Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Adenine-Driven Structural Switch from a Two- to Three-Quartet DNA G-Quadruplex

*Martina Lenarčič Živković, Jan Rozman, and Janez Plavec**

anie_201809328_sm_miscellaneous_information.pdf

Supporting Information

Table of Contents

1. Materials and Methods

Sample preparation

All isotopically unlabeled and residue-specific $^{15}N/^{13}C$-labeled oligonucleotides were synthesized on K&A Laborgeraete GbR DNA/RNA Synthesizer H-8 using standard phosphoramidite chemistry in DMT-off mode with the exception of Ran4A5T oligonucleotides that were synthesized in DMT-on mode. Oligonucleotides synthetized in DMT-off mode were cleaved from the solid support by treatment with concentrated aqueous ammonia at 55 °C overnight. Oligonucleotides synthesized in DMT-on mode were purified using reverse-phase HPLC chromatography, followed by removal of DMT group with reaction in 80% AcOH for 30 min. Oligonucleotides were then extracted with pure ethanol and 8 M LiCl. Samples were after rinsing with 70% ethanol dissolved in pure water. The pH value of samples was adjusted to around 7 using LiOH. Samples were subsequently desalted using ultrafiltration cell with 1 kDa cellulose membrane. Oligonucleotides synthesized in DMF-off mode were desalted using 3K Amicon® Ultra-15 centrifugal filter units. Concentration of desalted samples was determined by UV absorption at 260 nm using UV/VIS Spectrophotometer Varian CARY-100 BIO UV–VIS. Extinction coefficients were determined by nearest neighbour method. 10% of ^2H_2O was added and the pH of the samples was adjusted to 7.0 using LiOH prior to heating them at 90 °C for 5 min. After heating, KCl and potassium phosphate buffer with pH 7.0 were added to the samples, followed by quick cooling on ice.

CD and UV experiments

Circular dichroism (CD) spectra were recorded on an Applied Photophysics Chirascan CD spectrometer at 25 °C using a quartz cuvette with a 1.0 mm path length. The wavelength range was from 220 to 320 nm. CD samples were prepared at 30 μM oligonucleotide concentrations in 70 mM KCl and 15 mM potassium phosphate buffer. A blank sample containing only 70 mM KCl and 15 mM potassium phosphate buffer was used for baseline correction. Three scans were averaged to obtain each CD spectrum.

UV melting experiments were performed on a Varian CARY-100 BIO UV–VIS spectrophotometer with the Cary Win UV Thermal program using a 0.5 cm (for 53 μM sample) and 1.0 cm (for lower concentrations) pathlength cells. Ran4 samples for UV measurements were prepared at 3, 13, 26 and 53 μM oligonucleotide concentrations in 70 mM KCl and 15 mM potassium phosphate buffer. Samples of modificated oligonucleotides were prepared at 15 μM oligonucleotide concentration in 70 mM KCl and 15 mM potassium phosphate buffer. Samples were heated at 0.1 °C min^{-1} from 10 °C to 90 °C and absorbance at 295 nm was measured at 0.5 °C steps. Mineral oil was used to prevent evaporation and sample loss due to high temperatures. A stream of nitrogen was applied throughout the measurements to prevent condensation at lower temperatures. T_m was determined from the A_{295} versus temperature plot.

NMR experiments and structure calculations

NMR experiments were recorded on Agilent-Varian VNMRS 600 MHz and 800 MHz spectrometers at 5 °C and 25 °C. Double-pulsed field gradient spin echo (DPFGSE) pulse sequence was used to suppress the water signal. Identification of guanine imino and H8 protons in residue-specific 6% ^{15}N, ^{13}C-labeled samples of Ran4 was performed by 1D ^{15}N-edited and ^{13}C-edited HSQC NMR experiments, respectively. Adenine H8 and H2 protons were identified by performing 2D ^{13}C-HSQC experiments. Exchangeable and non-exchangeable proton resonances of Ran4 were assigned using 2D NOESY spectra (mixing times of 50, 100, 150, 200, 250, 300 ms) recorded in 90% H_2O/10% ^2H_2O. The 2D DQF-COSY and TOCSY (mixing time of 40 ms) spectra were used to estimate sugar conformations. Resonance assignment and topology determination of Ran4A5T were achieved using 2D NOESY (mixing times of 50 and 250 ms) and TOCSY (mixing time of 40 ms). Imino and H8 protons of 5 guanines in *syn* glycosidic conformation in Ran4A5T (*i.e.* G1,

G6, G12, G13 and G18) were identified by performing 1D ^{15}N-edited and 2D ^{13}C-HSQC NMR experiments on residue-specific 6% ^{15}N, ^{13}C-labeled samples.

All spectra were processed in VNMRJ program (Agilent Technologies). Resonance assignment and integration were achieved using SPARKY software (UCSF).[1] NOE distance restraints for non-exchangeable protons were obtained from 2D NOESY spectrum recorded at 5 °C with a mixing time of 150 ms. Only non-overlapping cross-peaks were integrated and used for distance restraints calculation. An average volume of intraresidual H8-H1' NOE correlation was used as the reference volume. Because residues G7, G10, G16, G19 and G20 did not overlap with any other cross-peak and clearly exhibited *anti* glycosidic conformations, we could reference the averaged volume to the value of 3.9 Å. With the help of this reference, we classified the remaining signals as strong (1.8 – 3.6 Å), medium (2.6 – 5.0 Å) and weak (3.5 – 6.5 Å). In this way, all intraresidual H8-H1' NOE correlations of the guanine residues with *anti* glycosidic conformation fell in the class of medium distance restraints. Moreover, H8-H1' cross-peaks of G6, G13 and G18 were classified as strong, confirming their *syn* glycosidic conformation. Although H8-H1' cross-peak of G1 in NOESY spectra at 5 °C was weaker than expected, its *syn* glycosidic conformation was confirmed by NOESY spectrum at 25 °C, where H8-H1' cross-peak of G1 exhibited strong signal. Moreover, downfield chemical shift for C8 carbon atoms of G1, G6, G13 and G18 were all above 140 ppm typical for residues in *syn* glycosidic conformation.

NOE distance restraints for exchangeable protons were obtained from 2D NOESY spectrum recorded at 5 °C with a mixing time of 300 ms. Cross-peaks were classified as medium (2.6 – 5.0 Å) or weak (3.5 – 6.5 Å) based on their intensities. Torsion angle χ along glycosidic bond was restrained to a range between 25 and 95° for *syn* and between 200 and 280° for *anti* guanine residues. Torsion angle χ restrains for all adenine residues were set between 200 and 280° and between 170 and 310° for thymine and cytosine residues describing their *anti* glycosidic conformation. Torsion angle v_2 was used to restrain sugar pucker to C2'-endo conformation for all residues.

Structure calculations were performed with AMBER 14 software[2] using parmbsc0 force field[3] with parmχ_{OL4}[4] and parm$\varepsilon\zeta_{OL1}$[5] modifications. The initial extended structure of Ran4 oligonucleotide was obtained with LEAP module of the AMBER 14 program. A total of 100 structures were calculated in 120 ps of NMR restrained simulated annealing (SA) simulations using the generalized Born implicit model to account for solvent effects. The cut-off for non-bonded interactions was 20 Å and the SHAKE algorithm for hydrogen atoms was used with the tolerance of 0.0005 Å. SA calculations were initiated with random velocities. After 5 ps at 300 K, the temperature was raised to 1000 K in the next 10 ps and held constant at 1000 K for 30 ps. Temperature was scaled down to 100 K in the next 45 ps and reduced to 0 K in the last 30 ps. Restraints included in the calculation involved NOE-derived (force constant 50 kcal mol^{-1} A^{-2}) and hydrogen bond distance restraints (force constant 50 kcal mol^{-1} A^{-2}), torsion angle χ and v_2 restraints (force restraints 200 kcal mol^{-1} A^{-2}) and planarity restraints between residues constituting G-quartets and base-triads (force constant 20 kcal mol^{-1} A^{-2}). Planarity restraints were excluded in the last 30 ps of SA. A family of 10 structures was selected based on the lowest energy and subjected to energy minimization with a maximum of 20 000 steps. Figures were visualized and prepared with UCSF Chimera software.[6] The coordinates of the Ran4 G-quadruplex have been deposited in the Protein Data Bank with the accession code 6GZN. Chemical shifts list has been deposited in the Biological Magnetic Resonance Data Bank with the accession code 34297.

2. Results

Table S1. NMR and structure statistics.

NMR restraints	
NOE-derived distance restraints	
Total	364
Intra-residual	221
Inter-residual	143
Sequential	100
Long-range	43
Hydrogen bond restraints	23
Torsion angle restraints	34
Planarity restraints	36
Structure Statistics	
Violations	
Mean NOE restraint violation (Å)	0.053 ± 0.011
Max. NOE restraint violation (Å)	0.069
Max. torsion angle restraint violation (°)	0
Deviations from idealized geometry	
Bonds (Å)	0.0121 ± 0.0001
Angles (°)	2.40 ± 0.03
Pairwise heavy atom RMSD (Å)	
Overall	1.33 ± 0.35
G-quartets and base-triads	0.58 ± 0.15

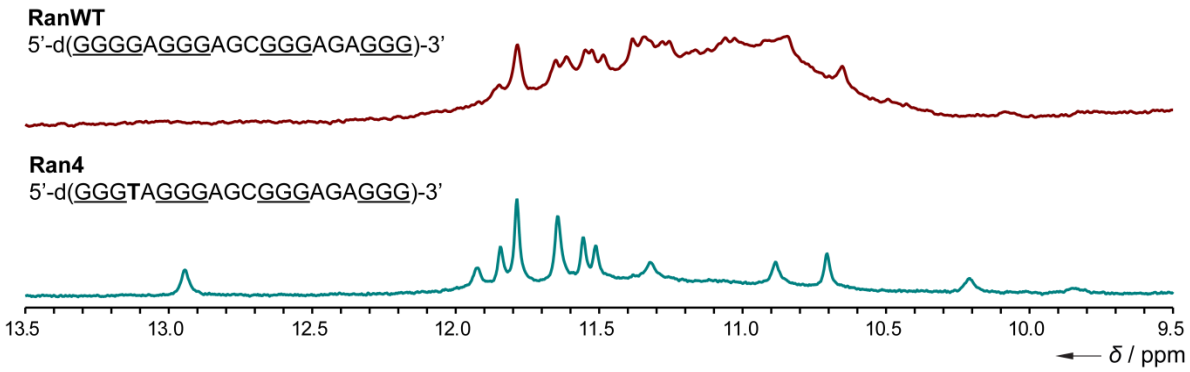

Figure S1. Imino region of 1D ^{1}H NMR spectra of wild type G-rich sequence from regulatory region of RANKL gene (RanWT) and its G4-to-T4 modified sequence (Ran4). Spectra were recorded on Agilent DD2 600 MHz spectrometer at 5 °C in 90% H_2O, 10% $^{2}H_2O$, 70 mM KCl, 15 mM potassium phosphate buffer with pH 7.0. Oligonucleotide concentrations were 0.6 mM.

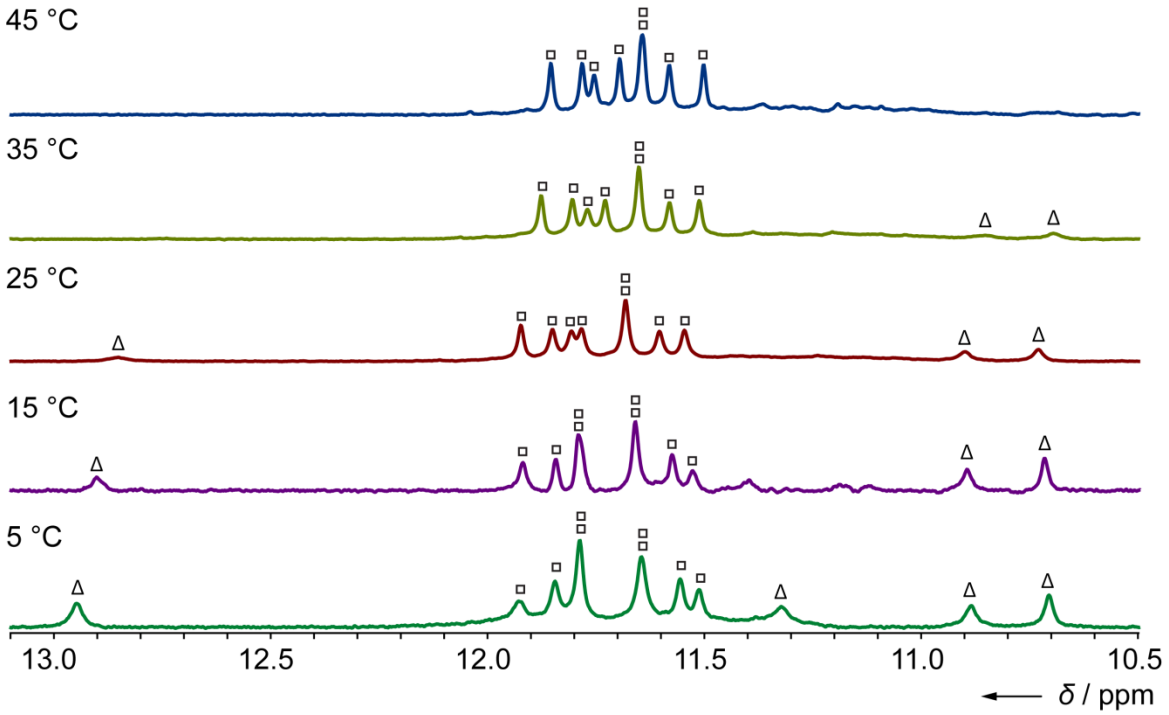

Figure S2. Imino region of 1D ^{1}H NMR spectra of Ran4 at different temperatures. Signals of imino protons belonging to residues from G-quartets are marked with squares, while imino protons of base-triad-forming residues are marked with triangles. Spectra were recorded on Agilent DD2 600 MHz spectrometer in 90% H_2O, 10% $^{2}H_2O$, 70 mM KCl and 15 mM potassium phosphate buffer with pH 7.0. Oligonucleotide concentration was 0.6 mM.

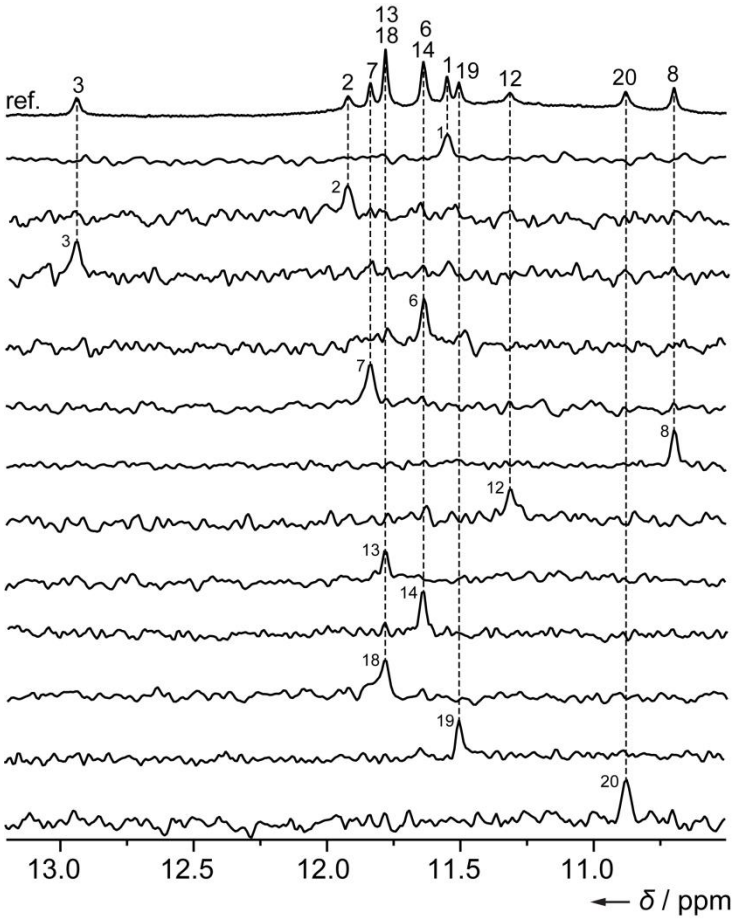

Figure S3. Unambiguous assignment of imino proton resonances of Ran4 was achieved by recording 1D ^{15}N-edited HSQC spectra on partially 6% residue-specific ^{15}N/^{13}C-labeled oligonucleotides of Ran4. Imino region of 1D ^{1}H NMR spectrum of Ran4 (ref.) and assignment of imino resonances are shown on top. Spectra were recorded on Agilent DD2 600 MHz spectrometer at 5 °C in 90% H_2O, 10% $^{2}H_2O$, 70 mM KCl, 15 mM potassium phosphate buffer with pH 7.0. Oligonucleotide concentrations were 0.6 mM.

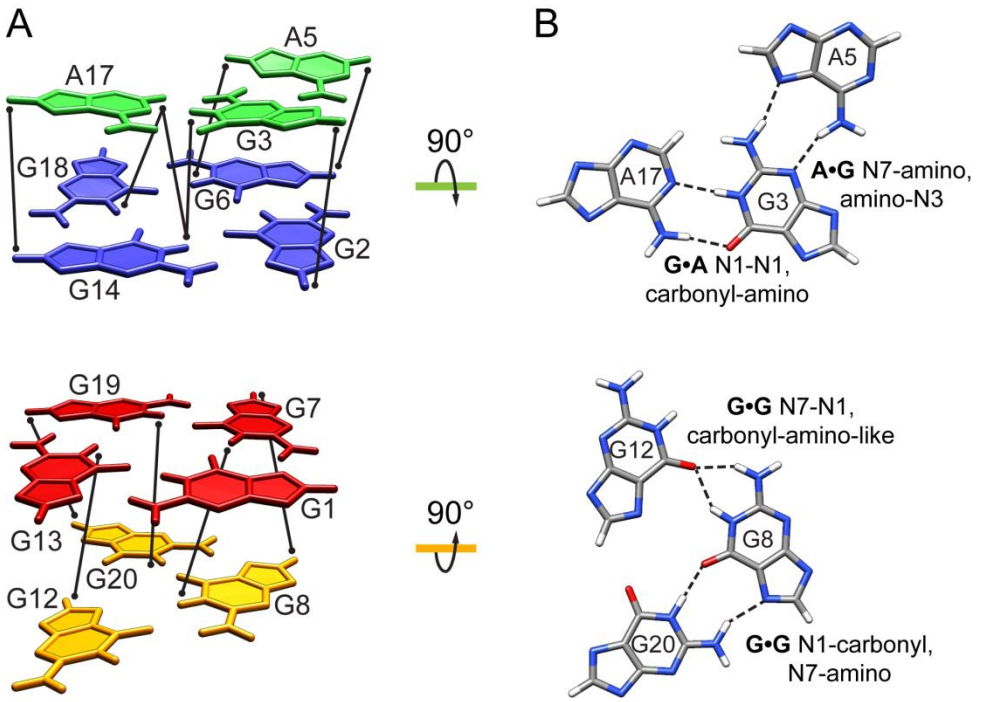

Figure S4. (A) NOE contacts between aromatic and imino protons of base-triads-forming residues and residues from G-quartets of Ran4 G-quadruplex. Residues comprising A5•G3•A17 base-triad, G2•G6•G18•G1 quartet, G1•G13•G19•G7 quartet and G20•G8•G12 base-triad are coloured green, blue, red and orange, respectively. (B) Base-pairing arrangements within both base-triads. Hydrogen bonds are depicted as black dotted lines.

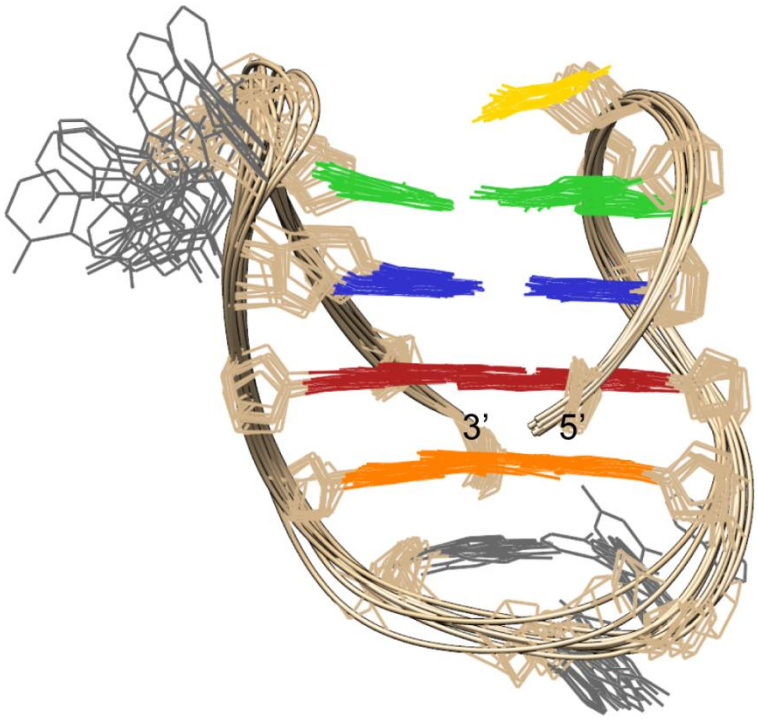

Figure S5: An ensemble of 10 lowest energy structures of Ran4 G-quadruplex. G1•G13•G19•G7 and G2•G6•G18•G14 quartets are shown in red and blue, respectively. A5•G3•A17 and G20•G8•G12 base-triads are presented in green and orange, respectively. T4 is shown in yellow, while other loop residues are coloured grey.

In addition to G-quartets and base-triads, T4 is well defined and stacked above A5•G3•A17 base-triad. In contrast, residues A15 and G16 are exposed to solvent and represent the most dynamic part of the molecule. Interestingly, residues constituting diagonal loop are less flexible. A9 is with its H2 proton oriented towards the G20•G8•G12 base-triad. Purine moiety of G10 is stacked on A9 as evidenced by NOE connectivities between H8 protons of A9 and G10. Its sugar moiety is oriented towards C11 which is supported by NOEs between sugar protons of G10 and H5/H6 protons of C11. C11 is positioned under the G20•G8•G12 base-triad and stacked on G20.

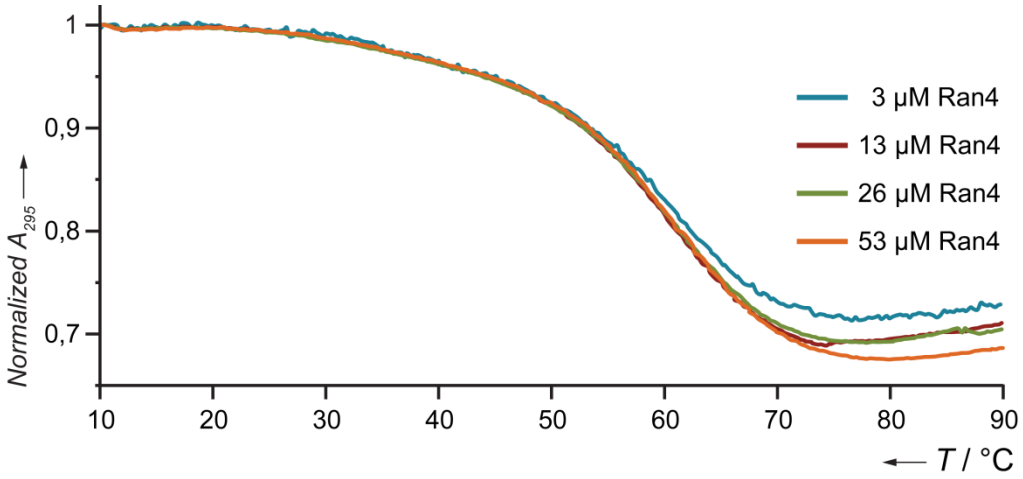

Figure S6. UV melting temperature experiments at different Ran4 oligonucleotide concentrations. Concentration-independent melting temperature is consistent with the unimolecular structure of the Ran4 G-quadruplex.

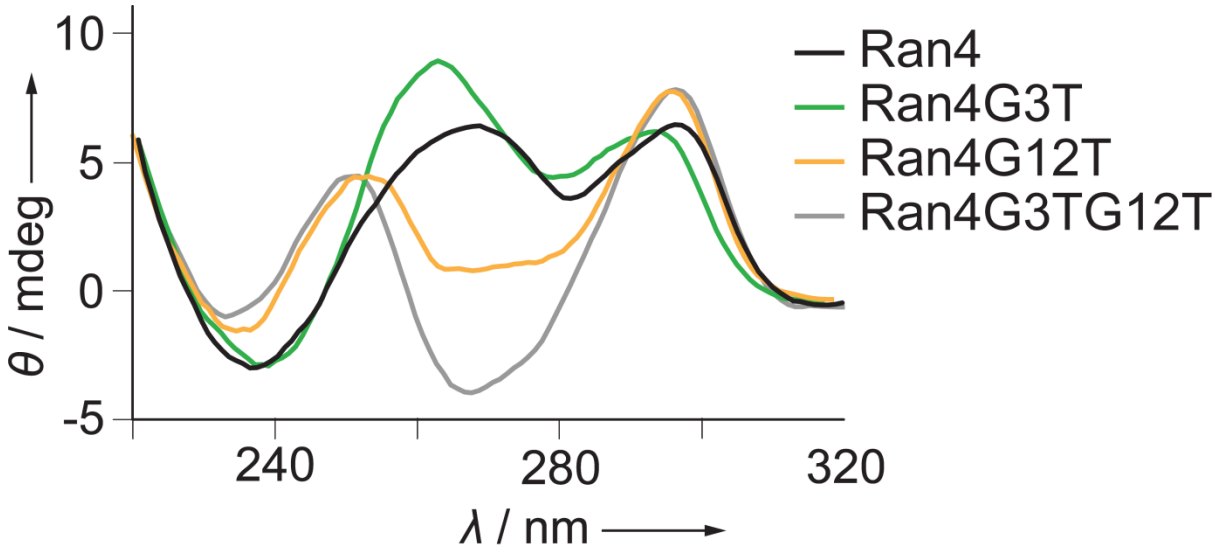

Figure S7. Comparison of CD spectra of Ran4 and modified sequences Ran4G3T, Ran4G12T and Ran4G3TG12T. All spectra were recorded at 25 °C in 70 mM KCl, 15 mM potassium phosphate buffer with pH 7.0. Oligonucleotide concentrations were 30 μM.

CD spectrum of Ran4 G-quadruplex shows a minimum at ~235 nm and two maxima at 265 and 295 nm, which are characteristic for antiparallel fold with extensive stacking of loop residues.[7,8] On the other hand, CD spectrum of Ran4G3TG12T displays a minimum at ~265 nm and two maxima at ~250 and 295 nm typical for antiparallel topologies without extra G-quartet stacking interactions[9]. While CD profile of Ran4G3T is similar to that of Ran4, the distribution of signals in CD spectrum of Ran4G12T resembles Ran4G3TG12T with almost superimposable maxima at ~250 and 295 nm. However, shallow minimum that spans between 260 and 280 nm in CD spectra of Ran4G12T suggests that, while stacking in Ran4G12T involves guanines with the same type of glycosidic conformations, Ran4G3TG12T is characterized by *syn-anti* glycosidic conformations along all G-strands.[8] In agreement, stacking of consecutive G2 and G3 that both adopt *anti* glycosidic conformations is observed in Ran4G12T. Moreover, similar melting temperatures of Ran4G12T (T_m 48.0 °C) and Ran4G3TG12T (T_m 50.0 °C) imply bigger contribution of G20•G8•G12 in comparison to A5•G3•A17 base-triad (Tm of Ran4G3T 54.4 °C) to the thermal stability of Ran4 (Table 1). Although G3 and G12 play important role in formation of base-triads, our results show they may not be critical for formation of the two-quartet G-quadruplex adopted by Ran4.

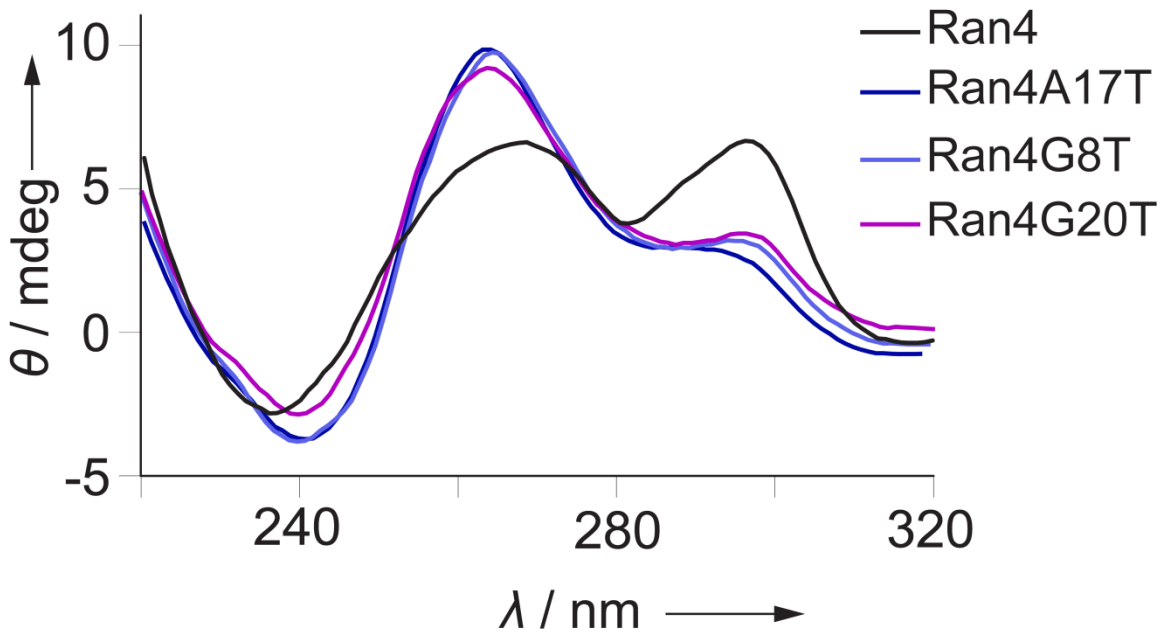

Figure S8. Comparison of CD spectra of Ran4 and modified sequences Ran4A17T, Ran4G8T and Ran4G20T. All spectra were recorded at 25 °C in 70 mM KCl, 15 mM potassium phosphate buffer with pH 7.0. Oligonucleotide concentrations were 30 μM.

CD spectra for all the three oligonucleotides display minimum at 240 nm, maximum at 260 nm and shoulder from 280 to 300 nm. Such distribution of CD signals is together with weak overlapping signals in imino region of ^{1}H NMR spectra (Figure 3A) typical for polymorphic samples with contribution of several DNA conformations.

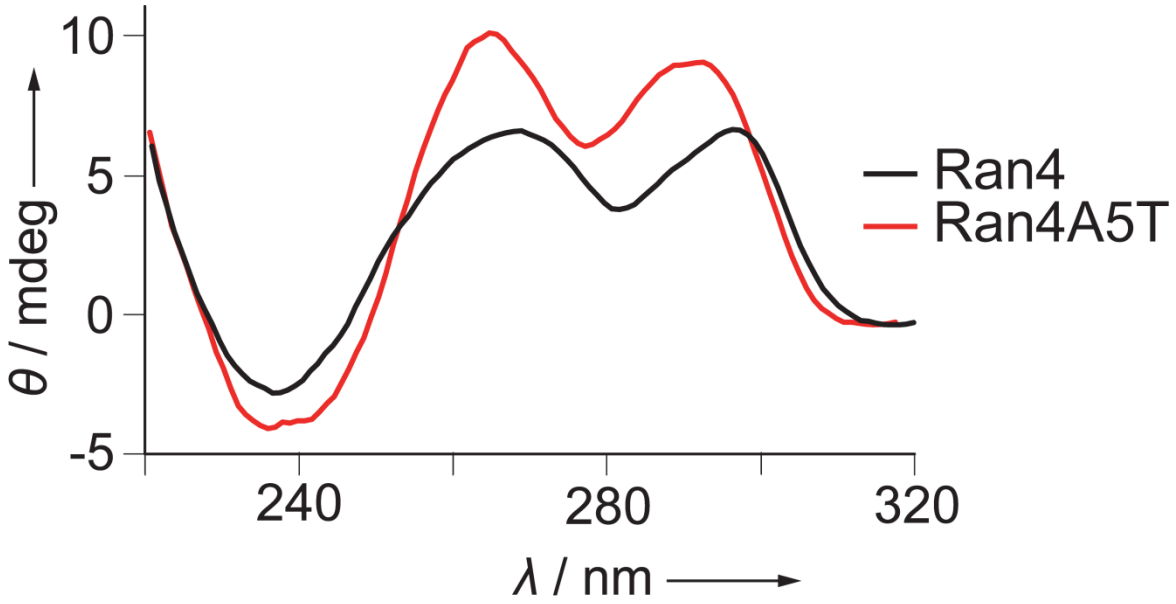

Figure S9. Comparison of CD spectra of Ran4 and modified sequence Ran4A5T. Spectra were recorded at 25 °C in 70 mM KCl, 15 mM potassium phosphate buffer with pH 7.0. Oligonucleotide concentrations were 30 μM.

(3+1) hybrid G-quadruplex topology adopted by Ran4A5T is in accord with CD profile with minimum in ~235-242 nm range and two maxima at ~265 and 290 nm.

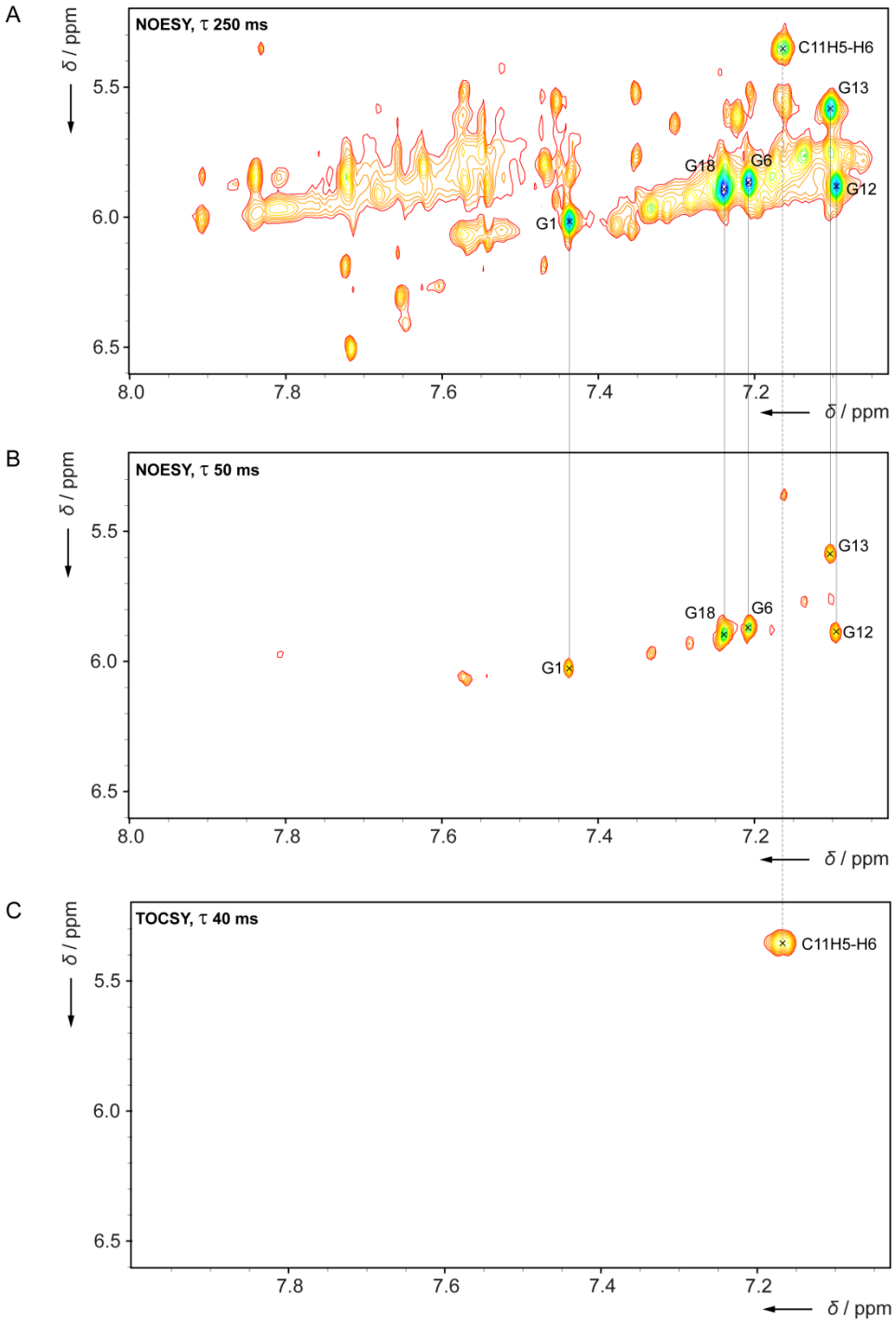

Figure S10. Anomeric-aromatic regions of NOESY NMR spectra of Ran4A5T with mixing times of 250 ms (A) and 50 ms (B) and TOCSY NMR spectrum with mixing time of 40 ms (C). Spectra were recorded on Agilent 800 MHz spectrometer at 25 °C in 90% H_2O, 10% ^2H_2O, 70 mM KCl, 15 mM potassium phosphate buffer with pH 7.0. Oligonucleotide concentration was 1.5 mM.

NOESY spectrum with mixing time of 250 ms (A) shows 6 intense signals in the anomeric-aromatic region. Five of them belong to guanines with *syn* glycosidic conformations (B), while sixth signal corresponds to C11 H5-H6 (C). Unambiguous assignment of aromatic H8 protons of G1, G6, G12, G13 and G18 was achieved by 2D ^{13}C-HSQC spectra acquired on residue-specific 6% ^{15}N/^{13}C-labeled oligonucleotides (Figure S9).

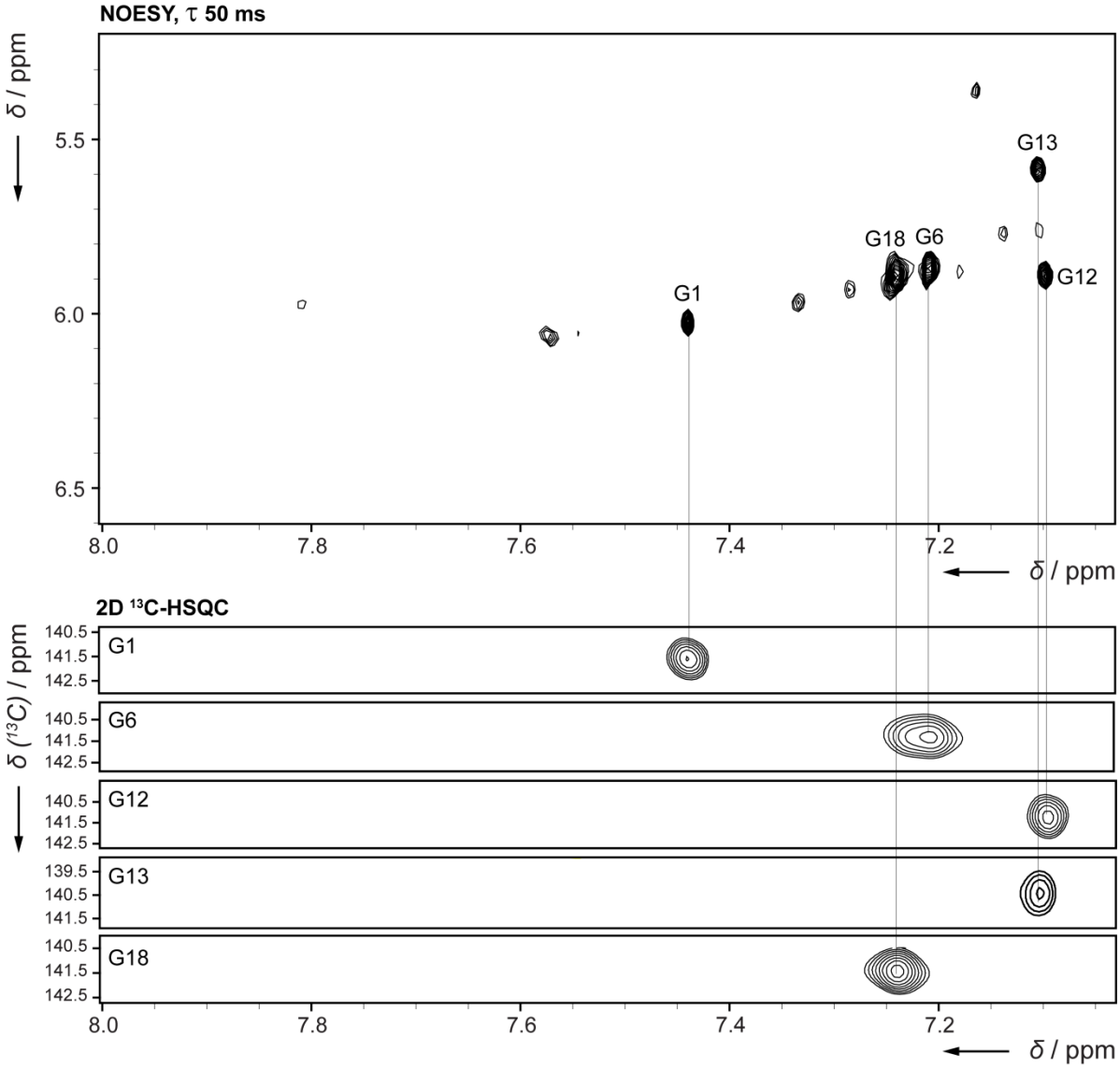

Figure S11. Unambiguous assignment of aromatic H8 protons of guanines of Ran4A5T in *syn* glycosidic conformations was achieved by recording 2D ^{13}C-HSQC NMR spectra on partially 6% ^{15}N/^{13}C-labeled oligonucleotides (below). ^{13}C-HSQC spectra were aligned to 2D NOESY spectrum with mixing time of 50 ms (above) for easier comparison. Chemical shift of C8 carbon atoms of G1, G6, G12, G13 and G18 in ^{13}C-HSQC spectra (δ_C ~140 ppm) corroborate their *syn* glycosidic conformations. All spectra were recorded on Agilent 800 MHz spectrometer at 25 °C in 90% H_2O, 10% ^2H_2O, 70 mM KCl and 15 mM potassium phosphate buffer with pH 7.0. Oligonucleotide concentration of the sample used for NOESY experiment was 1.5 mM, while ^{13}C-HSQC spectra were recorded on samples with 0.6 mM oligonucleotide concentrations.

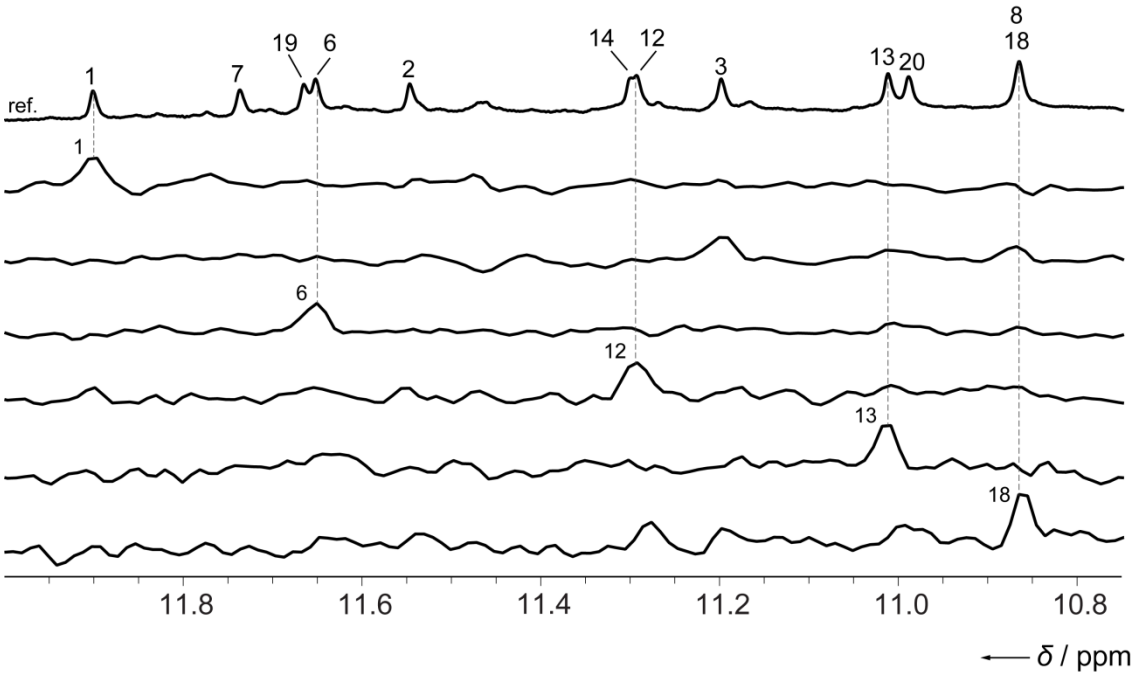

Figure S12. Imino resonances of guanines in *syn* glycosidic conformations (*i.e.* G1, G6, G12, G13 and G18) were unambiguously assigned by recording 1D ^{15}N-edited HSQC spectra on partially 6% residue-specific ^{15}N/^{13}C-labeled oligonucleotides. Imino region of 1D ^{1}H NMR spectrum of Ran4A5T (ref.) and assignment of imino resonances are shown on top. Spectra were recorded on Agilent DD2 600 MHz spectrometer at 25 °C in 90% H_2O, 10% $^{2}H_2O$, 70 mM KCl and 15 mM potassium phosphate buffer with pH 7.0. Oligonucleotide concentrations were 0.6 mM. Imino resonances of other G-quartet-forming guanines (*i.e.* G2, G3, G7, G8, G14, G19 and G20) were assigned based on imino-imino and imino-aromatic NOE connectivities in NOESY spectra of Ran4A5T (Figure S11).

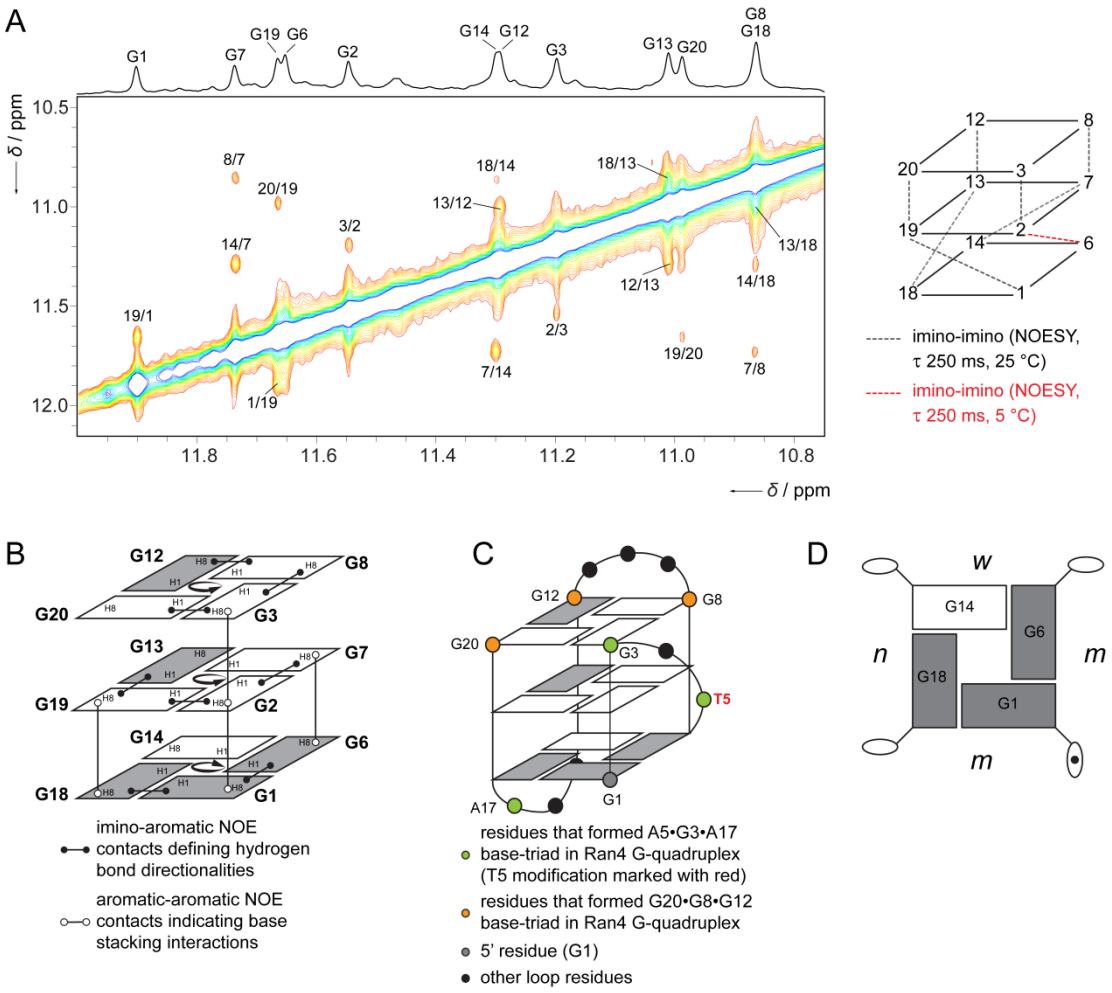

Figure S13. Determination of topology of Ran4A5T G-quadruplex. (A) Imino-imino region of NOESY NMR spectrum of Ran4A5T with mixing time of 250 ms at 25 °C in 90% H_2O, 10% ^2H_2O, 70 mM KCl, 15 mM potassium phosphate buffer with pH 7.0 (left). Oligonucleotide concentration was 1.5 mM. Schematic presentation of imino-imino NOEs observed at 25 °C (right), additional G2 H1 – G6 H1 signal was observed at 5 °C (marked with red). (B) Imino-aromatic (lines with black dots) and aromatic-aromatic (lines with white dots) NOE connectivities (from NOESY spectra with mixing time of 250 ms). (C) (3+1) hybrid topology of Ran4A5T G-quadruplex with three G-quartets. First two G-strands are connected with a propeller loop progressing anti-clockwise, while the other two loops both adopt anti-clockwise progressing lateral-type orientation (-(pll)[8a]). (D) The structure of Ran4A5T G-quadruplex is defined by one wide (between antiparallel segments G6-G8 and G12-G14), one narrow (between antiparallel segments G12-G14 and G18-G20) and two medium grooves (between parallel segments G1-G3 and G6-G7 and G1-G3 and G18 and G20, respectively).

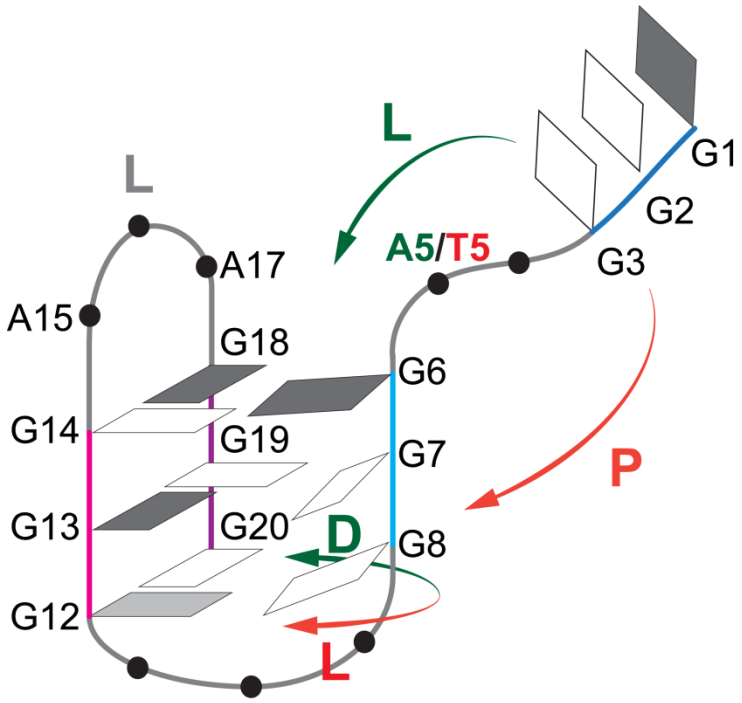

Figure S14. Possible folding pathways of Ran4 and Ran4A5T G-quadruplexes. Residues in *syn* and *anti* glycosidic conformations are represented as dark grey and white rectangles, respectively. G12 that flips its conformation from *anti* to *syn* upon A5-to-T5 modification is represented as light grey rectangle. G-strands I, II, III and IV are coloured dark blue, cyan, magenta and violet, respectively. L, P and D stand for lateral, propeller and diagonal loops, respectively. Possible folding pathway proposed for Ran4 G-quadruplex is marked with dark green arrows, while red arrows follow potential folding process of Ran4A5T G-quadruplex.

Proposed folding models start from similar hairpins formed by G-strands III and IV followed by formation of G-triplexes involving G-strands II, III and IV, whose geometries are on this stage of folding not yet final. Involvement of A5 in interactions with G3 defines the relative positions of G-strands I and II and determine the final geometries of Ran4 and Ran4A5T G-quadruplexes. In the case of A5•G3 base-pair, G-strands I and II connect in antiparallel manner. This pushes G-strand II towards IV forming diagonal middle loop (dark green arrows). On the other hand, absence of interactions between loop residues connecting G-stands I and II allows G3 to engage in the G-quartet. Consequently, two nucleotides long (*i.e.* T4-T5) propeller loop is formed between G-strands I and II. G-strand II is pushed in proximity of G-strand III resulting in lateral middle loop formation (red arrows).

References

[1] T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco.

[2] D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman, AMBER 14, University of California, San Francisco, **2014**.

[3] A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, M. Orozco, *Biophys. J.* **2007**, *92*, 3817.

[4] M. Krepl, M. Zgarbova, P. Stadlbauer, M. Otyepka, P. Banas, J. Koca, T. E. Cheatham, P. Jurecka, J. Sponer, *J. Chem. Theory Comput.* **2012**, *8*, 2506.

[5] M. Zgarbova, F. J. Luque, J. Sponer, T. E. Cheatham, M. Otyepka, P. Jurecka. *J. Chem. Theory Comput.* **2013**, *9*, 2339.

[6] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605.

[7] a) K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, A. T. Phan, *J. Am. Chem. Soc.* **2009**, *131*, 4301; b) L. Hu, K. W. Lim, S. Bouaziz, A. T. Phan, *J. Am. Chem. Soc.* **2009**, *131*, 16824.

[8] a) A. I. Karsisiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew. Chem. Int. Ed.* **2011**, *50*, 10645; b) J. Dickerhoff, K. Weisz, *Angew. Chem. Int. Ed.* **2015**, *54*, 5588.

[9] a) S. Amrane, R. W. L. Ang, Z. M. Tan, C. Li, J. K. C. Lim, J. M. W. Lim, K. W. Lim, A. T. Phan, *Nucleic Acids Res.* **2009**, *37*, 931; b) M. Trajkovski, P. Šket, J. Plavec, *Org. Biomol. Chem.* **2009**, *7*, 4677; c) M. Vorlíčková, I. Kejnovská, J. Sagi, D. Renčiuk, K. Bednářová, J. Motlová, J. Kypr, *Methods* **2012**, *57*, 64.