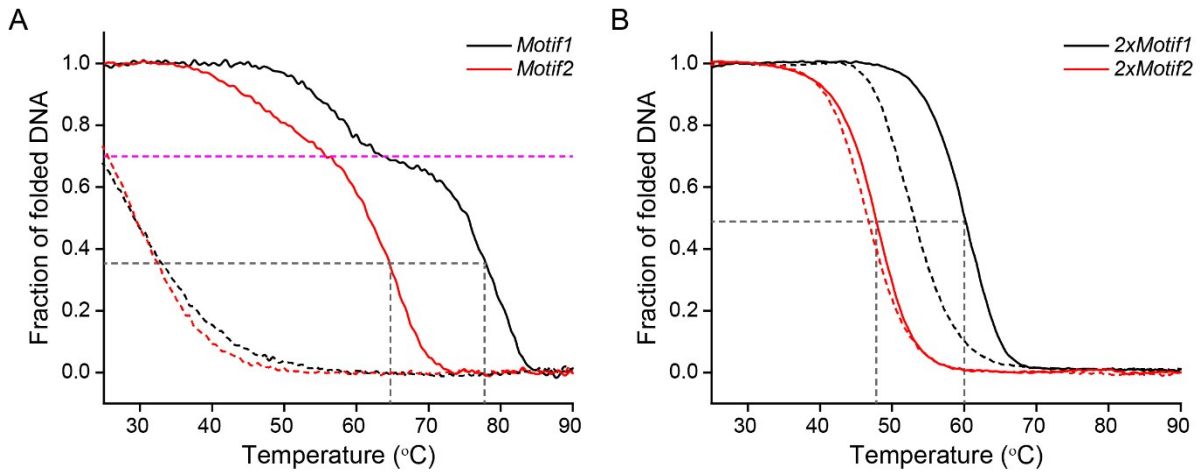

Figure S10: UV melting and cooling of the sequences (A) *Motif1* and *Motif2* and (B) *2xMotif1* and *2xMotif2*. The solid and dashed lines indicate heating and cooling profiles, respectively. Both sequences *Motif1 and Motif2* that form dimeric G4s display two-step melting profile and the magenta dashed line in (A) indicates the approximate starting of the second step melting. The grey dashed lines indicate the melting temperatures (T_m) which are obtained from the heating curves for comparison.

REFERENCES

1. Tickle, I. J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., Bricogne, G. , STARANISO **2018**.
2. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 4), 486-501.
3. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.
4. Joosten, R. P.; Long, F.; Murshudov, G. N.; Perrakis, A., The PDB_REDO server for macromolecular structure model optimization. *IUCrJ* **2014,** *1* (Pt 4), 213-20.
5. Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D., Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **2012,** *68* (Pt 4), 352-67.
6. Bakalar, B.; Heddi, B.; Schmitt, E.; Mechulam, Y.; Phan, A. T., A Minimal Sequence for Left-Handed G-Quadruplex Formation. *Angew Chem Int Ed Engl* **2019,** *58* (8), 2331-2335.