Supplementary Data

c-kit2 G-quadruplex stabilized via a covalent probe: Exploring G-quartet asymmetry

Kateřina Peterková[1,2,3], Ivo Durník[2,4], Radek Marek[2,4,5], Janez Plavec[1,3,6] and Peter Podbevšek[1,*]

[1] Slovenian NMR Centre, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia

[2] National Centre for Biomolecular Research, Faculty of science, Masaryk University, Kamenice 5, CZ-625 00 Brno, Czechia

[3] Faculty of Chemistry and Chemical Technology, University of Ljubljana, Večna pot 113, SI-1000 Ljubljana, Slovenia

[4] CEITEC-Central European Institute of Technology, Masaryk University, Kamenice 5, CZ-62500 Brno, Czechia

[5] Department of Chemistry, Faculty of Science, Masaryk University, Kamenice 5, CZ-62500 Brno, Czechia

[6] EN-FIST Centre of Excellence, Trg OF 13, SI-1000 Ljubljana, Slovenia

TABLE OF CONTENTS

COMPUTATIONAL DETAILS

Force Field Parameterization

Because OL15 force field has not been primarily developed for non-standard DNA nucleotides, we started by performing preliminary simulations assessing U^{py} conformational behaviour (Supporting Data, Testing Molecular Dynamics). We determined that the crucial conformational motion of U^{py} is rotation of pyrene moiety with respect to the linker and decided to test how accurately is it described by comparing results with quantum mechanical data.

First, the U^{py} nucleotide was extracted from the DNA model and the sugar-phosphate moiety was replaced by a methyl group (referred to as U^{py}Me). Relaxed potential energy surface scan was then performed at the PBE0-D3BJ/def2-TZVPP level of theory (1–4) in the program Gaussian (5). Pseudo-torsion angle τ describing conformation on the linker was defined between C4, C5, CP1, and CP11 atoms (Figure S1). Scan involved 35 steps by 10°. Corresponding scan using OL15 force field was obtained via in-house code combining gradient calculation in sander from AMBER (6) with Gaussian optimizer.

Resulting potential energy surface is shown in Figure S2 together with structures of detected energy minima (**M1** and **M2**). We found that the original OL15 force field underestimates the energy barrier for conformational change between **M1** and **M2** by 2.0 kcal mol^{-1}. Additionally, energy penalty for **M2** was underestimated by approximately 0.5 kcal mol^{-1}. To provide more accurate description in simulated annealing, unbiased and biased molecular dynamics, we fitted this torsion term to the quantum mechanical data by employing FFDevel (https://ffdevel.ncbr.muni.cz). OL15 force field optimized by adding τ pseudo-torsion term gave potential energy profile also shown in Figure S2 and was used for the production calculations.

Partial atomic charges for U^{py}Me, U^{py}1, and U^{py}21 were derived using standard RESP procedure (7, 8) implemented in the program Antechamber (9) employing electrostatic potential calculated at the HF/6-31G* level in the program Gaussian. Initial AMBER atom types were assigned using PyRED server (10–12).

NMR Structure Determination

Implicit Solvent Simulated Annealing

Linear scaling was used when the restraint violation exceeded 0.5 Å and 1.0° for distance and dihedral angle restraints, respectively. Initial velocities were derived from a random seed, which was different for each repeat. Further, tight coupling Langevin thermostat (γ = 5.0 ps^{-1}), 0.5 fs integration time step, and SHAKE algorithm (13) acting on bonds with hydrogen atoms were used.

Explicit Solvent Refinement

Final refinement of the NMR ensemble was performed in an explicit solvent under the periodic boundary conditions employing the truncated octahedral box. Long-range interactions were treated with the particle-mesh Ewald method (14), with a direct summation cut-off set to 8.0 Å. The same cut-off was used for Lennard-Jones interactions. Equations of motions were integrated with a time step of 2 fs and lengths of bonds containing hydrogen atoms were constrained by SHAKE algorithm (13).

Each system was equilibrated by geometry optimization followed by heating (100 ps) to 300 K at a constant volume employing the Langevin thermostat with a collision frequency (γ) of 1.0 ps^{-1}. Finally, the proper density was adjusted by short simulation (500 ps) at the constant temperature (the same thermostat as in the previous step) and pressure maintained by the barostat set to 100 kPa with a feedback time constant (t_p) of 1.2 ps. Integration time step of 1 fs was employed during equilibration.

After equilibration, MD simulations were performed at a constant temperature of 300 K (Berendsen thermostat, t_T = 5 ps) and a pressure of 100 kPa (weak coupling barostat, tp = 6 ps). Integration time step of 2 fs was used. Trajectory of 1 μs was collected for each model and in the last step, the systems were cooled to 0 K during 10 ns at constant volume and geometry optimized (5 000 steps). After stripping solvent, the structures of DNA including two intercalated potassium cations were used as the final NMR ensemble.

Testing Molecular Dynamics

To get some estimate on the conformational behaviour of c-kit2_py21, two starting structures were prepared by substituting previously solved NMR structure of c-kit2 (PDB ID 2KYP, model 1) using Nemesis (https://nemesis.ncbr.muni.cz/). U^{py} was stacked onto G4-G6-G16-G20 quartet in two coplanar conformations (Figure S4A). Two potassium cations were placed in between the adjacent G-quartets. Resulting models will be denoted as c-kit2_py21A, and c-kit2_py21B. Parent structure was prepared alongside and was altered only by adding K^+ cations. Each DNA was then immersed into a truncated octahedral box filled by the TIP3P water and 19 sodium cations to maintain electroneutrality. Molecular dynamics simulations were performed in the Amber 16 package using GPU and pmemd-cuda 14.0. Temperature and pressure were kept at 300 K and 100 kPa, respectively. SHAKE algorithm acting on bonds with hydrogen atoms, 2 fs integration timestep were used, and 10 μs long trajectory for each system was obtained.

Overall structure of c-kit2_py21A and c-kit2_py21B was found to be stable during the entire simulation. However, several conformational changes were recorded for the $U^{py}21$ nucleotide. After 1 μs of the MD, uracil nucleobase of $U^{py}21$ rotated by 180° around N-glycosidic bond and around linker in both models and remained in this orientation for the rest of the simulation (Figures S4B and S4C). Because of that, we excluded first μs of the MD from the analysis as it appeared that the systems were still under equilibration. One additional conformational change was observed in the c-kit2_py21B model (Figure S4D). Here, free 3'-end rotated so that the -OH group started to interact with the preceding phosphate.

Complex conformational behaviour of $U^{py}21$ is caused by the covalent linkage to the G-quadruplex. While there are only two possible conformations for the U^{py} itself, which can be described by pseudo-torsion angle τ (Supporting Data, Force Field Parameterization), the number of conformations increases when U^{py} is coupled to the G-quadruplex, due to reduced symmetry. Based on this, pseudo-torsion angle τ between uracil nucleobase and pyrene moiety could not been used to explore all the possible conformations. For example, c-kit2_py21B would have same value of τ (Figure S4B) as sugar-base flipped ckit2_py21A (Figure S4D), even though they are in inequivalent states.

Therefore, we employed different set of geometrical parameters to correctly separate these motions. They were based on torsion angles between three independent vectors τ_{py} and τ_{bs}, which allowed us to individually reference rotations of uracil nucleobase and pyrene moiety with respect to the linker (Figure S7). Time evolution of these parameters during the testing MD is shown in Figure S5. Because all observed motions could be separated using τ_{py} and τ_{bs}, we decided to employ these parameters in subsequent free energy calculations to get more information on their coupling and corresponding free energies.

We also monitored time evolution of torsion angle χ on N-glycosidic bond defining mutual orientation of the uracil nucleobase and sugar (Figure S6) and conformational switching comparable to NMR data was observed. For c-kit2_py21A, χ of $U^{py}21$ was sampling two substates, where one belonged to *anti* and second to *syn*. On the contrary, $U^{py}21$ in c-kit2_py21B and T21 in c-kit2 both demonstrated *anti* conformation (Figure S6). Fast exchange between *syn-anti* substates suggests that these motions play a secondary role in the separation of main stacking modes of $U^{py}21$. We believe that sugar can find the most favourable orientation on the timescale of MD simulations and that this conformation is rather enforced by position of the uracil nucleobase and pyrene moiety. Therefore, we decided that it is not necessary to bias or restrain χ during the free energy calculations.

Molecular Dynamics

Unbiased Molecular Dynamics Simulations

Unbiased molecular dynamics closely followed the setting used during explicit solvent refinement of the NMR structure including both equilibration and production phase. Starting structures were extracted from 10 μs of testing molecular dynamics, when systems were found to equilibrate (Supporting Data, Testing Molecular Dynamics). Resulting trajectories of ckit2 and c-kit2_py21 were 1 μs long and in case of c-kit2_py21 also served to generate starting configurations for subsequent biased MD.

Biased Molecular Dynamics Simulations

To avoid coherency between walkers, each walker started from different configuration taken from the unbiased MD simulations as a restart file with a minimum time separation of at least 10 ns from the other restart files. Moreover, a stochastic Langevin thermostat with a low collision frequency of $\gamma = 0.1\ ps^{-1}$ was employed during ABF simulations to avoid spontaneous synchronizations of walkers. Besides, each walker was run in a 5 ns batches with generating a new seed for the pseudo-random generator during the restart of MD simulation. Change of seed numbers was required for the proper functionality of Langevin thermostat in frequently restarted MD runs (15), and it further improved stochastic behaviour and avoided possible synchronizations of walkers.

Nucleus-Independent Chemical Shift (NICS)

Spatial NMR shielding/deshielding effect around pyrene moiety in c-kit2_py1 was demonstrated by the nucleus-independent chemical shift approach (16). First, the geometry of the pyrene moiety was optimized *in vacuo* via Gaussian 16.B1 at the PBE0-D3BJ/def2-TZVPP level of theory with applied symmetry. The chemical shifts were evaluated on regular grid around the pyrene by the PBE0-D3BJ/6-311G** approach. There were 110 points in *x*, 120 in *y* and 80 in *z* direction and they were spaced by 0.25 Å. Fortunately, only 1/8 of the space around pyrene had to be explicitly evaluated due to pyrene symmetry (*xy*, *xz*, *yz* mirror planes). Calculated chemical shifts were visualized using VMD (17) and are shown in Figure S30.

Pyrene moiety appeared to experience sliding on the G-quartet as it was not held by any specific interactions, but only by stacking and covalent linkage to the sugar-phosphate backbone. As these slight motions can be important for values of observed chemical shifts, we decided to use all structures from the NMR ensemble of c-kit2_py1 to calculate change in the chemical shift induced by the attachment of the pyrene moiety. While this would not be possible due to high computational cost of required quantum mechanical calculations, we overcame the problem by employing NICS data. In the approximation, each structure from the ensemble was re-oriented such that the pyrene from ckit2_py1 was best fitted to the pyrene used in the NICS calculation described above. Then, the NICS value was retrieved for each imino proton of the adjacent quartet from the volumetric map (Figure S30) and averaged across all of the models.

This approximation assumes that the pyrene is a rigid moiety, which is satisfied. Another approximation is that the remaining structure (linker to uracil, uracil nucleobase, and c-kit2) does not influence the NICS values of the pyrene. Similarly, additional errors arise from exclusion of solvent effects, which were not directly accounted for in the NICS calculation. Based on this, calculated values can be interpreted only in qualitative sense.

Summary of NICS results for c-kit2_py1 refined with and without experimentally derived restraints in comparison to experimental differences in chemical shielding of imino protons between c-kit2 and c-kit2_py1 is provided in Table S4. NICS values for both ensembles showed G6 as one of the most shielded imino protons. However, only the unrestrained ensemble showed G14 shielded more than G2 as observed in NMR spectra. G18 demonstrated almost no shielding or deshielding in agreement with the spectra. The major difference is seen for G2 that is predicted as more shielded contrary to the experiment by the restrained ensemble. Because pyrene is sliding from G2 over G6 to G14 (Figures 4B and 4C) and the population of different positions gives observed shift, positioning pyrene towards G14 would increase its shielding and decrease shielding of G2. Better performance of unrestrained ensemble suggests that pyrene is present above G14 during a transient state, which might not be captured well by the time-averaged restraints. Additionally, the discrepancy in absolute values of NICS towards NMR shifts can be attributed to employed approximations discussed above.

REFERENCES

1. Adamo,C. and Barone,V. (1999) Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.*, **110**, 6158–6170.

2. Weigend,F. and Ahlrichs,R. (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.*, **7**, 3297–3305.

3. Grimme,S., Antony,J., Ehrlich,S. and Krieg,H. (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*, **132**, 154104.

4. Grimme,S., Ehrlich,S. and Goerigk,L. (2011) Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.*, **32**, 1456–1465.

5. Frisch,M.J., Trucks,G.W., Schlegel,H.B., Scuseria,G.E., Robb,M.A., Cheeseman,J.R., Scalmani,G., Barone,V., Petersson,G.A., Nakatsuji,H., *et al.* (2016) Gaussian 16, Revision A.03 Gaussian Inc., Wallingford CT.

6. Case,D.A., Babin,V., Berryman,J.T., Betz,R.M., Cai,Q., Cerutti,D.S., Cheatham III,T.E., Darden,T.A., Duke,R.E., Golhke,H., *et al.* (2016) Amber 16 University of California, San Francisco.

7. Bayly,C.I., Cieplak,P., Cornell,W. and Kollman,P.A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.*, **97**, 10269–10280.

8. Cieplak,P., Cornell,W.D., Bayly,C. and Kollman,P.A. (1995) Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem.*, **16**, 1357–1377.

9. Wang,J., Wang,W., Kollman,P.A. and Case,D.A. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.*, **25**, 247–260.

10. Vanquelef,E., Simon,S., Marquant,G., Garcia,E., Klimerak,G., Delepine,J.C., Cieplak,P. and Dupradeau,F.-Y. (2011) R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res.*, **39**, W511–W517.

11. Dupradeau,F.-Y., Pigache,A., Zaffran,T., Savineau,C., Lelong,R., Grivel,N., Lelong,D., Rosanski,W. and Cieplak,P. (2010) The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Phys. Chem. Chem. Phys.*, **12**, 7821–7839.

12. Wang,F., Becker,J.-P., Cieplak,P. and Dupradeau,F.-Y. (2013) R.E.D. Python: Object oriented programming for Amber force fields Jules Verne, Sanford Burnham Prebys Medical Discovery Institute, Université de Picardie.

13. Ryckaert,J.-P., Ciccotti,G. and Berendsen,H.J.C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, **23**, 327–341.

14. Darden,T., York,D. and Pedersen,L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.

15. Sindhikara,D.J., Kim,S., Voter,A.F. and Roitberg,A.E. (2009) Bad Seeds Sprout Perilous Dynamics: Stochastic Thermostat Induced Trajectory Synchronization in Biomolecules. *J. Chem. Theory Comput.*, **5**, 1624–1631.

16. Schleyer,P. von R., Maerker,C., Dransfeld,A., Jiao,H. and van Eikema Hommes,N.J.R. (1996) Nucleus-Independent Chemical Shifts: A Simple and Efficient Aromaticity Probe. *J. Am. Chem. Soc.*, **118**, 6317–6318.

17. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.

SUPPLEMENTARY FIGURES

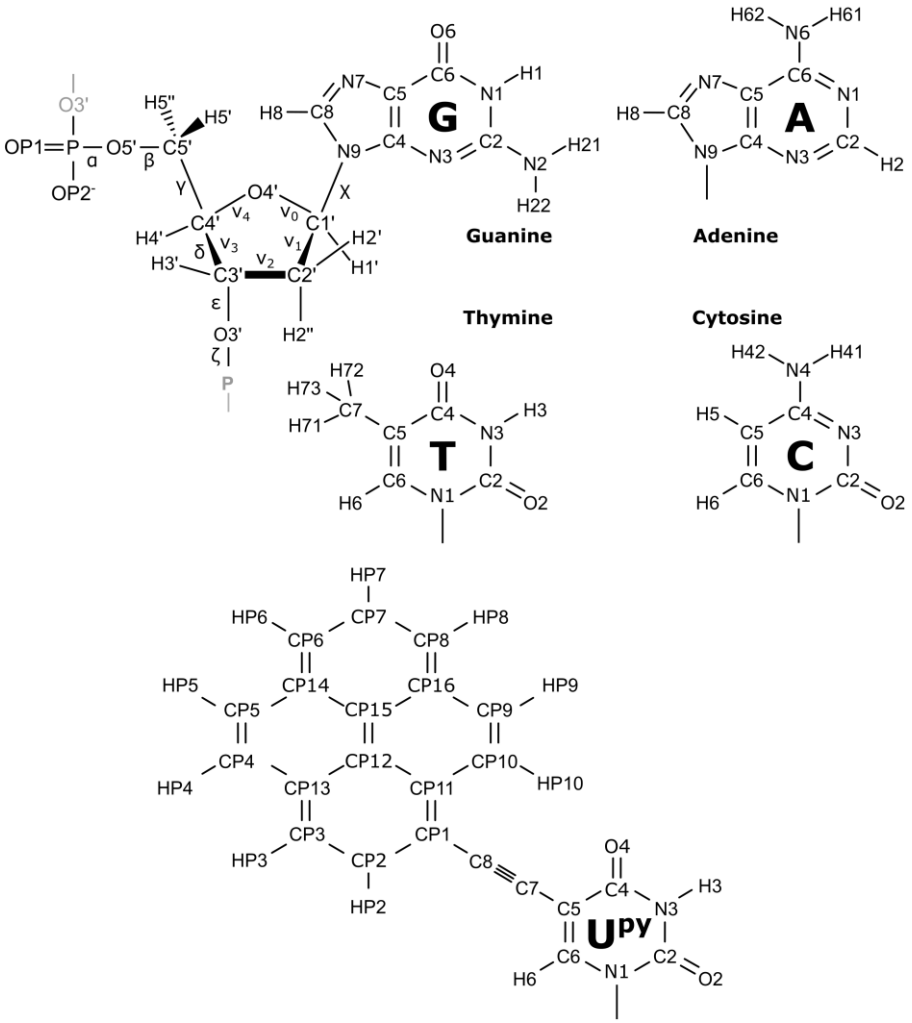

Figure S1. Naming of atoms, backbone torsion angles (α, β, γ, δ, ε, and ζ) and five torsion angles around the bonds in the ribose ring used to define sugar puckering (ν_0, ν_1, ν_2, ν_3, and ν_4).

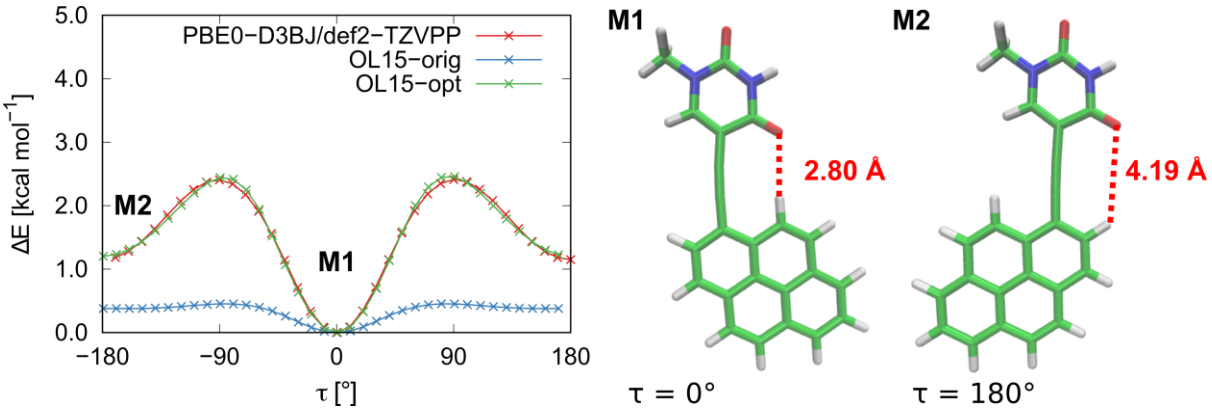

Figure S2. Potential energy surface for rotation of pyrene moiety with respect to the linker in U^{py}Me calculated by quantum mechanics, the original OL15 force field and optimized OL15 with added pseudo-torsion term. Torsion τ was defined between C4, C5, CP1, and CP11 atoms (Figure S1). Barrier to the rotation is underestimated by the original force field, which rather features a plateau for conformation **M2** with only slight rise in the energy. Parametrized and added torsion term results in accurate representation of the quantum mechanical profile. Interaction between O6 and pyrene proton, marked by striped red line in the structures, gives a 1.2 kcal mol^{-1} preference for conformation **M1** over **M2**. Sugar-phosphate moiety was replaced by a CH_3 cap.

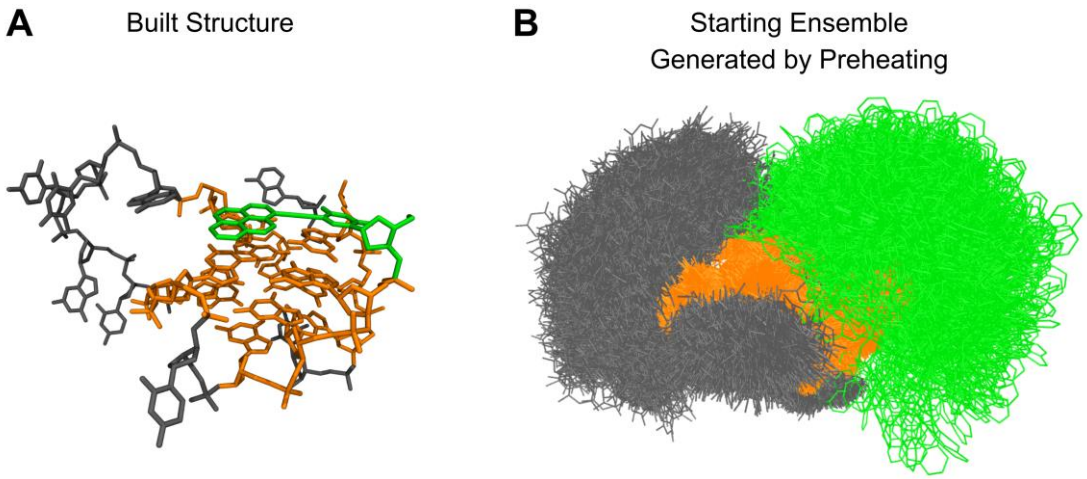

Figure S3. (A) Starting structure built by substituting U^{py} into position 1 of c-kit2 (PDB ID 2KYP) using Nemesis program, (B) ensemble of 1000 starting structures generated by preheating at 1200 K for 200 ps with all restraints except NOE derived distance restraints involving U^{py}1. G-quartets are in orange, U^{py}1 in green and loop nucleotides in grey. Hydrogen atoms and intercalated potassium cations are not shown for clarity.

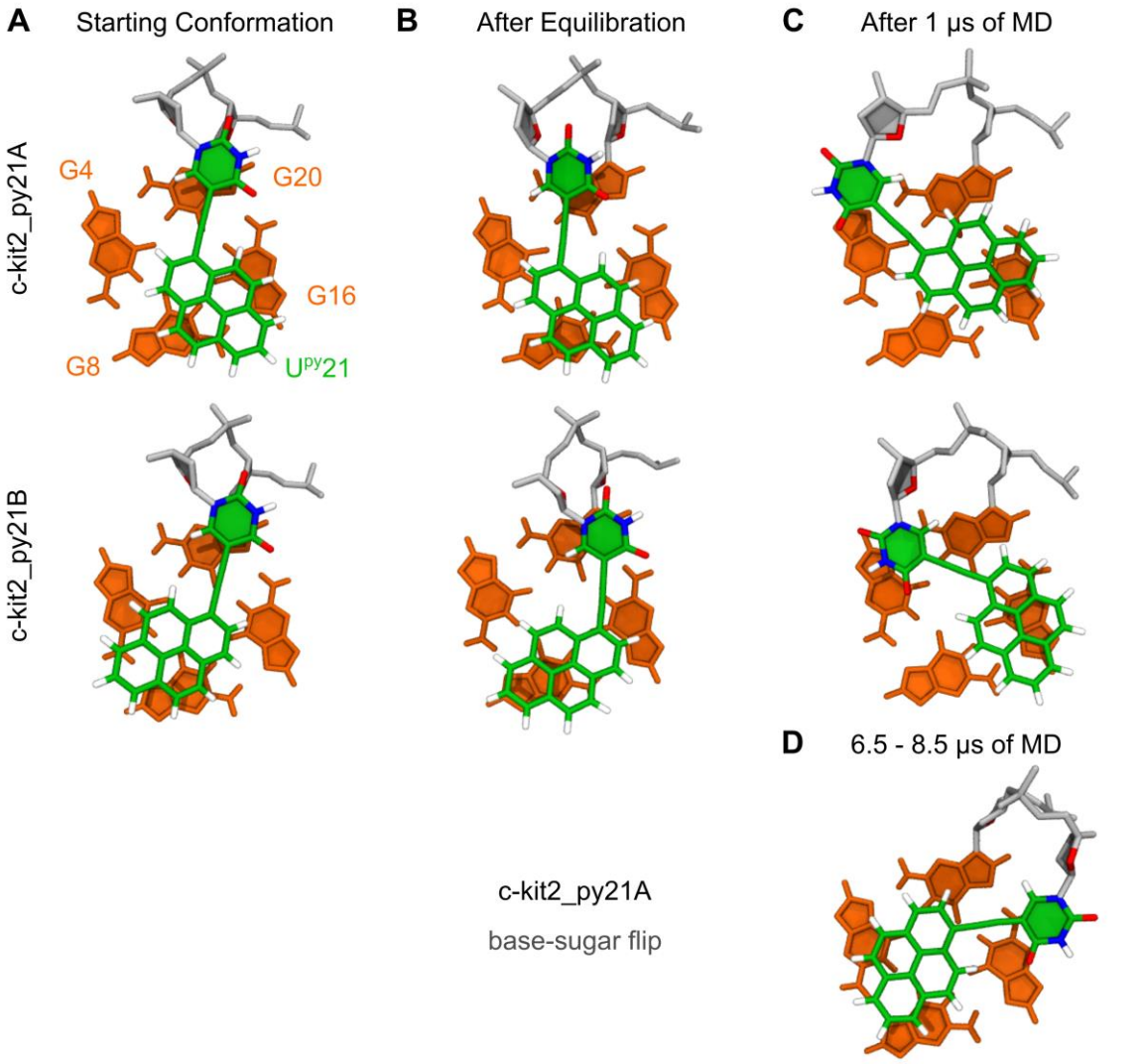

Figure S4. Structures observed during testing MD for c-kit2_py21A, and c-kit2_py21B: (A) starting structure, (B) after equilibration, (C) after 1 μs of molecular dynamics (D) example of c-kit2_py21A in second base-sugar orientation, which occurred between 6.5 and 8.5 μs (Figures S5 and S6).

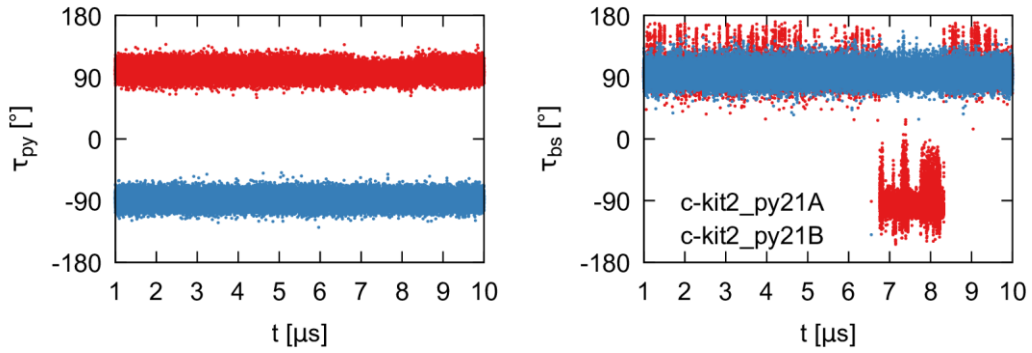

Figure S5. Analysis of pyrene and base-sugar orientation for c-kit2_py21A and c-kit2_py21B during 10 μs long unbiased testing molecular dynamics. Employed collective variables are shown in Figure S7. While both systems preserved their starting configuration of pyrene described by τ_{py}, one flip occurred in the base-sugar domain for c-kit2_py21A given by τ_{bs}. Free energy calculations were then designed to investigate how these motions are coupled (Figure S32). First μs was discarded from the analysis because systems were still under equilibration (Figure S4).

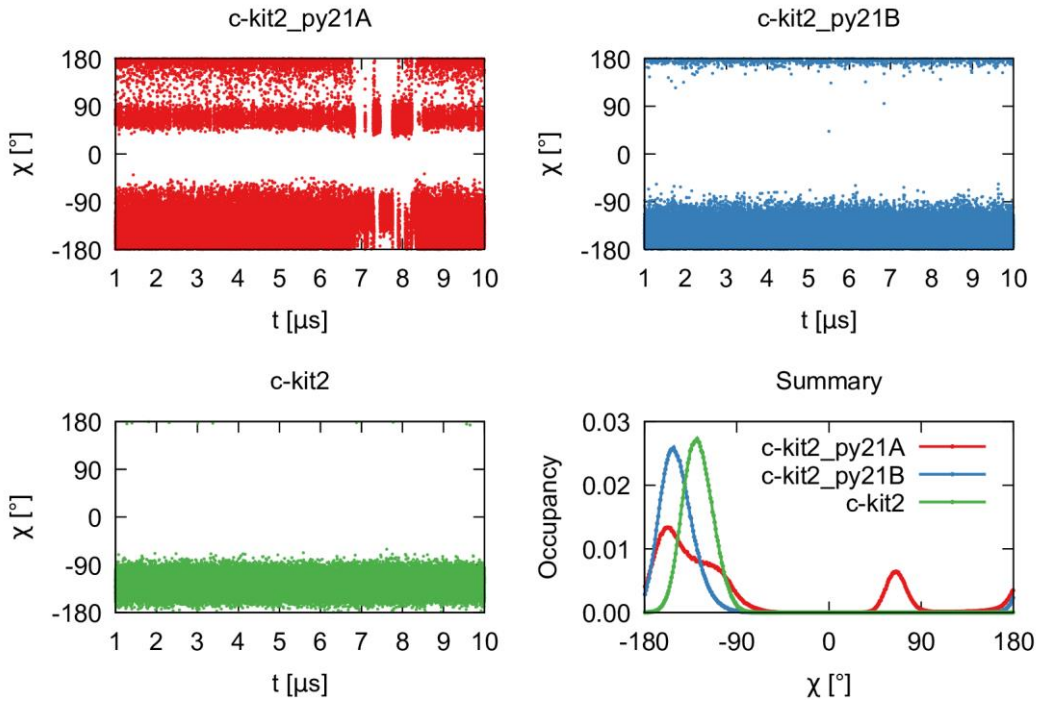

Figure S6. Analysis of χ dihedral of $U^{py}21$ (or T21) in c-kit2_py21A, c-kit2_py21, and c-kit2 models during 10 μs long unbiased testing molecular dynamics. In case of c-kit2_py21A, *syn-anti* switching took place, whereas c-kit2_py21B and c-kit2 both preferred *anti* conformation of the 3' terminal nucleotide. Histogram analysis is presented in right bottom corner. First μs was discarded from the analysis because systems were still under equilibration (Figure S4).

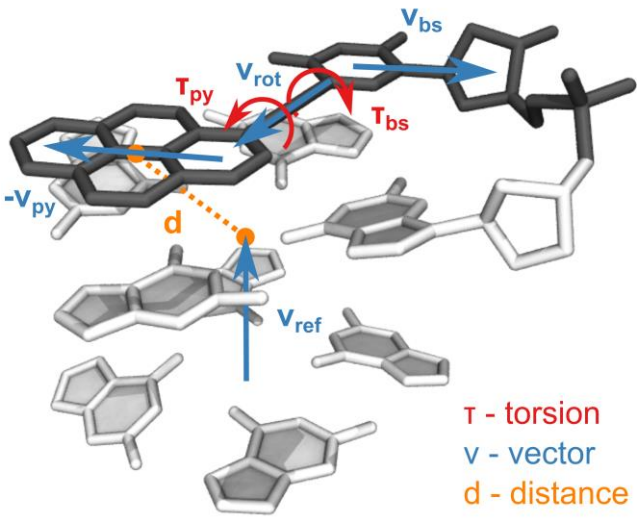

Figure S7. Geometrical parameters (collective variables) employed during biased molecular dynamics of ckit2_py21. Rotation of pyrene moiety with respect to the linker was described by torsion angle τ_{py} between three independent vectors $-v_{py}$, v_{rot}, and v_{ref}. A similar description was used for rotation of base-sugar moiety τ_{bs} (vectors v_{bs}, v_{rot}, and v_{ref}). First vector defined position of the rotating moiety, vector v_{rot} the axis of rotation, and v_{ref} referenced the rotation to the position of the G-core. Distance d marks wall restraint preventing irreversible destacking of U^{py} during performed rotations.

A

	5	10	15	20
c-kit2_py9	C **GGG** C **GGG** **U^{py}**GCTA **GGG** A **GGG** T			
c-kit2_py10	C **GGG** C **GGG** C**U^{py}**CTA **GGG** A **GGG** T			
c-kit2_py11	C **GGG** C **GGG** CG**U^{py}**TA **GGG** A **GGG** T			
c-kit2_py12	C **GGG** C **GGG** CGC**U^{py}**A **GGG** A **GGG** T			
c-kit2_py13	C **GGG** C **GGG** CGCT**U^{py}** **GGG** A **GGG** T			

B

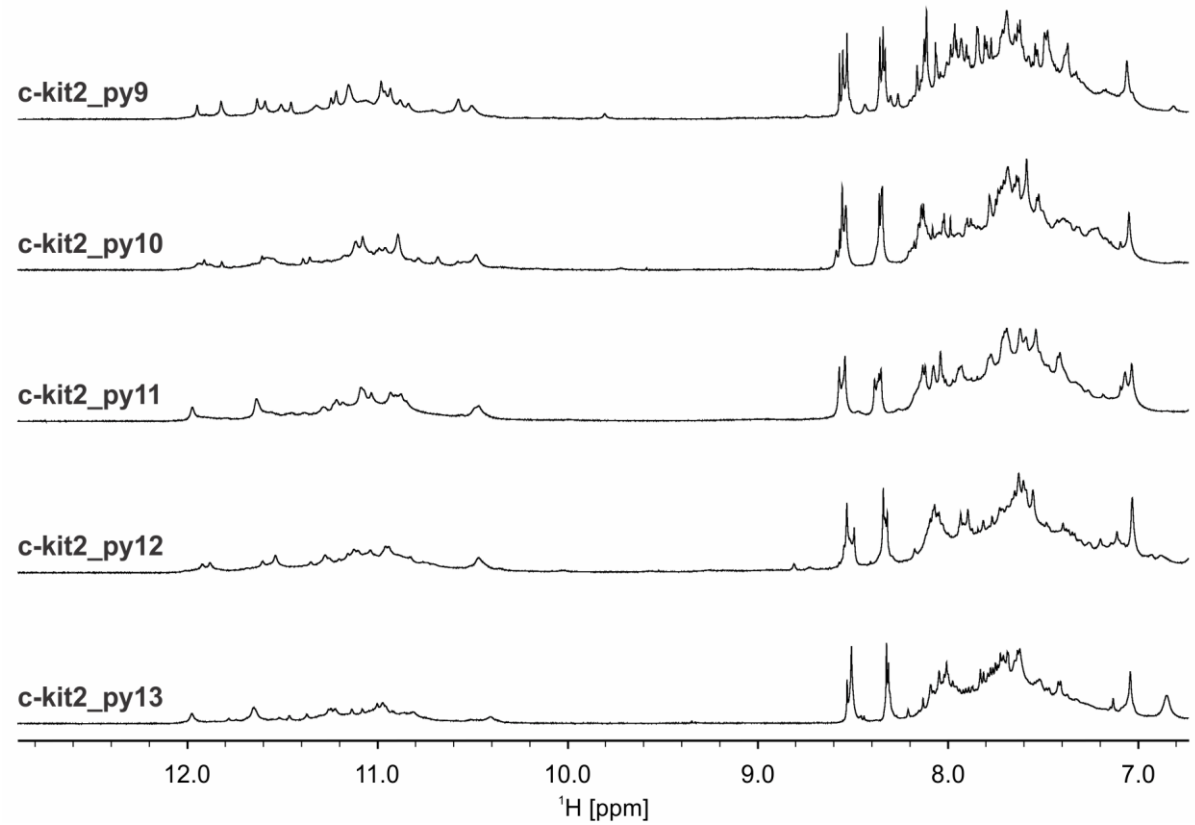

Figure S8. (A) Oligonucleotide sequences of loop U^{py} analogues (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of loop U^{py} analogues. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations varied between 0.7 and 0.9 mM per strand.

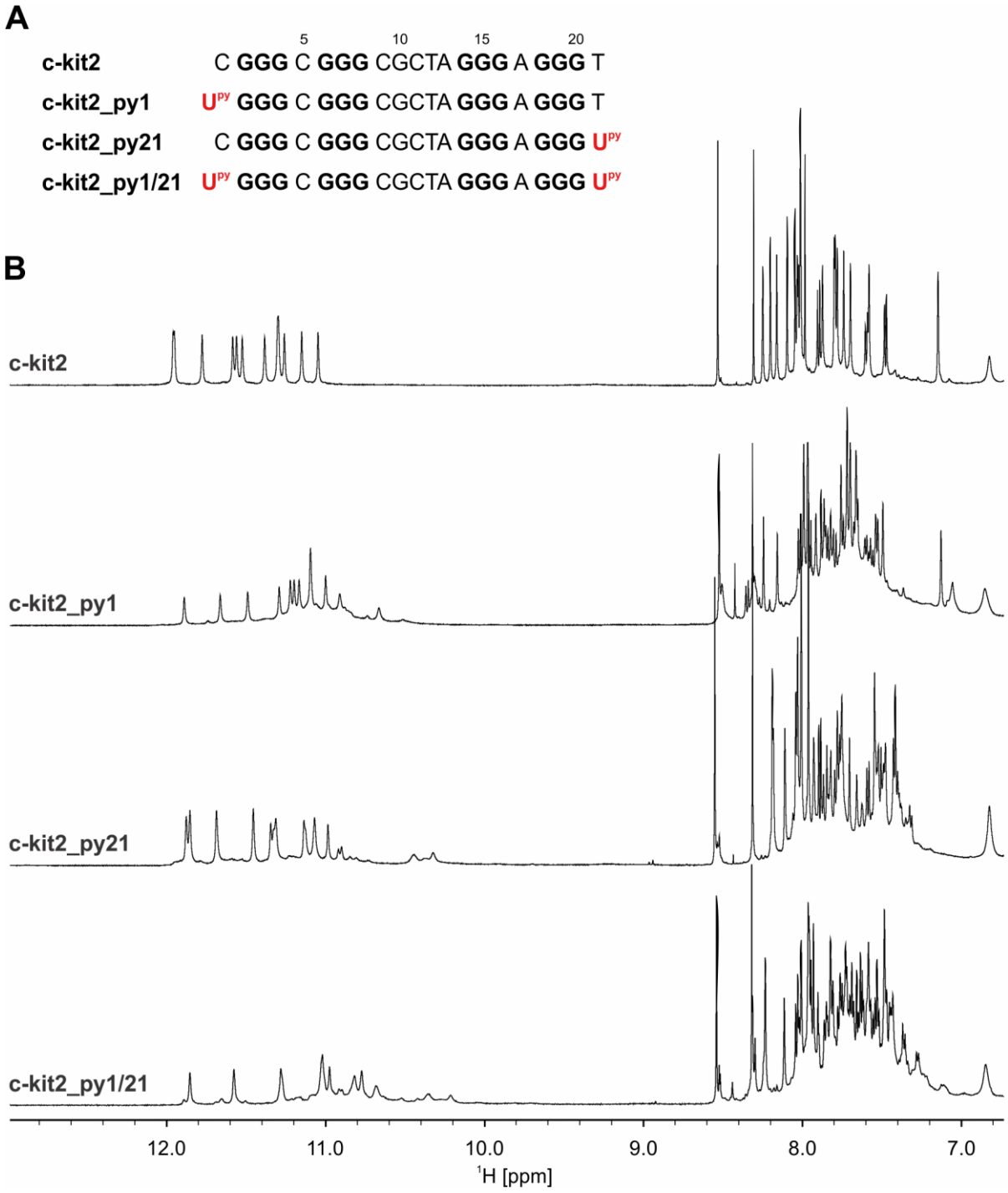

Figure S9. (A) Oligonucleotide sequences of c-kit2 and terminal U^{py} analogues (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of c-kit2 and terminal U^{py} analogues. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations varied from 0.7 to 1.0 mM.

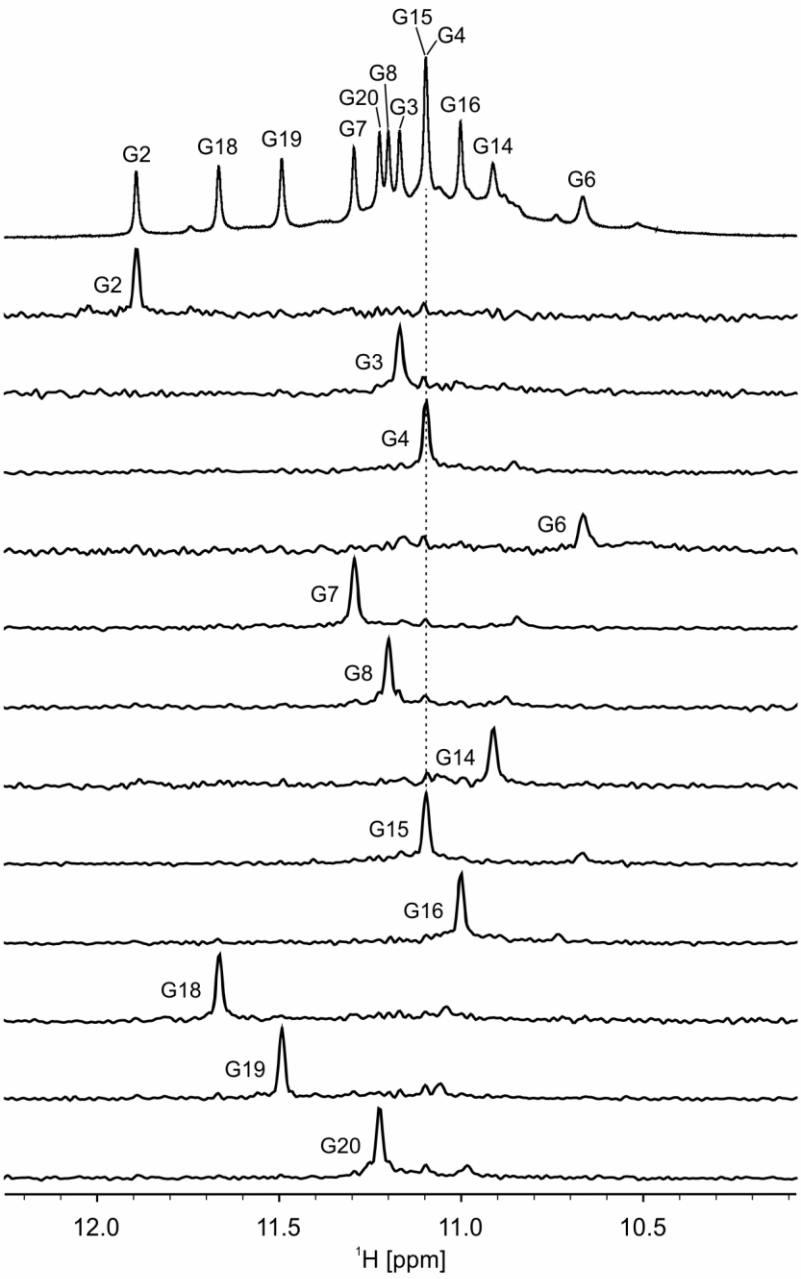

Figure S10. Imino region of 1D ^{1}H NMR spectrum above ^{15}N-edited HSQC spectra of c-kit2_py1 G-quadruplex. HSQC spectra were acquired on partially (10%) site-specific ^{15}N- and ^{13}C- labelled oligonucleotides. Assignment of H1 proton resonances is indicated next to the 1D signals. The dashed line signifies the assignment of G4 and G15 whose H1 protons are nearly isochronous. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations of labelled samples varied between 0.4 and 0.6 mM, oligonucleotide concentration of non-labelled sample was 1.0 mM.

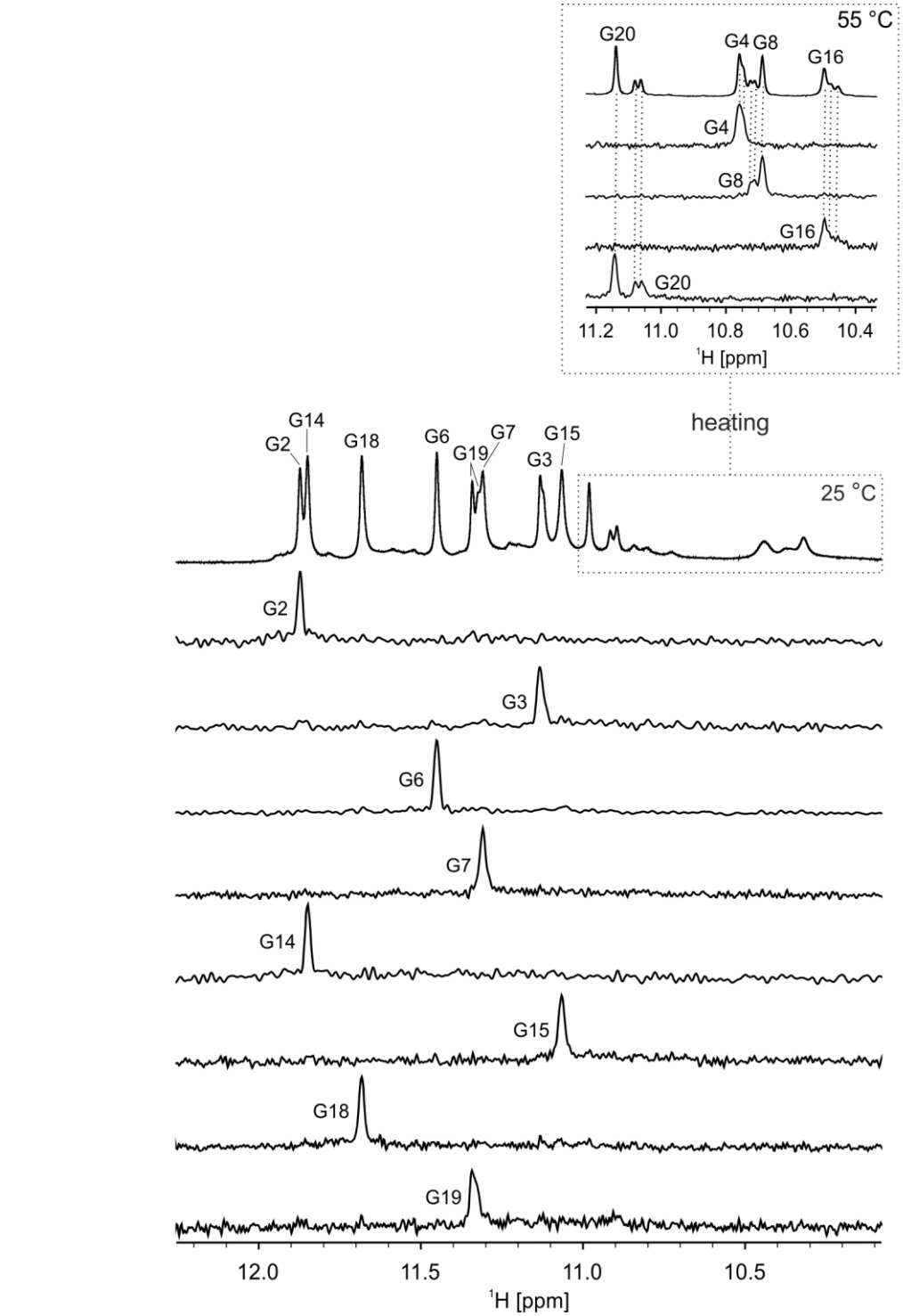

Figure S11. Imino region of 1D ^{1}H NMR spectrum above ^{15}N-edited HSQC spectra of c-kit2_py21 G-quadruplex. HSQC spectra were acquired on partially (10%) site-specific ^{15}N- and ^{13}C- labelled oligonucleotides. Assignment of H1 proton resonances is indicated next to the 1D signals. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 or 55 °C. Oligonucleotide concentrations of labelled samples varied between 0.5 and 0.8 mM, oligonucleotide concentration of non-labelled sample was 0.9 mM.

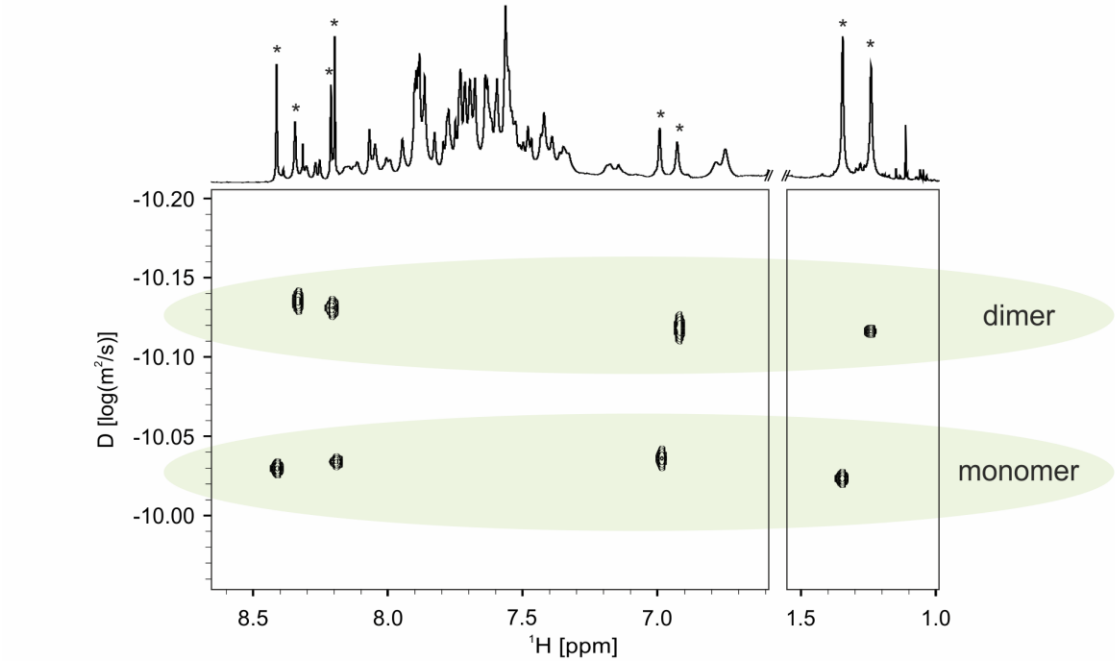

Figure S12. 1D ^{1}H NMR spectrum above the DOSY NMR spectrum of c-kit2_py1. Well-resolved proton signals used for calculation of diffusion coefficients are marked with asterisks in 1D ^{1}H NMR spectrum. Cross-peaks from monomeric and dimeric species are marked with green boxes. Spectra were acquired in 100% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 15 °C. Oligonucleotide concentration was 1.0 mM.

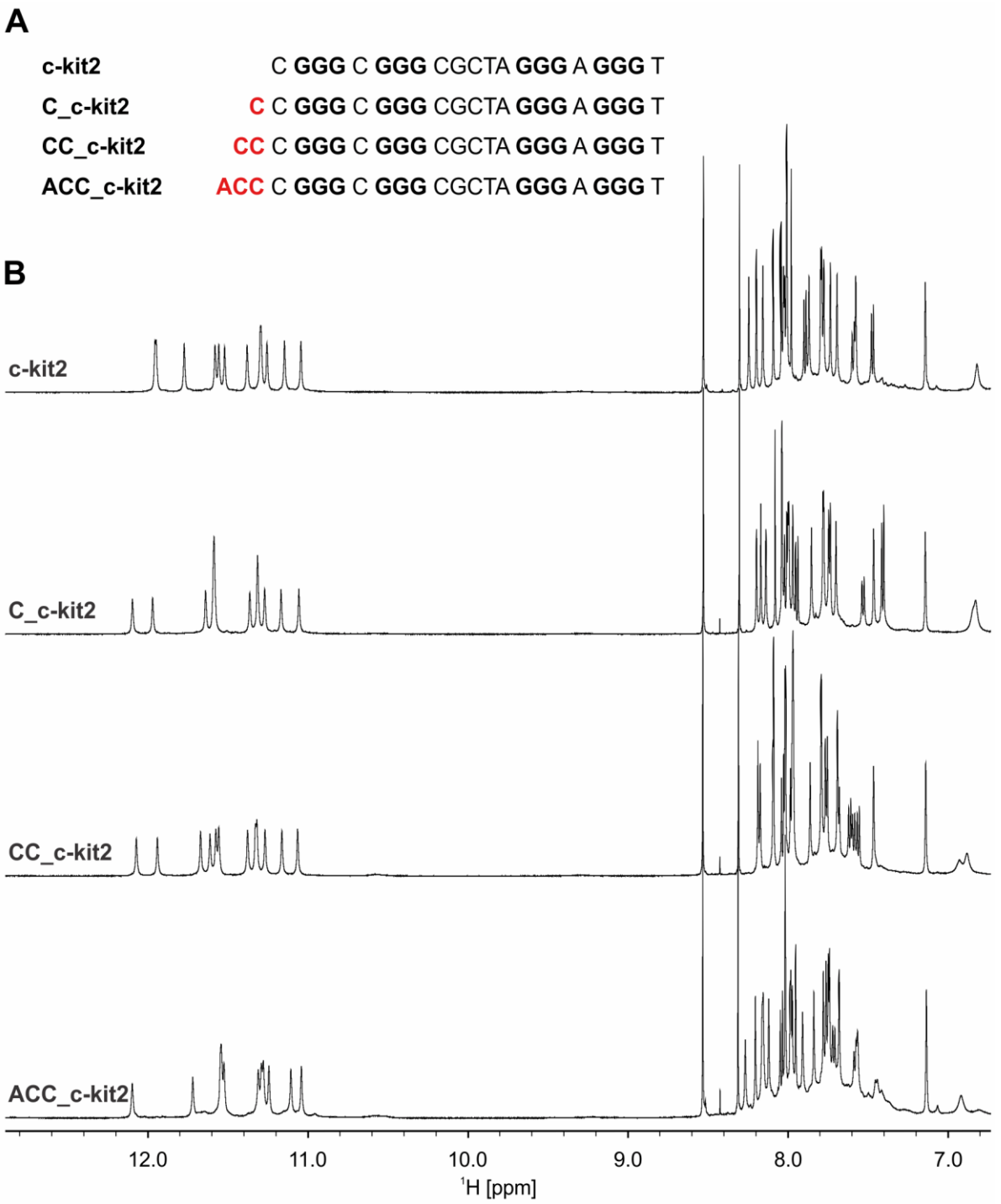

Figure S13. (A) c-kit2 oligonucleotide sequences with 1-3 nucleotide long 5'-overhang (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of c-kit2 and c-kit2 with 1-3 nucleotide long 5'-overhangs. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations were ~0.4 mM.

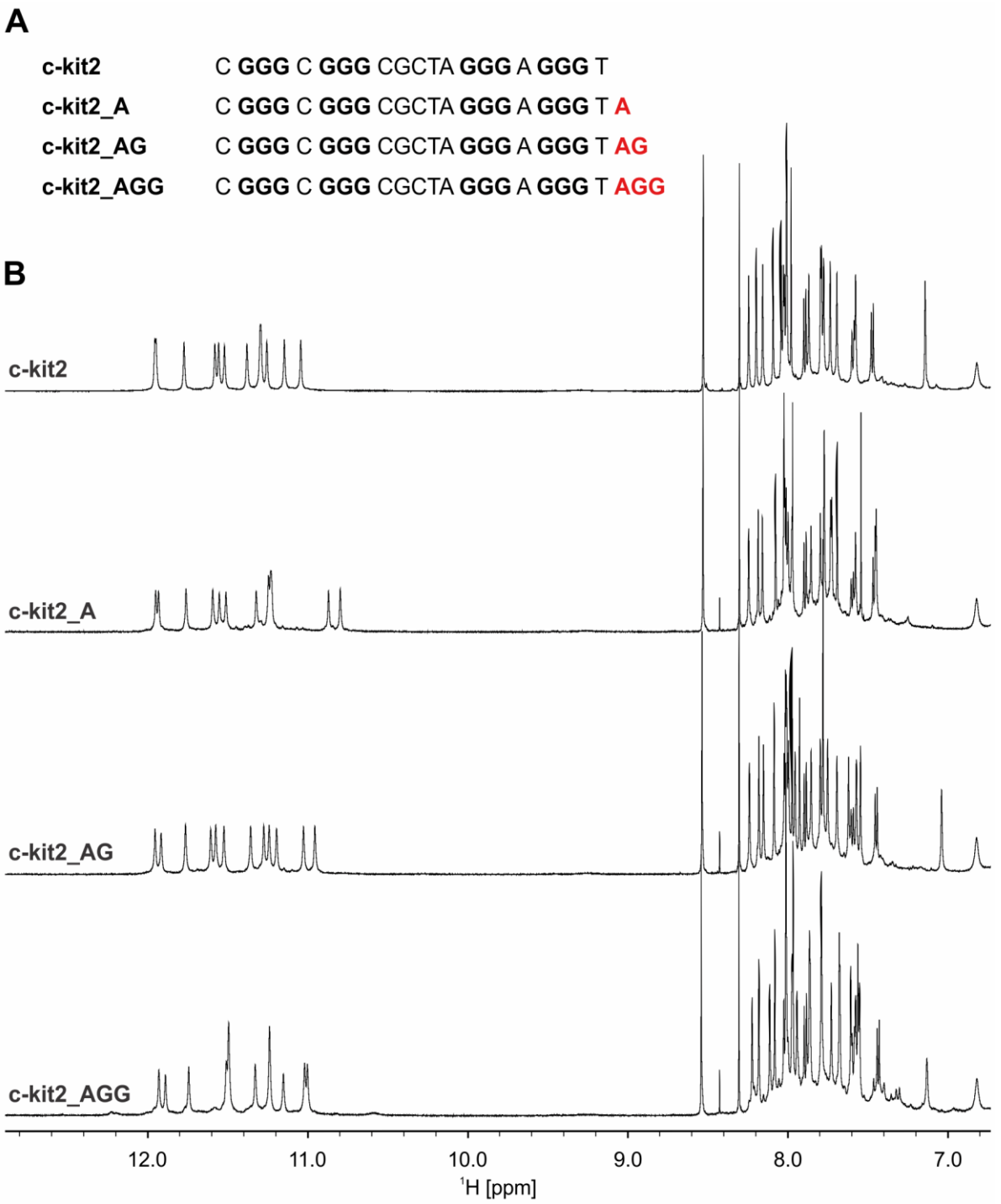

Figure S14. (A) c-kit2 oligonucleotide sequences with 1-3 nucleotide long 3'-overhang (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of c-kit2 and c-kit2 with 1-3 nucleotide long 3'-overhangs. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations were ~0.4 mM.

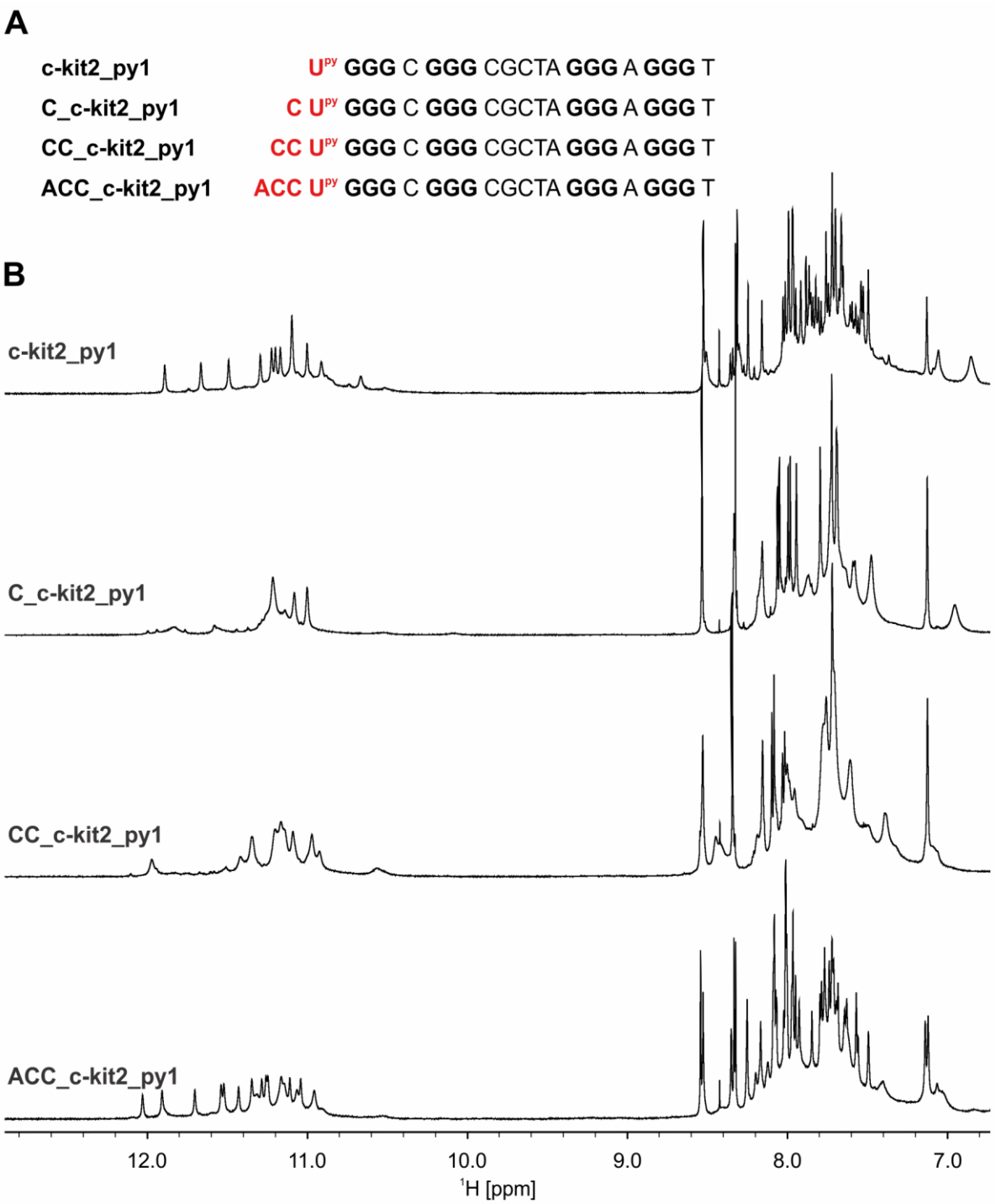

Figure S15. (A) c-kit2_py1 oligonucleotide sequences with 1-3 nucleotide long 5'-overhang (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of c-kit2_py1 and c-kit2_py1 with 1-3 nucleotide long 5'-overhangs. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations were ~0.4 mM.

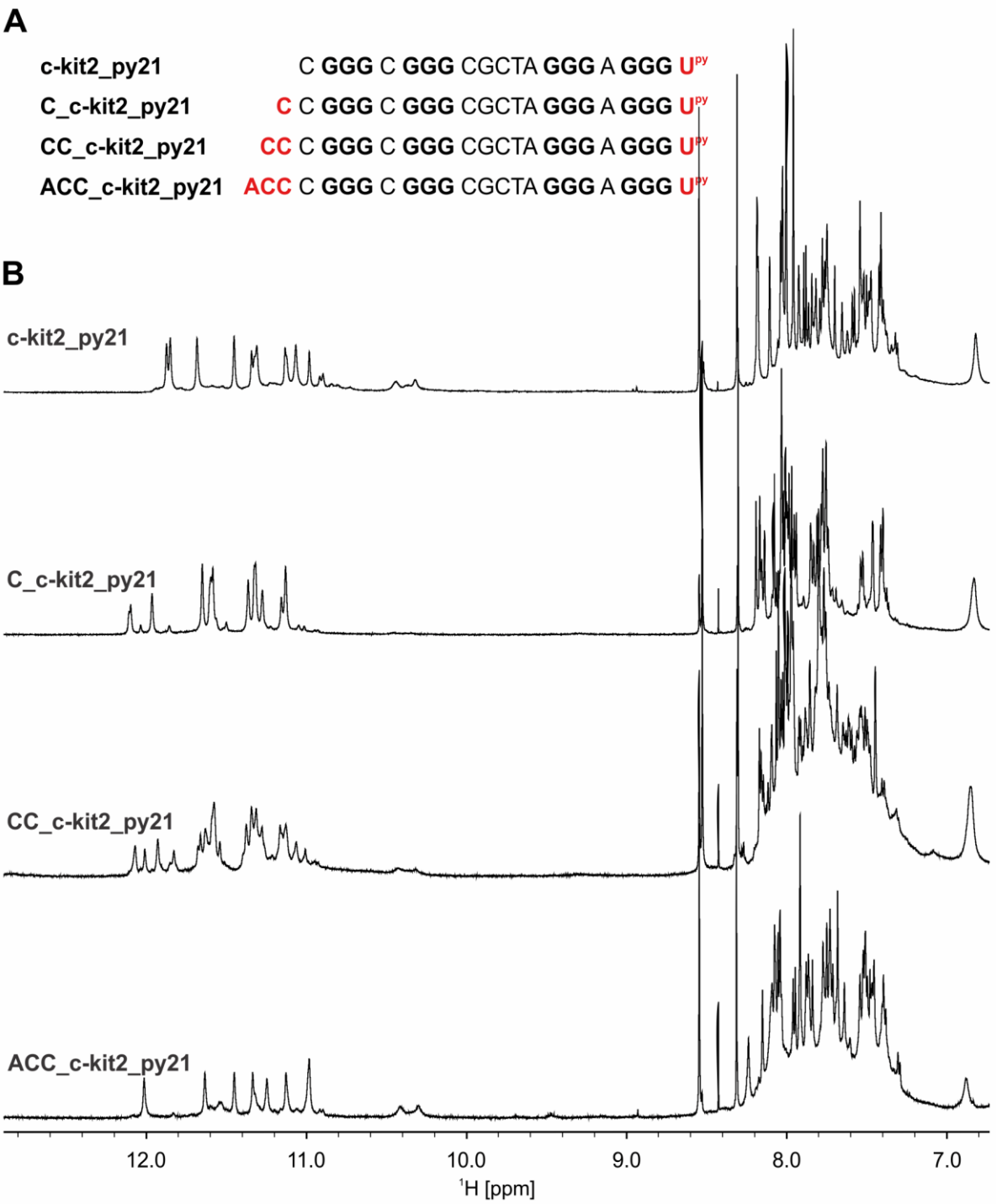

Figure S16. (A) c-kit2_py21 oligonucleotide sequences with 1-3 nucleotide long 5'-overhang (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of c-kit2_py21 and c-kit2_py21 with 1-3 nucleotide long 5'-overhangs. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations were ~0.4 mM.

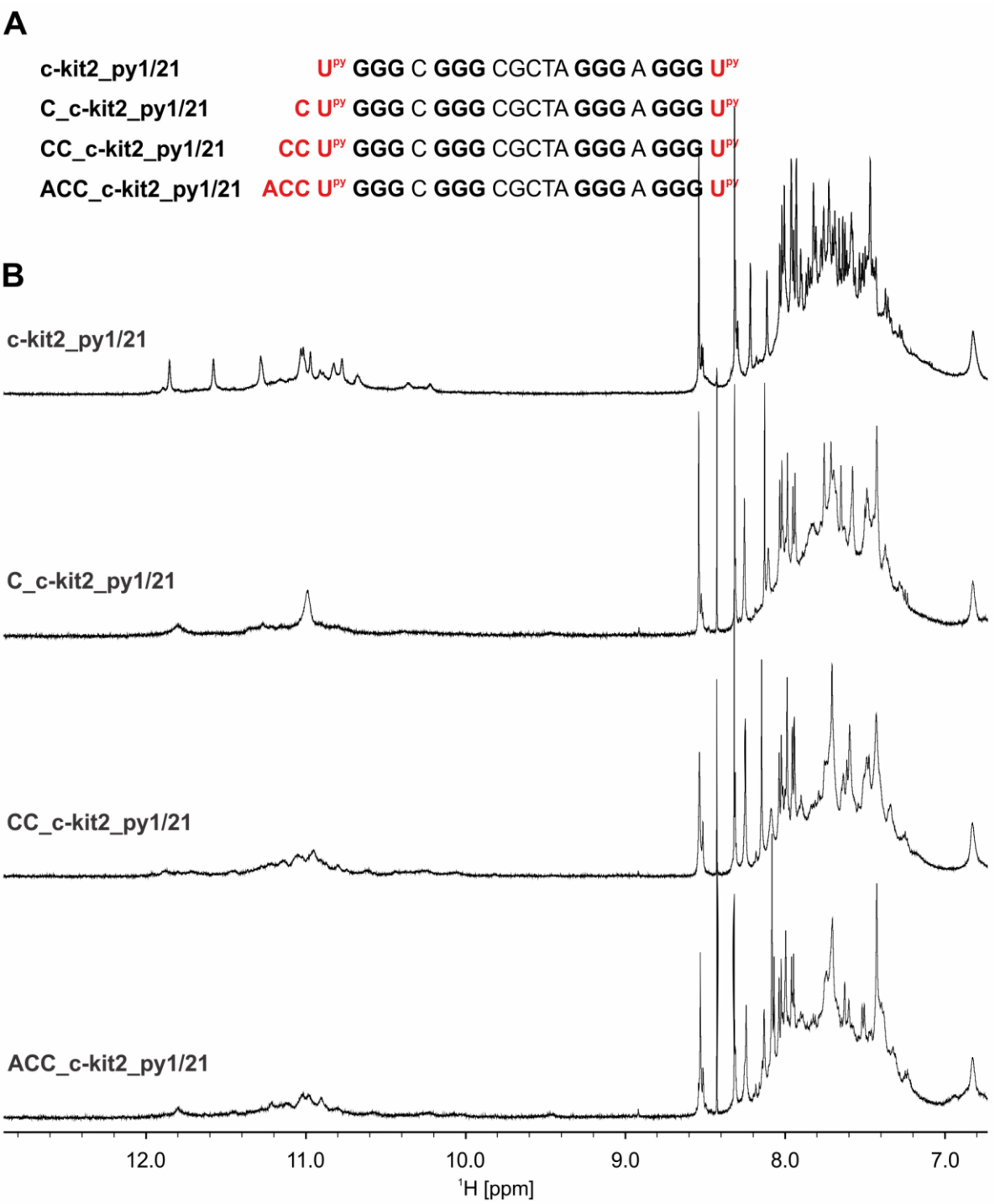

Figure S17. (A) c-kit2_py1/21 oligonucleotide sequences with 1-3 nucleotide long 5'-overhang (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of c-kit2_py1/21 and c-kit2_py1/21 with 1-3 nucleotide long 5'-overhangs. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations were ~0.4 mM.

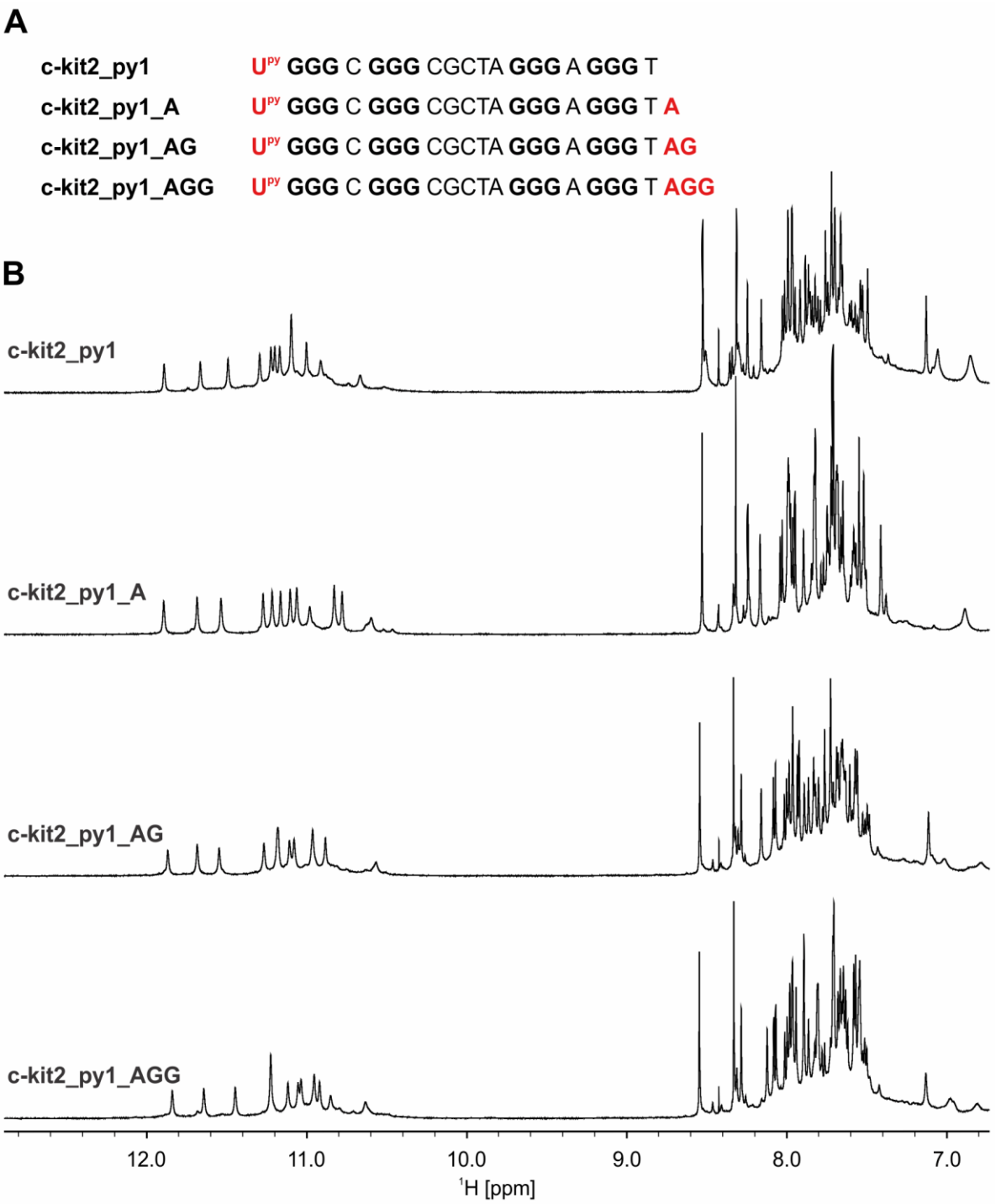

Figure S18. (A) c-kit2_py1 oligonucleotide sequences with 1-3 nucleotide long 3'-overhang (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of c-kit2_py1 and c-kit2_py1 with 1-3 nucleotide long 3'-overhangs. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations were ~0.4 mM.

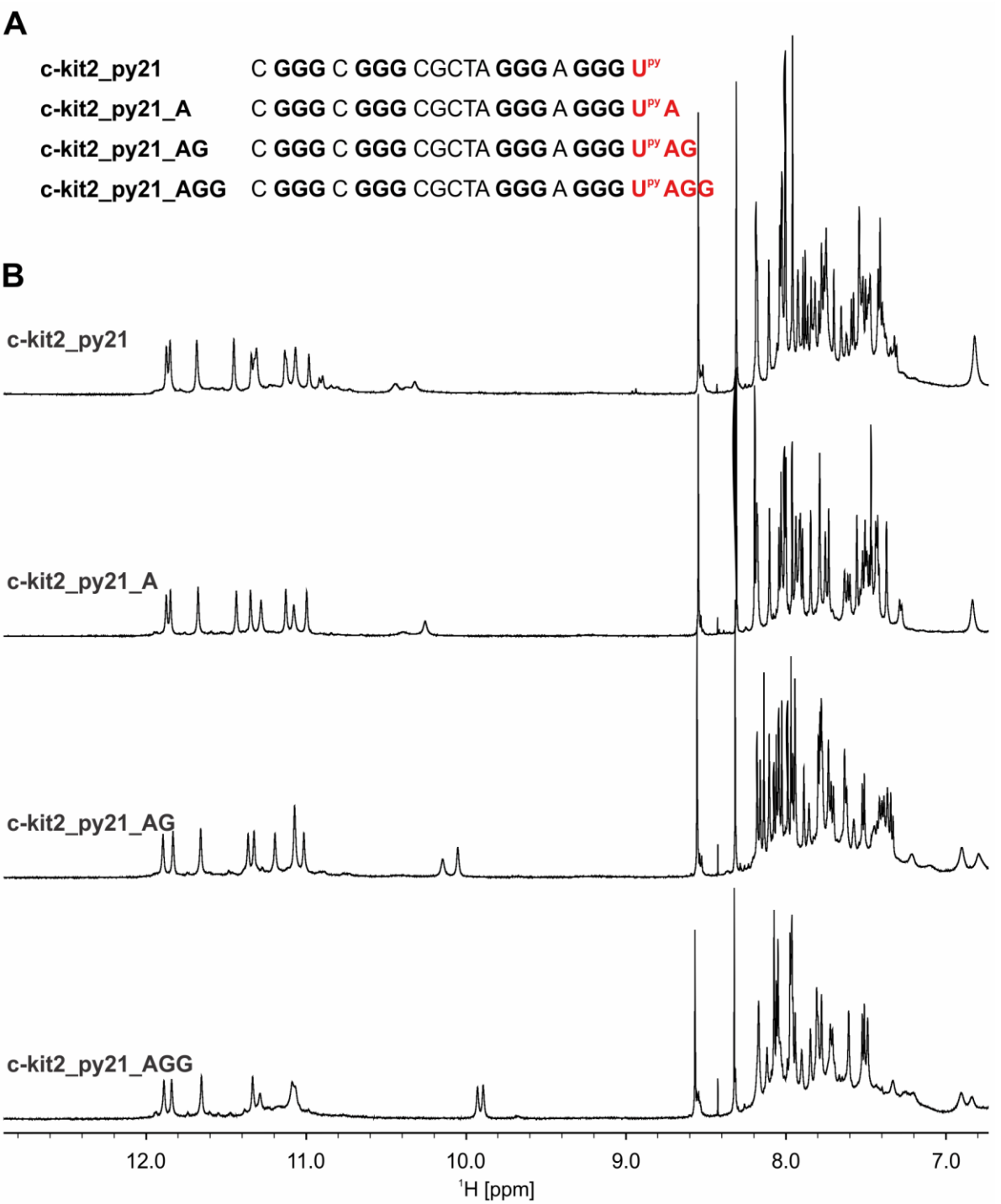

Figure S19. (A) c-kit2_py21 oligonucleotide sequences with 1-3 nucleotide long 3'-overhang (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of c-kit2_py21 and c-kit2_py21 with 1-3 nucleotide long 3'-overhangs. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations were ~0.4 mM.

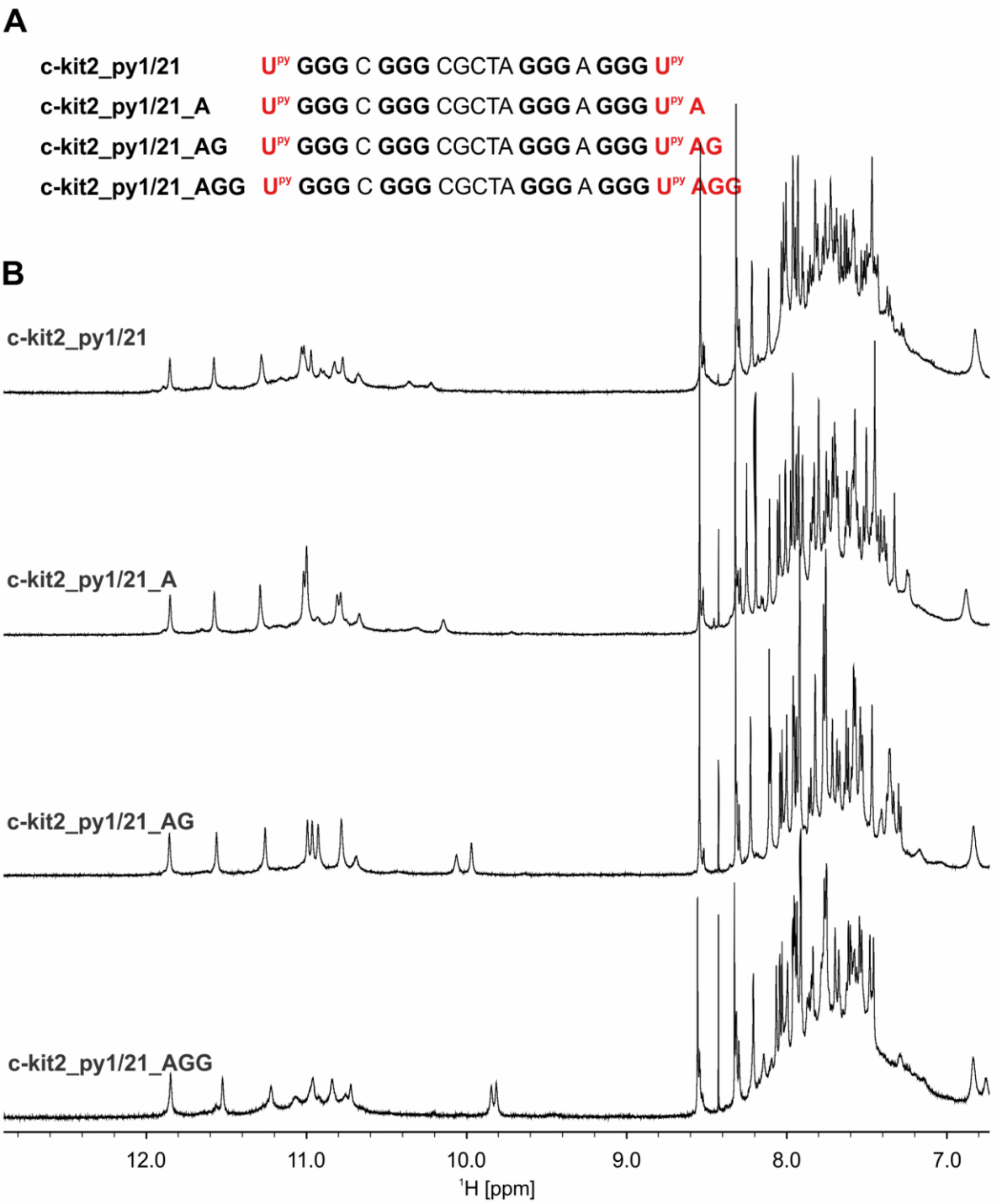

Figure S20. (A) c-kit2_py1/21 oligonucleotide sequences with 1-3 nucleotide long 3'-overhang (highlighted in red). Guanines involved in G-quartet formation are in bold. (B) Imino and aromatic regions of ^{1}H NMR spectra of c-kit2_py1/21 and c-kit2_py1/21 with 1-3 nucleotide long 3'-overhangs. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations were ~0.4 mM.

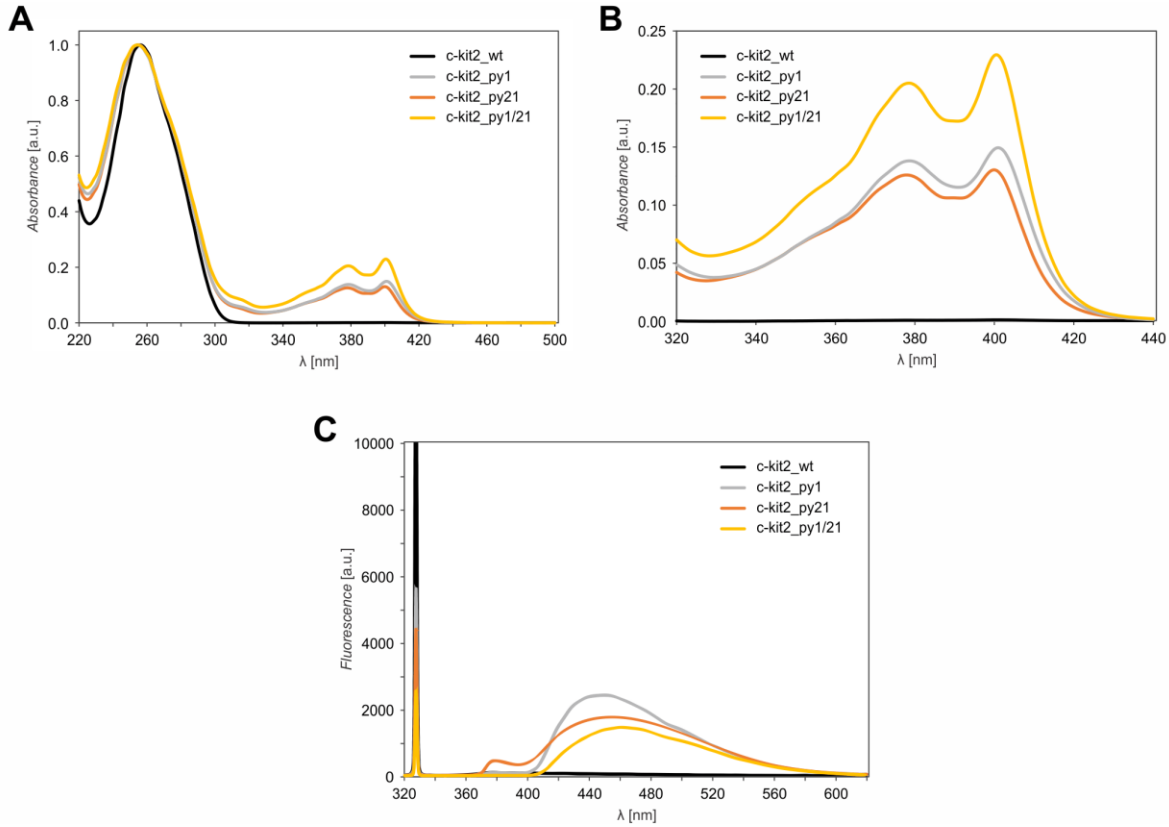

Figure S21. Normalized absorbance spectra of c-kit2 and terminal U^{py} analogues recorded in the range (A) from 220 nm to 500 nm, and (B) in the region typical for absorption of pyrene moiety, from 320 nm to 440 nm. (C) Fluorescence emission spectra of 150 µl c-kit2 and terminal U^{py} analogues in 20 mM KCl and 5 mM K-phosphate buffer at 25 °C. Excitation wavelength of 330 nm was used.

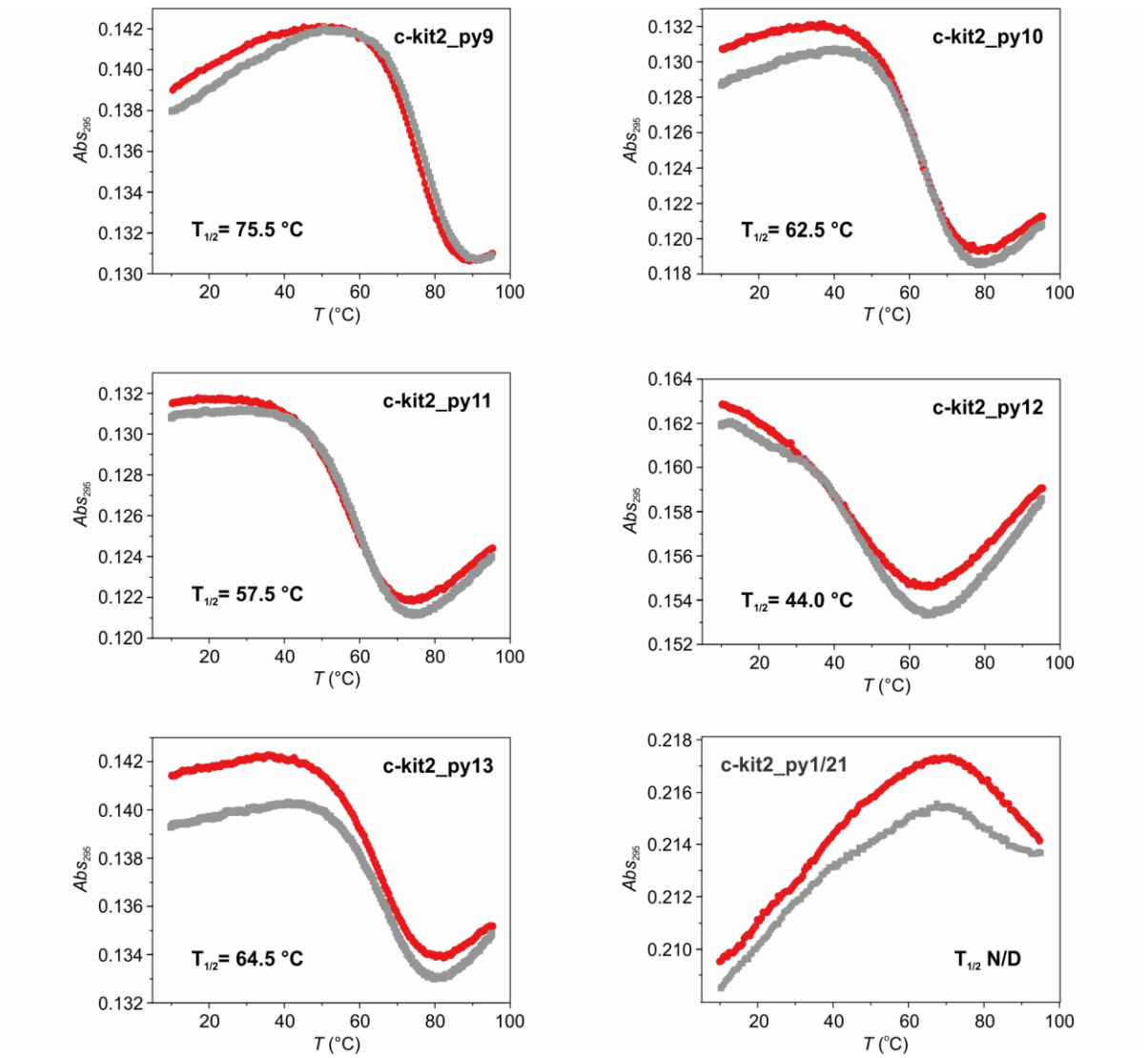

Figure S22. UV melting curves of loop U^{py} analogues and c-kit2_py1/21. Data were acquired for DNA with 20 mM KCl and 5 mM K-phosphate buffer, pH 7, oligonucleotide concentrations were ~10 μM. Thermal melting analysis was initiated by equilibrating samples at 10 °C for 10 min before being heated to 95 °C (red curve) then after 10 min at 95 °C temperature was reduced to 10 °C (grey curve) at a rate of 0.5 °C/min. Absorbance was followed at 295 nm.

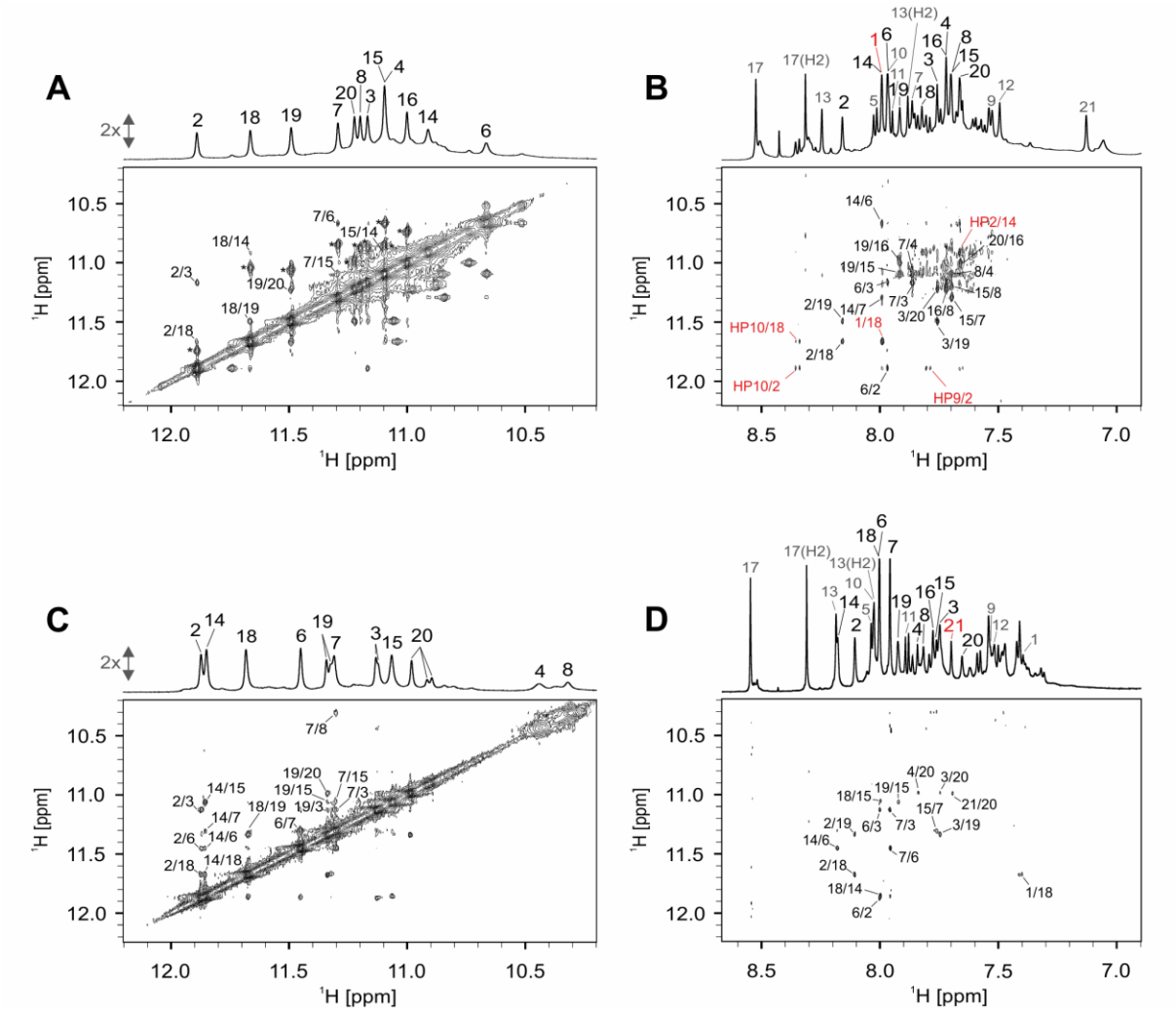

Figure S23. Imino-imino and aromatic-imino regions of 2D NOESY spectrum (τ_m = 250 ms) of (A, B) ckit2_py1 and (C, D) c-kit2_py21. Assignment of NOE contacts involving protons from $U^{py}1$ nucleotide are labelled in red. Exchange cross-peaks between monomeric and dimeric species of c-kit2_py1 in (A) are labelled with black asterisk. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer. Oligonucleotide concentrations varied from 0.9 to 1.0 mM.

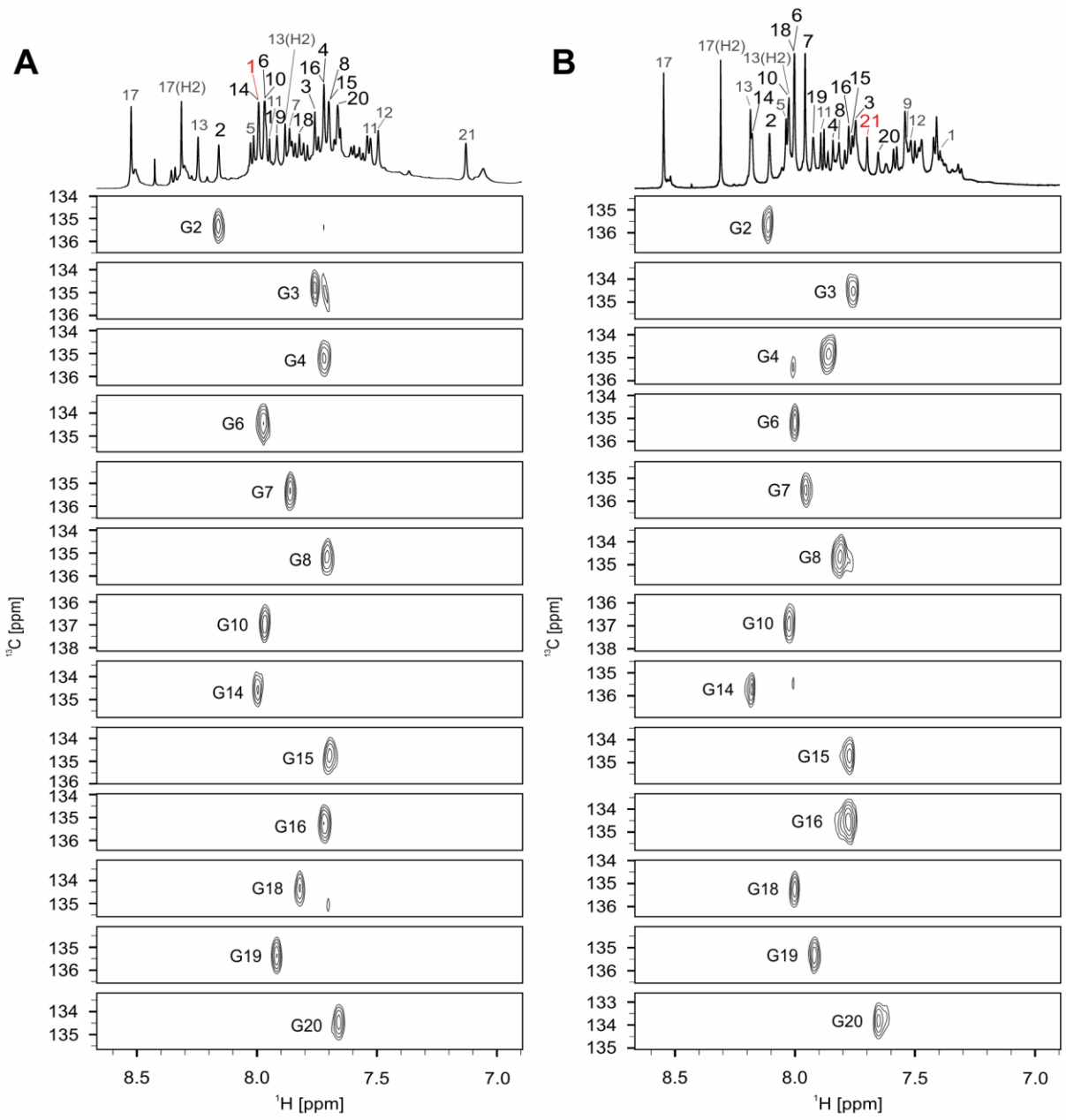

Figure S24. Aromatic region of 1D ^{1}H NMR spectrum above 2D ^{13}C-edited HSQC spectra of (A) c-kit2_py1 and (B) c-kit2_py21 G-quadruplex. HMQC spectra were acquired on partially (10%) site-specific ^{15}N- and ^{13}C-labelled oligonucleotides. Assignment of H8 proton resonances is indicated next to the 2D cross-peaks. Grey colour in 1D ^{1}H NMR spectrum indicates loop nucleotides, while U^{py} is in red. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations of labelled samples varied between 0.4 and 0.6 mM, oligonucleotide concentration of non-labelled sample was 1.0 mM.

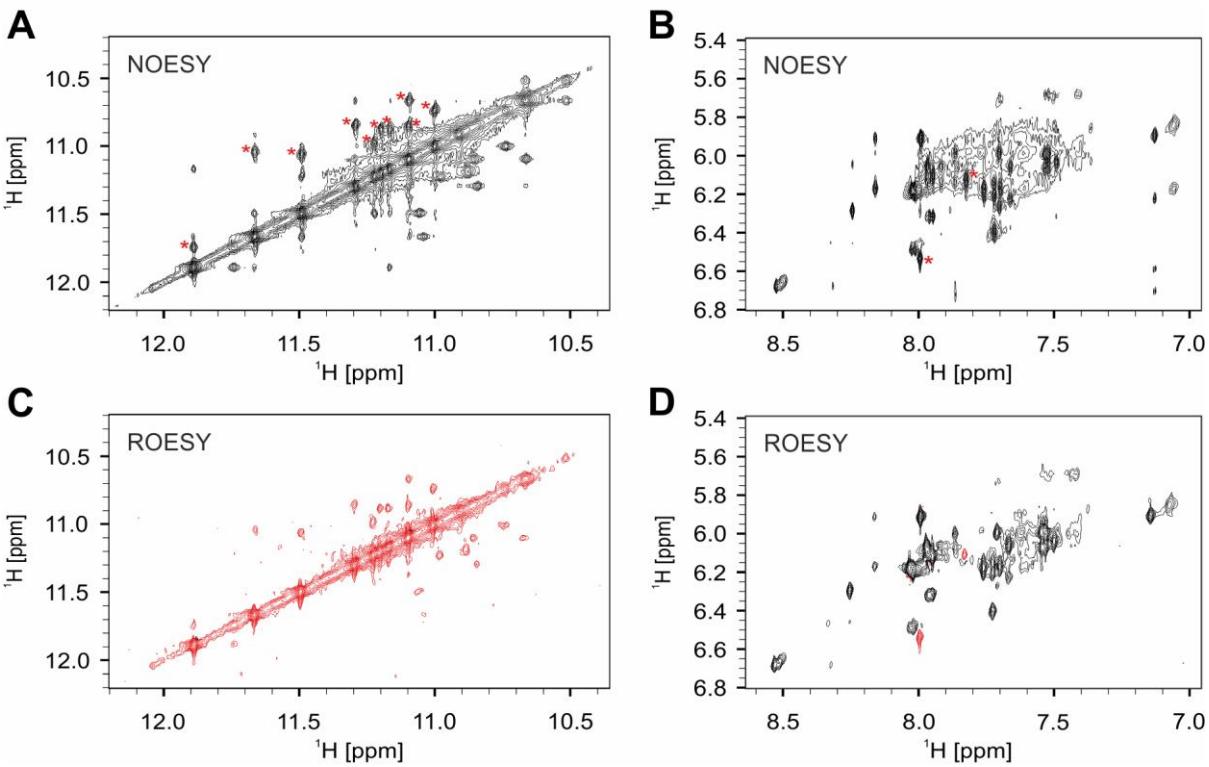

Figure S25. (A, C) Imino-imino and (B, D) aromatic-anomeric regions of (A, B) 2D NOESY spectrum (τ_m = 250 ms) of ckit2_py1 and (C, D) 2D ROESY spectrum (τ_m = 120 ms) of c-kit2_py1. Exchange cross-peaks between monomeric and dimeric species of c-kit2_py1 in (A, B) are labelled with red asterisk. Diagonal and ROE cross-peaks in (C, D) are depicted in red. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer. Oligonucleotide concentration was 1.0 mM.

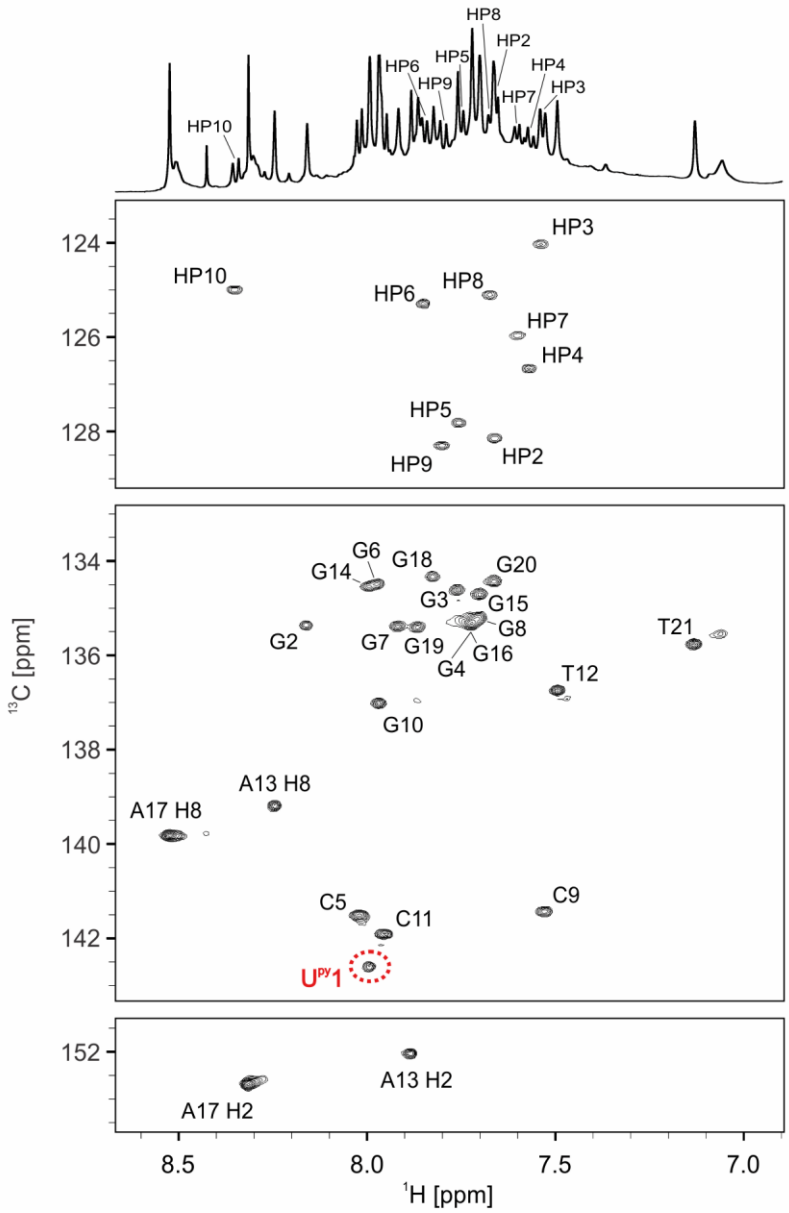

Figure S26. 1D ^{1}H NMR spectrum above 2D ^{1}H-^{13}C HSQC spectrum of c-kit2_py1. Assignment is indicated next to the 2D signals. Red dashed circle marks U^{py}1C6-H6 cross-peak with downfield chemical shift in carbon dimension indicating its *syn* conformation around glycosidic bond. Spectrum was acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations was 1 mM.

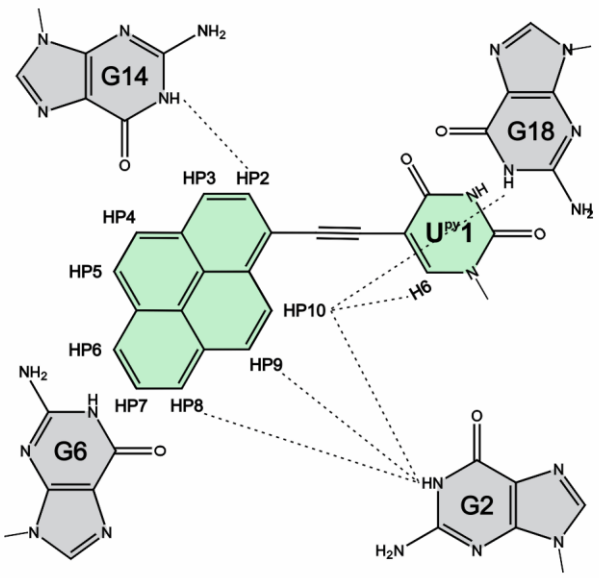

Figure S27. Schematic representation of key internucleotide NOE connectivities with the pyrene moiety in c-kit2_py1 G-quadruplex observed in a NOESY spectrum (τ_m = 250 ms). Guanine bases are in grey and U^{py} in green. NOE contacts are shown as dotted lines.

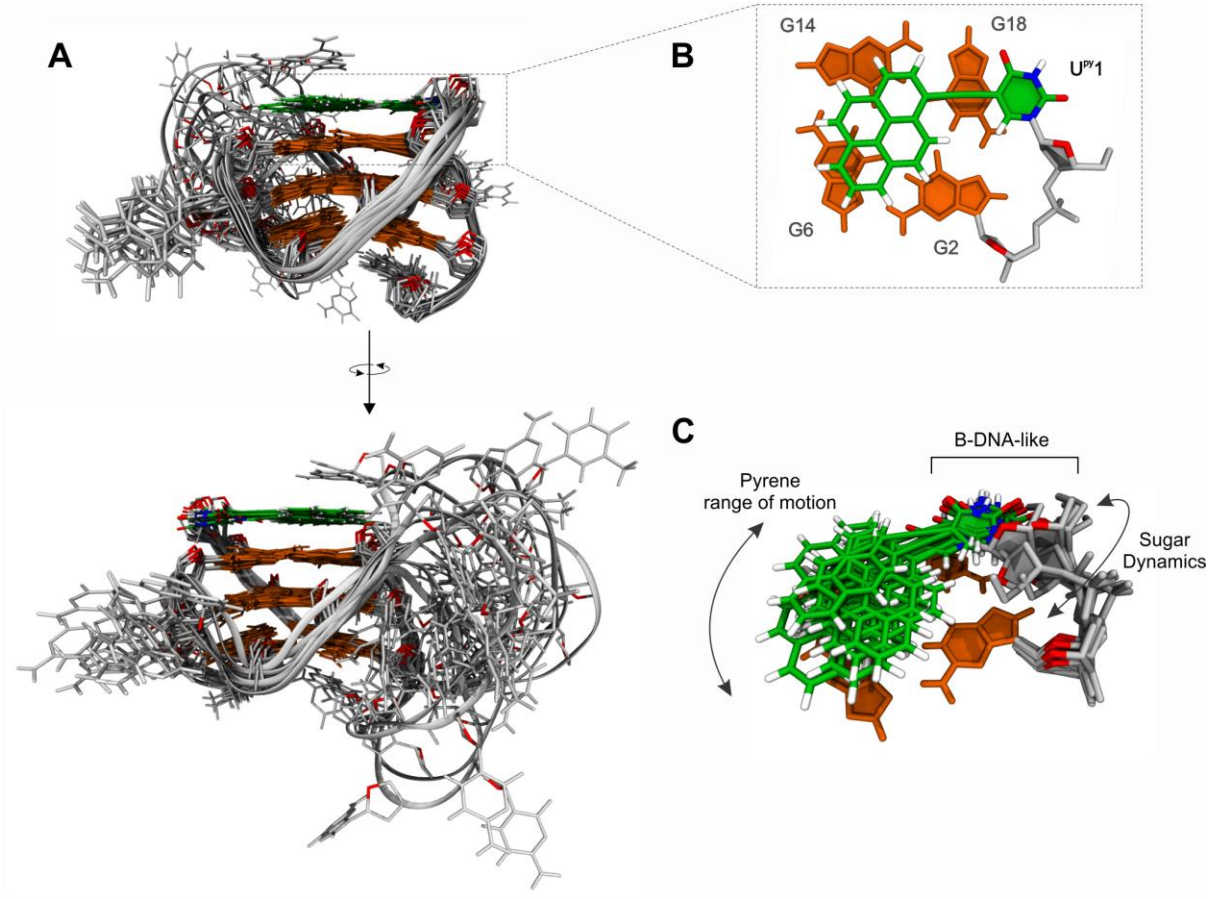

Figure S28. Solution structure of c-kit2_py1. (A) Superposition of 10 structures from the NMR ensemble refined in the explicit solvent without the use of restraints, $U^{py}1$ is shown in green, and G-quartets in orange. Sugar O4' atoms are highlighted in red and remaining loops and backbone are in light grey. (B) Details of $U^{py}1$ stacked on top of the G2-G6-G14-G18 quartet. (C) Superposition of $U^{py}1$ conformations demonstrating range of sliding motion sampled by the pyrene moiety, while the linked uracil adopts a B-DNA-like conformation. Additionally, $U^{py}1$ *syn* ↔ ***anti*** switching was observed contrary to the restrained simulation (Figure S29).

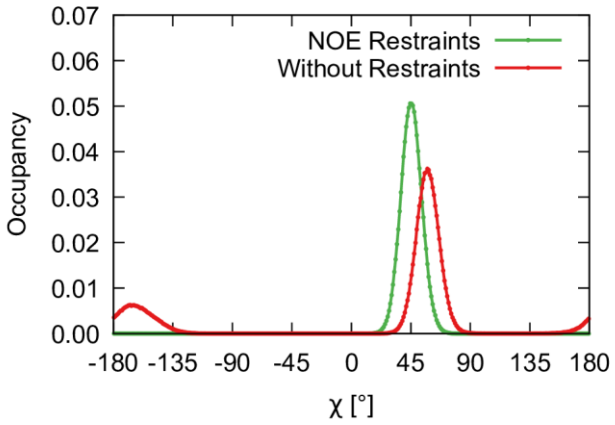

Figure S29. Populations of $U^{py}1$ χ glycosidic torsion values during refinement of the NMR structure by a 1 μs long explicit solvent molecular dynamics of c-kit2_py1 with NOE derived distance restraints and without any restraints. In case of the unrestrained simulations, fast *syn* ↔ *anti* transitions were observed in all models, contrary to the experimental data. This can be attributed to a higher energy state with lower population, or a force field artefact. However, the average value from the unrestrained simulations (28°) falls also into the experimentally observed *syn* range.

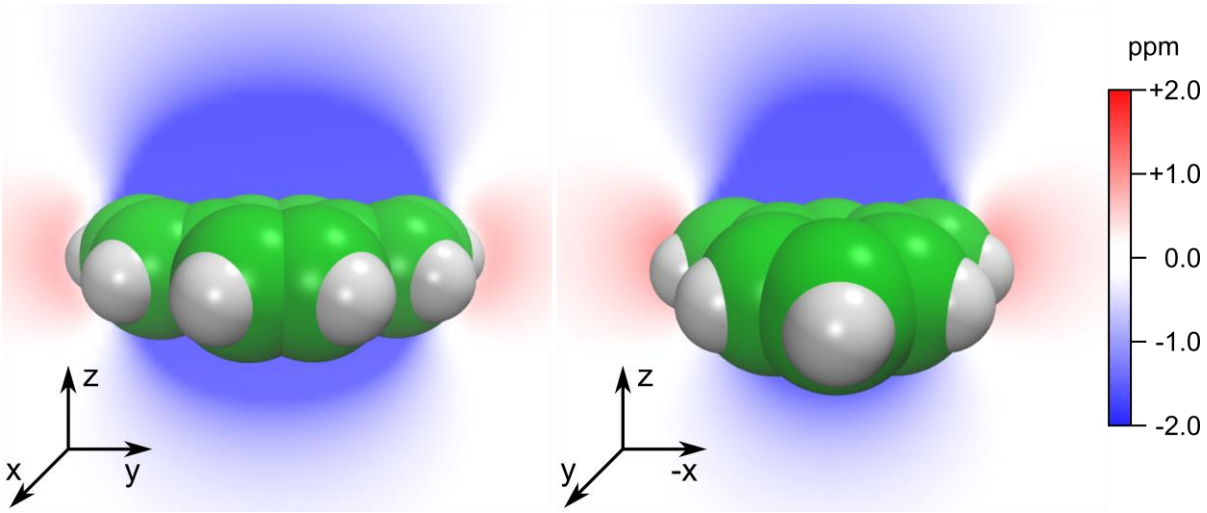

Figure S30. Slices through nucleus-independent chemical shift (NICS) volumetric data calculated for pyrene. Shielding is shown as blue region, whereas deshielding as red region. Pyrene is visualized using a vdW model.

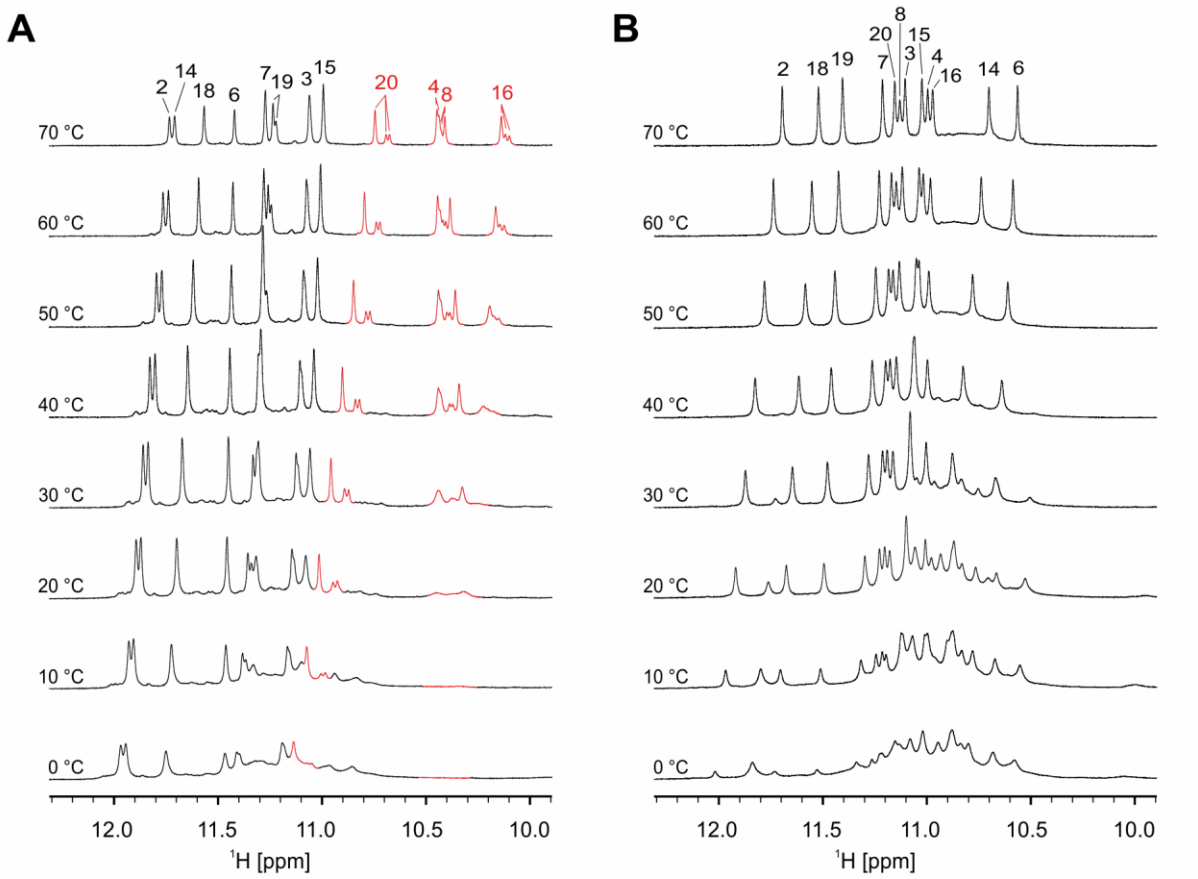

Figure S31. Temperature-dependent 1D ^{1}H NMR spectra of (A) c-kit2_py21 and (B) c-kit2_py1. Red colour in (A) highlights signals corresponding to imino protons of the G4-G8-G16-G20 quartet adjacent to the pyrene moiety. Assignment of H1 proton resonances is indicated above 1D signals in spectrum measured at (A) 70 °C and (B) 50 °C. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer. Oligonucleotide concentrations varied between 0.9 and 1.0 mM.

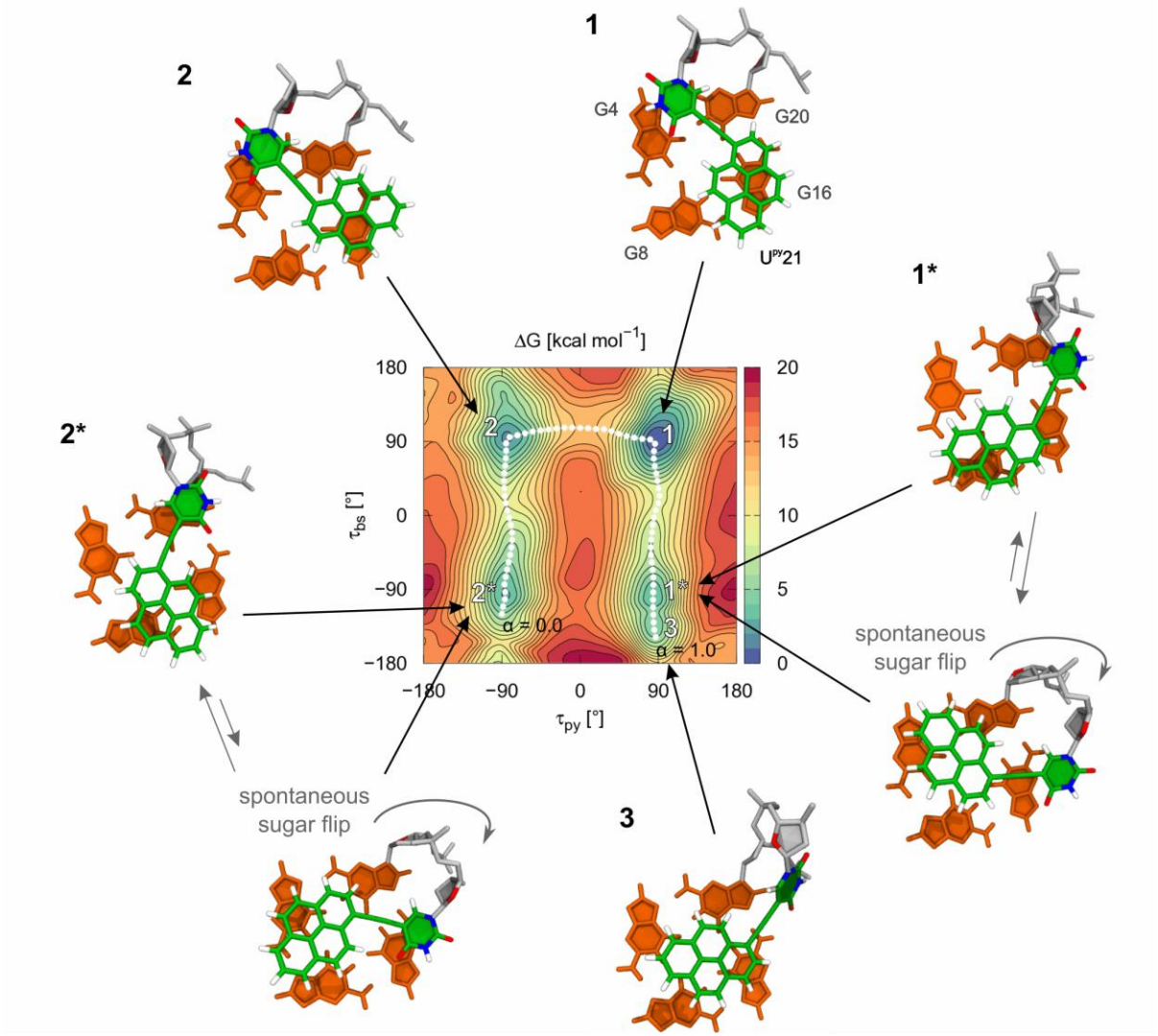

Figure S32. Free energy surfaces within τ_{py} and τ_{sb} for c-kit2_py21 with representative structures showing different stacking modes of pyrene moiety. White string of beads shows calculated minimum free energy path connecting found minima. Employed collective variables are depicted in Figure S7.

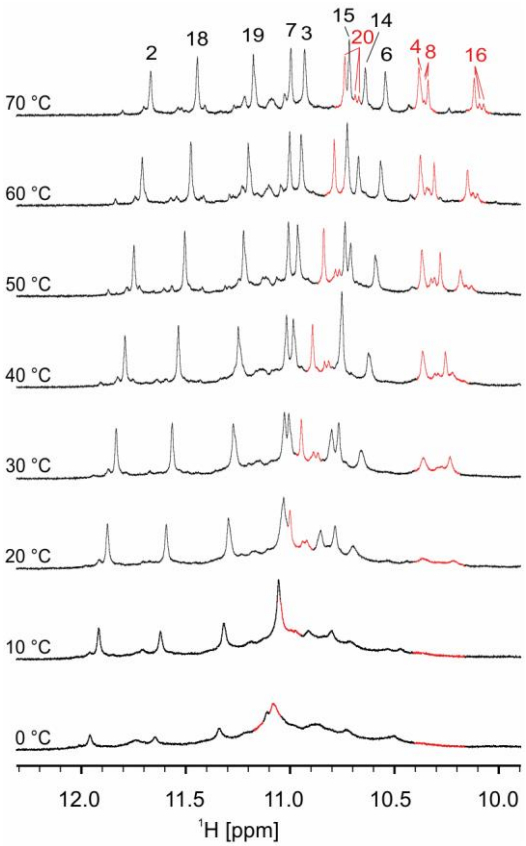

Figure S33. Temperature dependent 1D ^{1}H NMR spectra of c-kit2_py1/21. Red colour highlights signals corresponding to imino protons of the G4-G8-G16-G20 quartet adjacent to the pyrene moiety. Assignment of H1 proton resonances is indicated above 1D signals in spectrum measured at 80 °C. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer. Oligonucleotide concentrations was 0.7 mM.

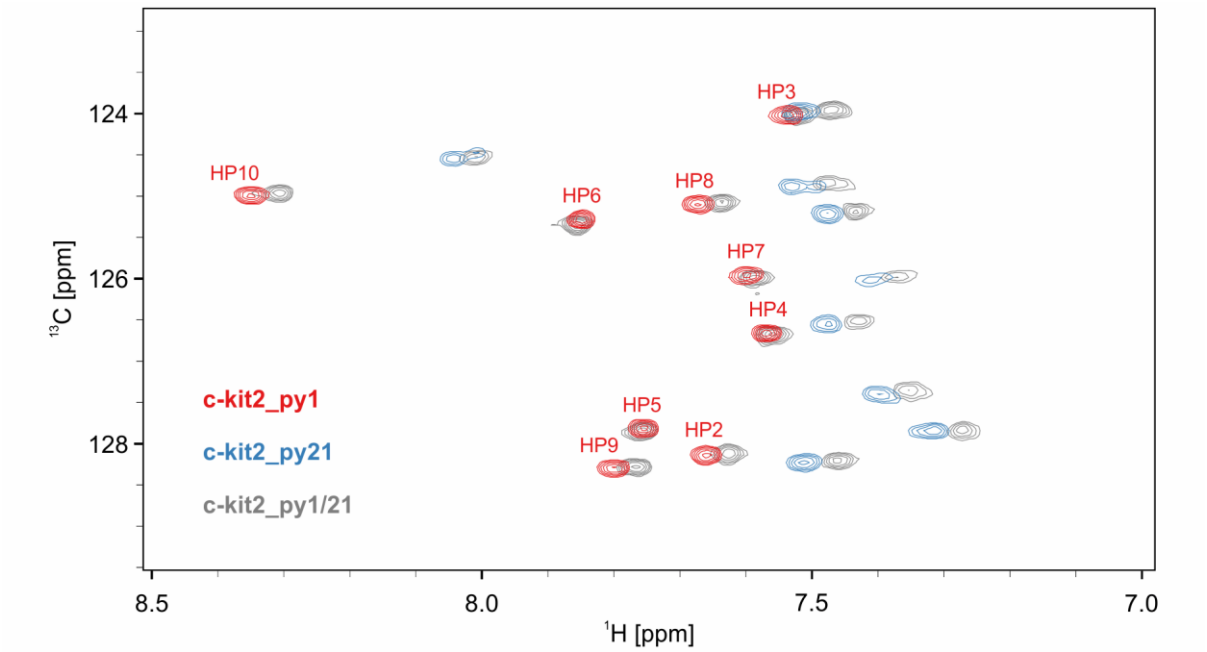

Figure S34. Overlay of 2D ^{1}H-^{13}C HSQC spectra of c-kit2_py1 (red), c-kit2_py21 (blue) and c-kit2_py1/21 (grey). Assignment of resonances from pyrene moiety located at 5'-end of the G-core is indicated next to the 2D signals. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations varied between 0.7 and 1.0 mM.

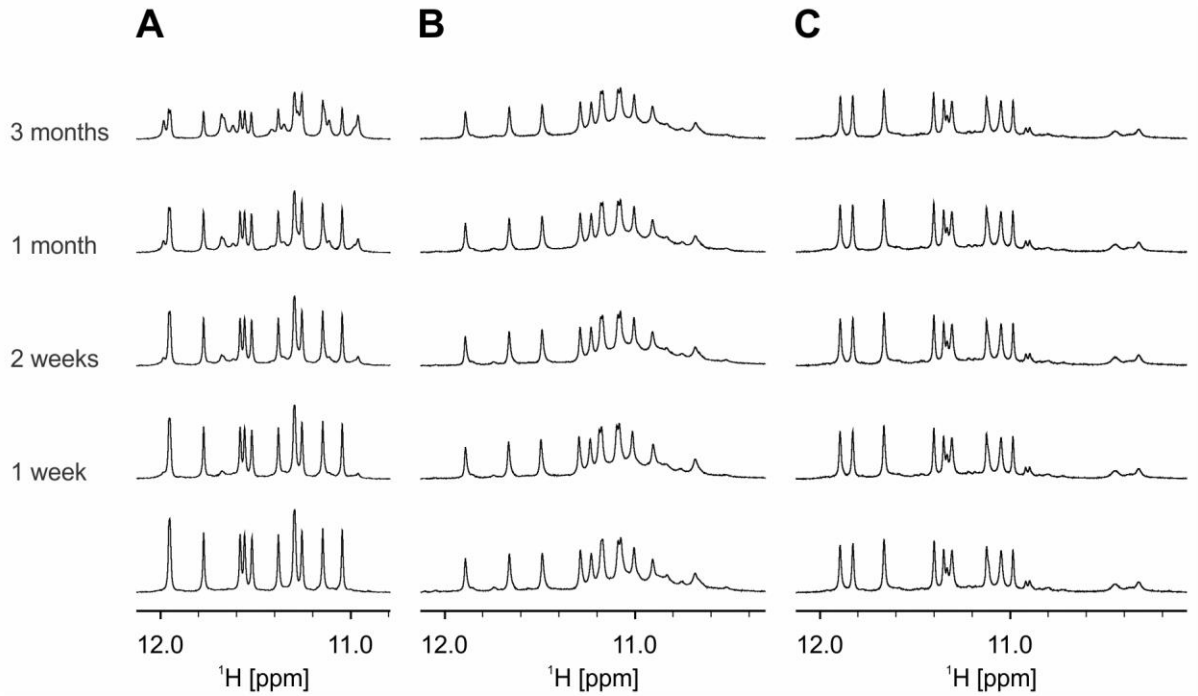

Figure S35. Effect of time on the fold of (A) c-kit2, (B) c-kit2_py1, and (C) c-kit2_py21 G-quadruplexes. Imino-proton regions of NMR spectra were recorded on the samples immediately after annealing, followed by 1 week, 2 weeks, 1 month and 3 months of incubation at room temperature. While the monomeric form of c-kit2 G-quadruplex gradually converted to the dimeric form, both analogues are stable even after several months. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations varied between 0.8 and 1.0 mM.

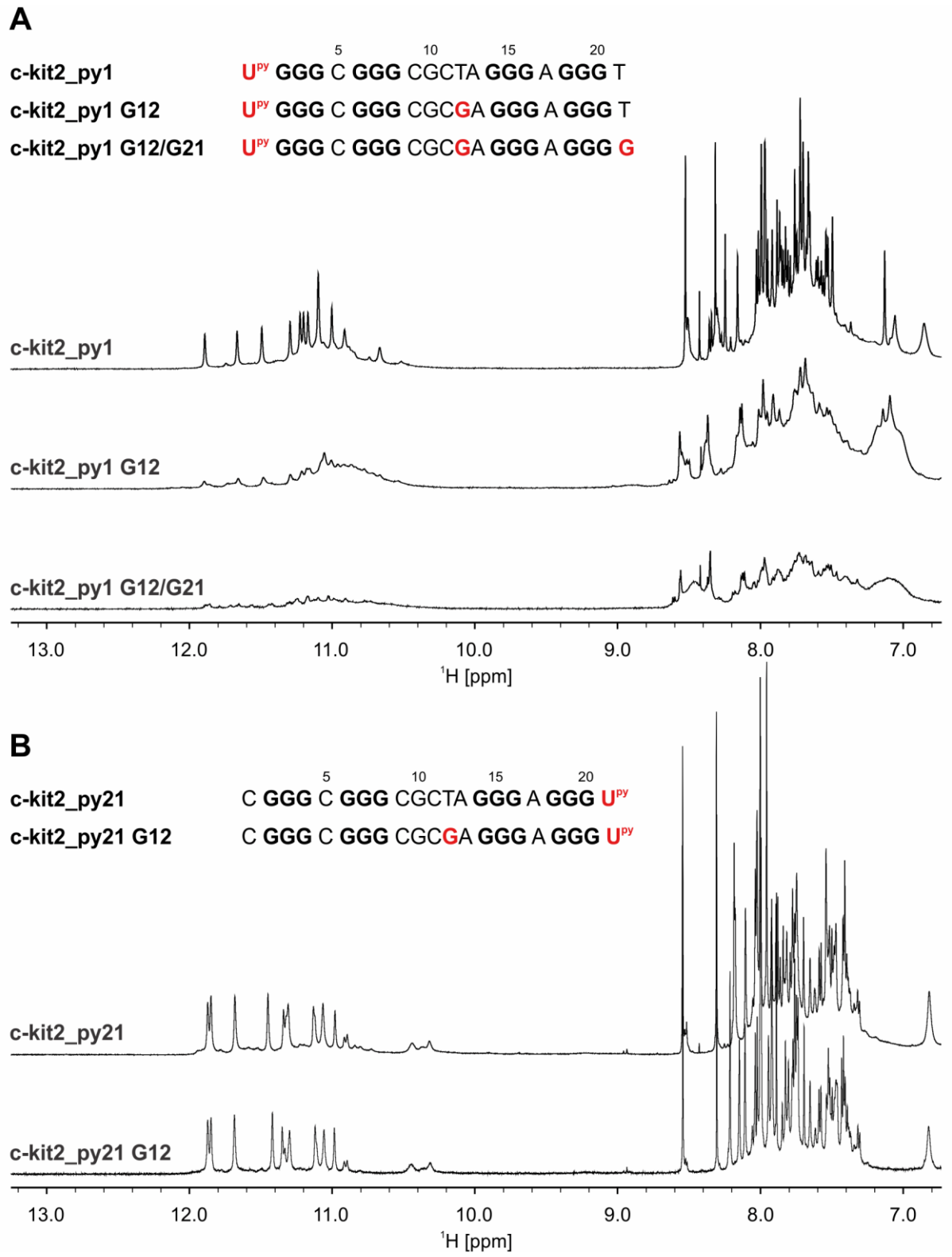

Figure S36. Imino and aromatic regions of 1D ^{1}H NMR spectra of modified analogues of (A) c-kit2_py1 and (B) c-kit2_py21. Spectra were recorded in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C. Oligonucleotide concentrations of c-kit2_py1 and c-kit2_py21 were ~0.9 mM per strand, oligonucleotide concentrations of their analogues were ~0.5 mM. Oligonucleotide sequences are reported above spectra.

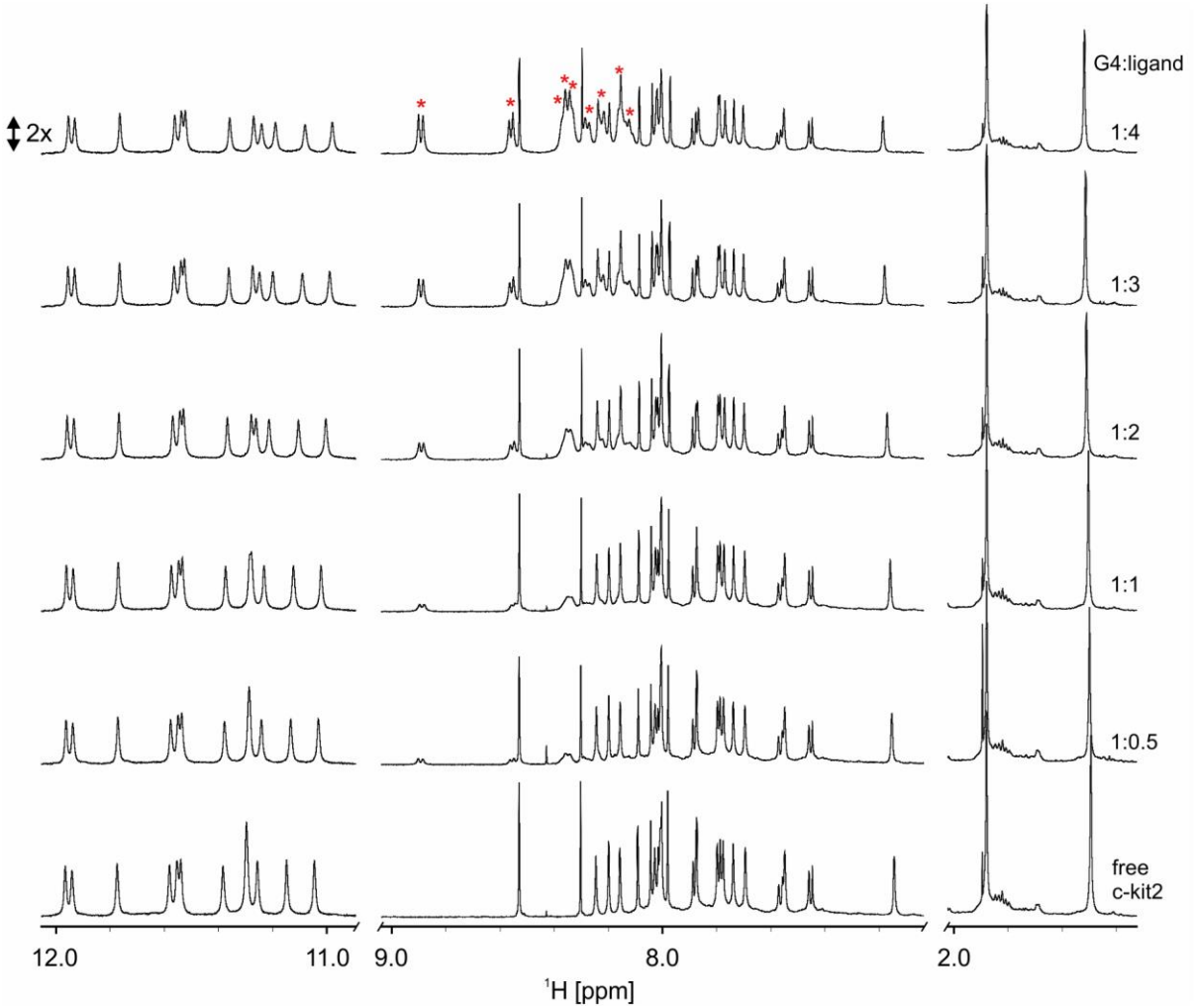

Figure S37. 1D ^{1}H spectra of 100 μM c-kit2 sample with increasing additions of sodium salt of 1-pyrensulfonic acid. Proton signals of pyrene group are marked with red asterisks. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C.

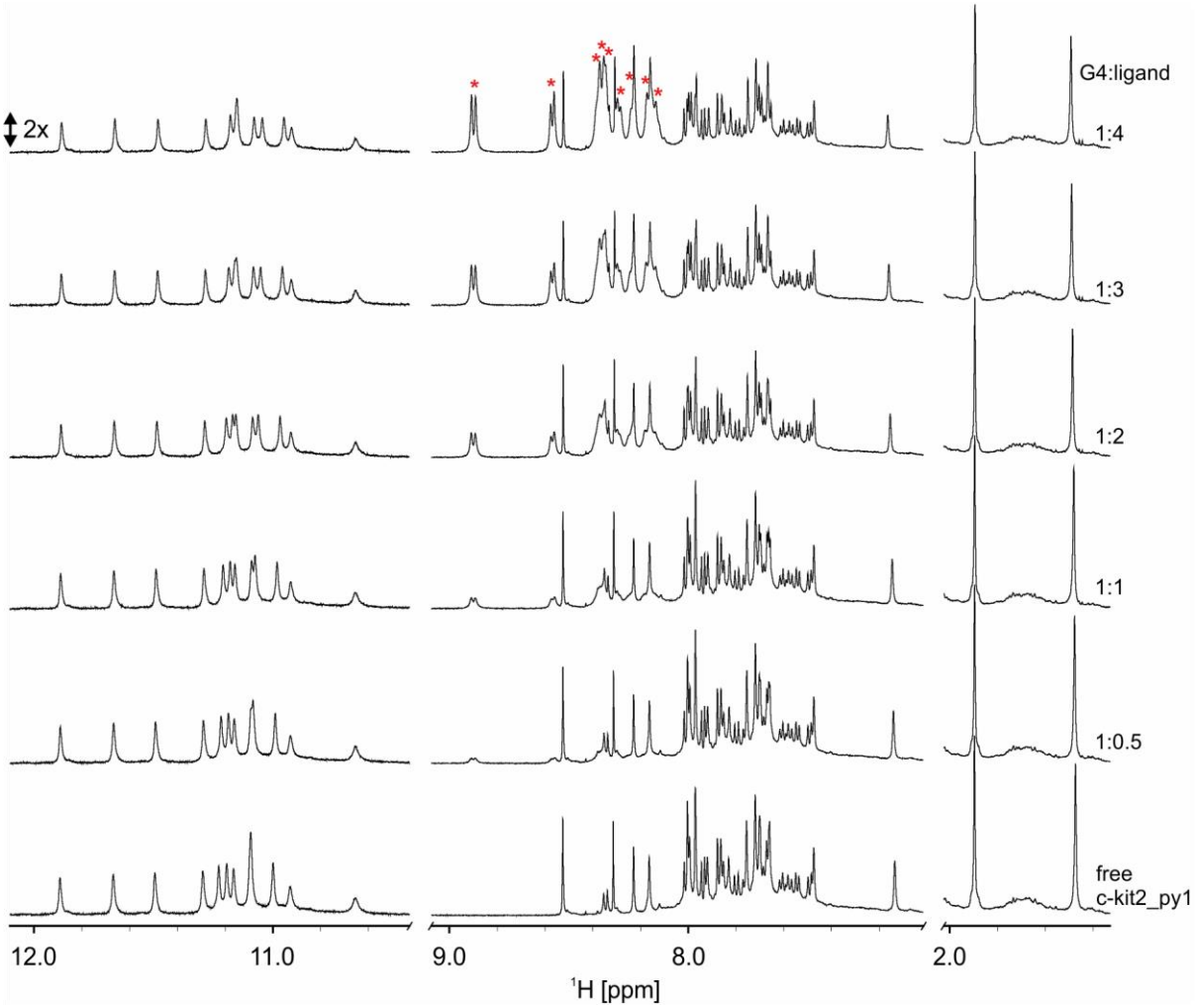

Figure S38. 1D ^{1}H spectra of 100 μM c-kit2_py1 sample with increasing additions of sodium salt of 1-pyrensulfonic acid. Proton signals of pyrene group are marked with red asterisks. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C.

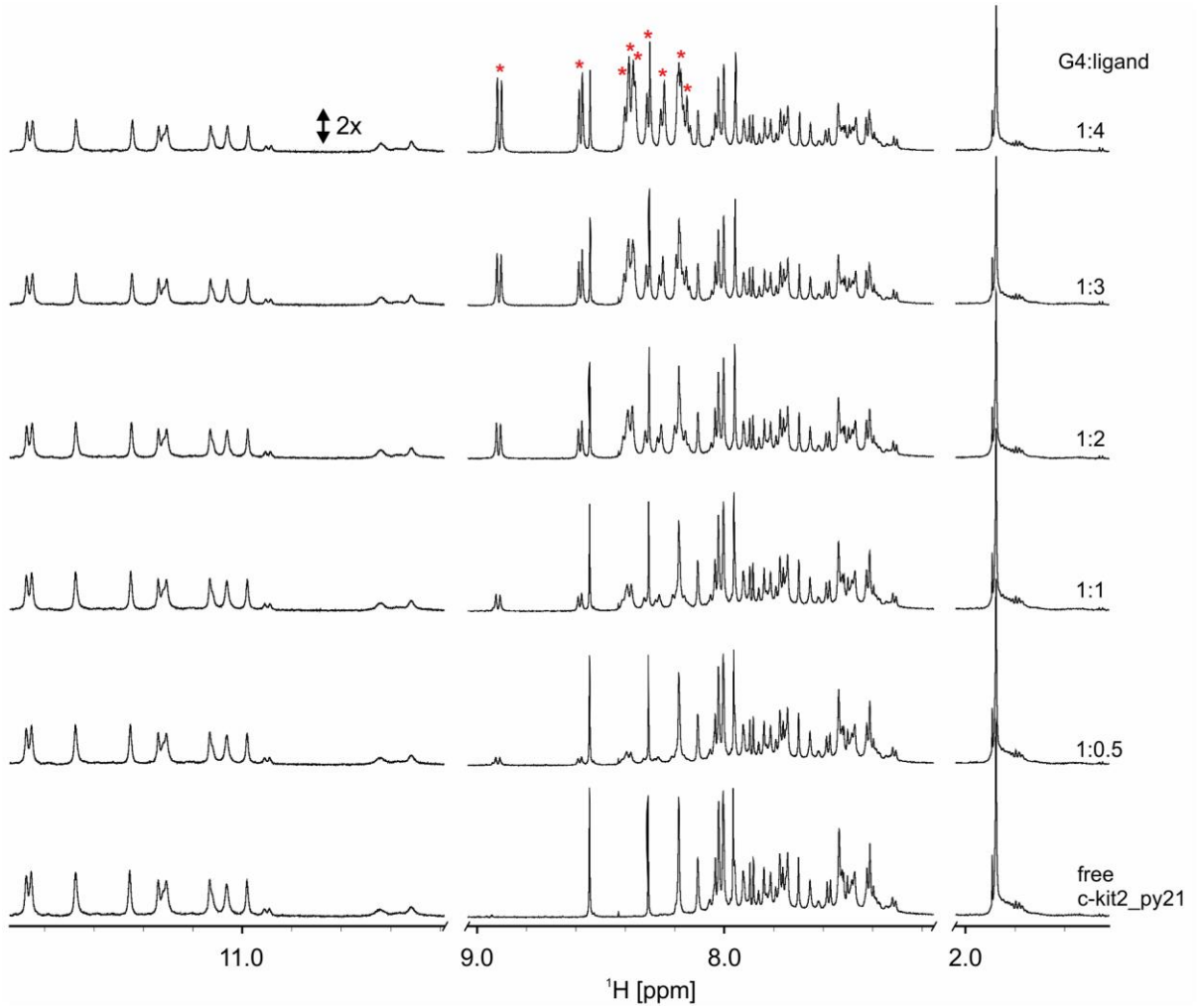

Figure S39. 1D ^{1}H spectra of 100 μM c-kit2_py21 sample with increasing additions of sodium salt of 1-pyrensulfonic acid. Proton signals of pyrene group are marked with red asterisks. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C.

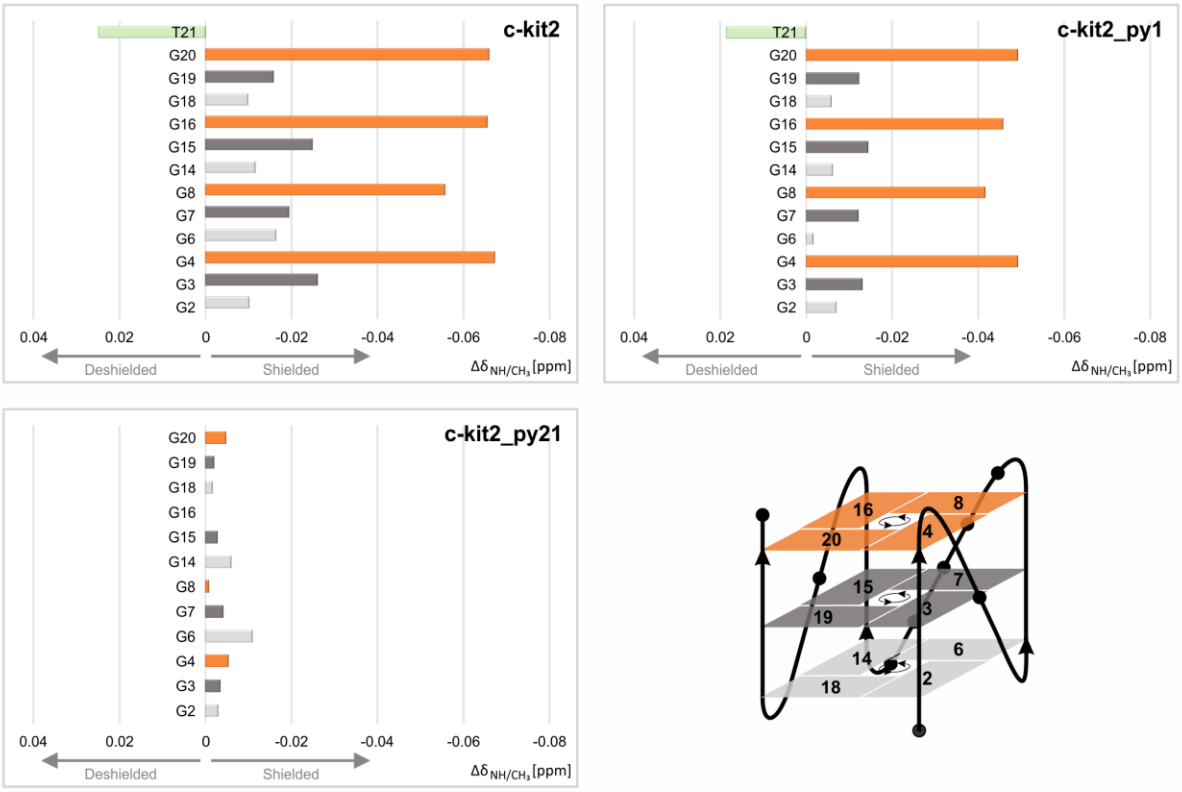

Figure S40. Histograms of NMR chemical-shift perturbations induced in c-kit2, c-kit2_py1, and c-kit2_py21 in the presence of four molar equivalents of sodium salt of 1-pyrensulfonic acid. c-kit2 and c-kit2_py1 exhibit similar profiles of chemical shift changes with the most significant changes being observed for nucleotides of the G4-G8-G16-G20 quartet. Differences in the shift changes between these two complexes may be attributed to differences in affinities. On the contrary, no interaction was observed in case of c-kit2_py21. Schematic representation of the c-kit2 G-quadruplex (right bottom) with G-quartet colour coding used in histograms. Green colour corresponds to the chemical shift perturbation of methyl group from nucleotide 21.

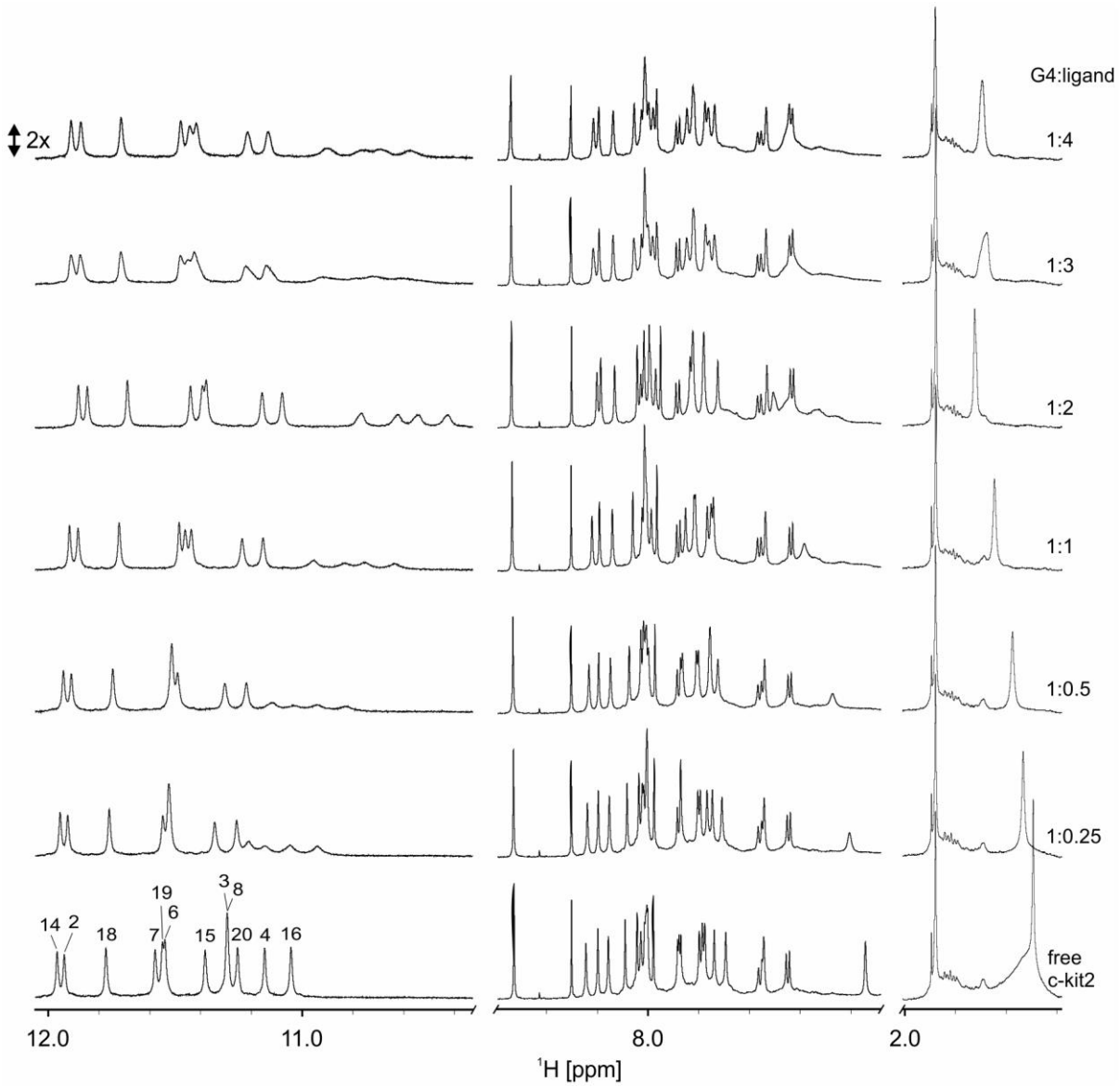

Figure S41.1D ^{1}H spectra of 100 µM c-kit2 sample with increasing additions of hydroxypyrene. Proton signals of hydroxypyrene were not observed. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C.

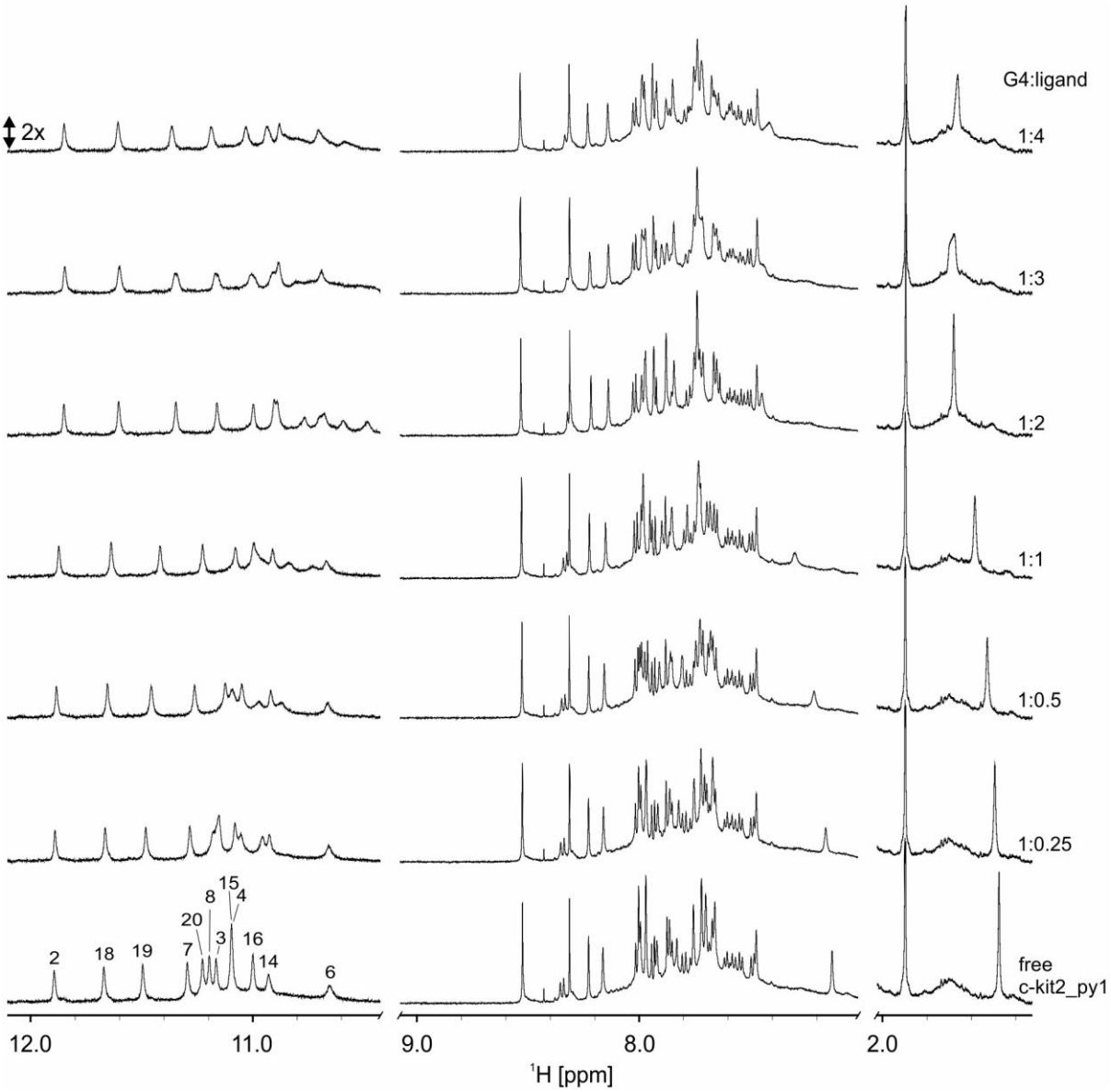

Figure S42.1D ^{1}H spectra of 100 μM c-kit2_py1 sample with increasing additions of hydroxypyrene. Proton signals of hydroxypyrene were not observed. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C.

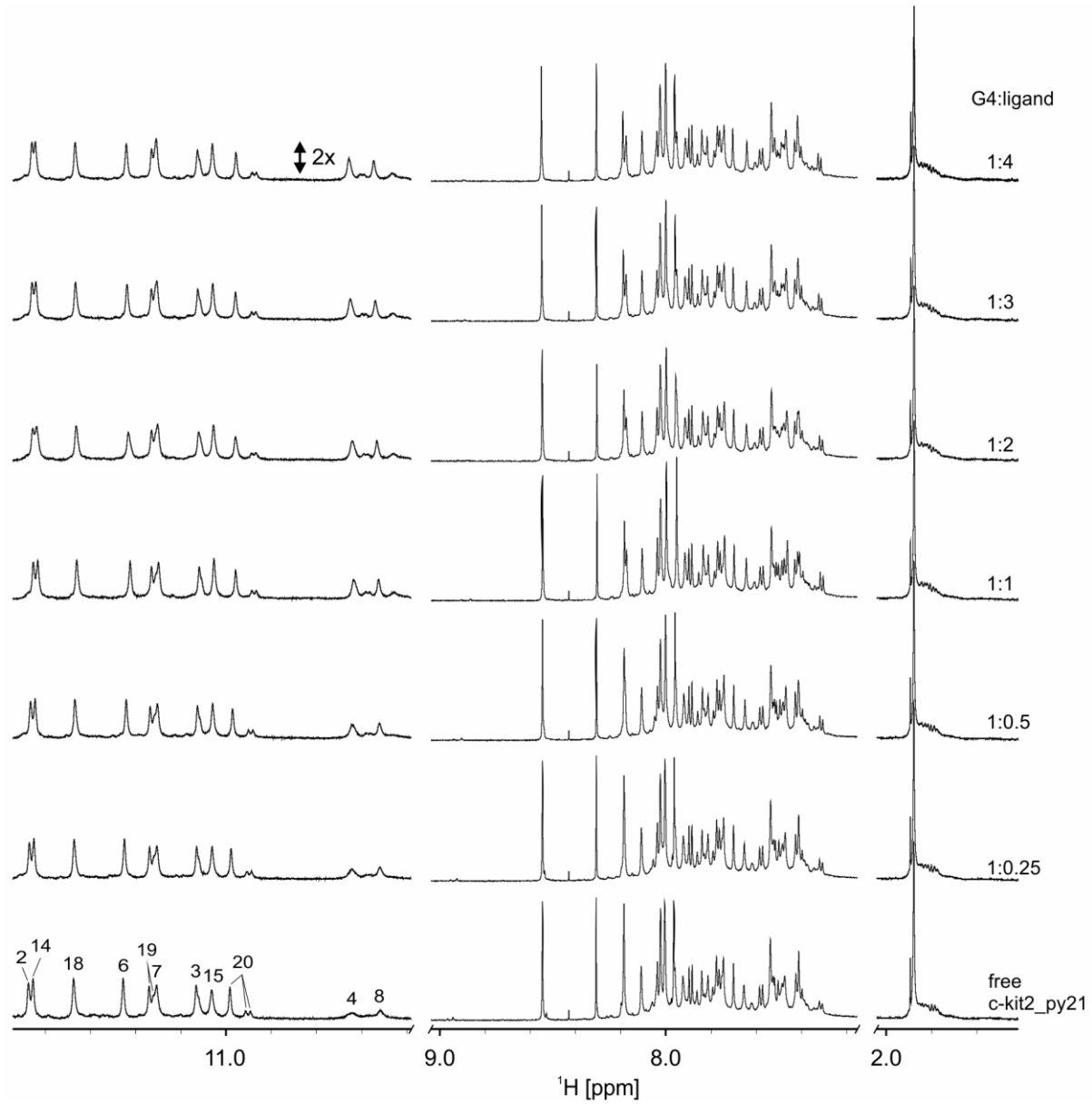

Figure S43.1D ^{1}H spectra of 100 µM c-kit2_py21 sample with increasing additions of hydroxypyrene. Proton signals of hydroxypyrene were not observed. Spectra were acquired in 90% H_2O and 10% D_2O, 20 mM KCl, and 5 mM K-phosphate buffer, pH 7, on a 600 MHz NMR spectrometer at 25 °C.

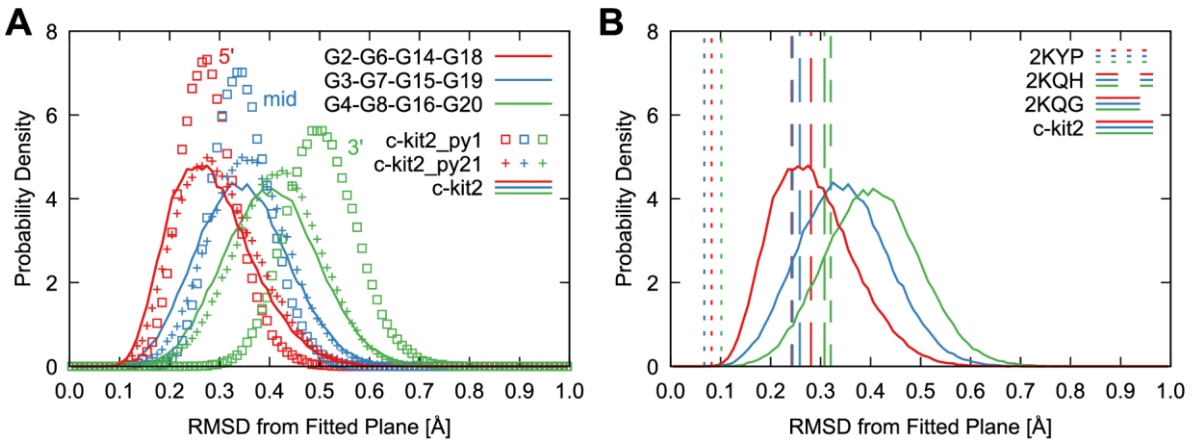

Figure S44. (A) Comparison of root-mean-square deviation (RMSD) from a plane, which was best-fitted individually to each G-quartet in c-kit2, c-kit2_py1, and c-kit2_py21 molecular dynamics. The 5' G-quartet turned out as the most planar, whereas the 3' G-quartet is the least planar. (B) Comparison of published NMR structures and our molecular dynamics. All G-quartets were identically and almost ideally planar in the c-kit2 NMR structure 2KYP. While 2KQH and 2KQG ensembles showed fluctuations closer to our molecular dynamics, clear separation of G-quartets was still not observed. Average values are reported for the NMR ensembles, because number of samples is insufficient to draw reasonable histograms and are marked by vertical lines. Structures were taken from a 1 μs unrestrained molecular dynamics in case of c-kit2 and c-kit2_py21 and from the last 100 ns of NMR refinement of ten model of c-kit2_py1 ensemble.

SUPPLEMENTARY TABLES

Table S1. Effect of pyrene incorporation on thermal stability

Oligonucleotides	**$T_{1/2\ (major)}$ [°C]**	**ΔT [°C]** [#]	**ΔH_m [kcal·mol^{-1}]**
c-kit2	59.6 ± 0.1		44.7 ± 0.7
c-kit2_py1	87.4 ± 0.2	+27.8	57.2 ± 1.2
c-kit2_py21	80.5 ± 0.2	+20.9	48.0 ± 0.8

[#]Difference compared to the melting temperature of c-kit2.

Table S2. Pseudoration phases for all sugars of c-kit2_py1 NMR ensemble refined with experimentally derived restraints.

	Pseudorotation Phase [°]										
Nucleotide / Model	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Avg.**
$U^{py}1$	162.5	140.5	178.4	169.9	170.8	160.0	158.8	150.0	152.6	160.8	160.4
G2	185.7	167.3	180.8	184.4	185.8	171.0	186.0	182.6	185.3	134.8	176.4
G3	164.4	134.6	132.8	136.2	140.3	159.2	136.2	179.4	146.6	131.0	146.1
G4	187.9	162.4	151.3	162.4	149.2	163.4	153.1	161.2	187.5	158.4	163.7
C5	179.8	186.5	163.7	146.9	178.9	178.9	132.9	151.4	143.4	158.1	162.1
G6	158.7	151.2	156.2	154.9	169.8	141.9	154.5	154.6	159.3	162.4	156.4
G7	128.0	132.3	134.1	130.3	135.9	133.9	133.8	130.8	136.4	137.9	133.3
G8	186.1	162.4	155.4	154.2	146.7	167.0	155.0	180.4	144.7	171.6	162.4
C9	135.4	168.4	138.6	177.3	150.2	153.7	160.8	172.1	135.2	143.1	153.5
G10	139.7	186.7	165.7	140.4	175.7	187.0	183.2	186.0	163.6	178.2	170.6
C11	166.8	157.6	158.9	130.6	176.7	177.0	162.3	153.1	171.4	163.2	161.8
T12	149.8	144.5	138.2	172.6	146.3	147.0	146.4	139.7	139.7	144.1	146.8
A13	156.0	137.0	147.1	164.5	142.9	146.0	139.2	158.1	137.1	144.4	147.2
G14	181.2	161.3	175.7	172.9	165.8	180.5	168.9	184.0	184.0	151.3	172.6
G15	144.3	141.1	146.9	137.8	141.9	144.1	138.2	148.9	143.0	139.9	142.6
G16	156.0	154.5	155.0	148.0	157.8	134.7	169.5	152.5	139.5	153.9	152.1
A17	186.4	186.1	155.2	180.9	180.7	189.6	156.7	139.4	175.2	180.9	173.1
G18	170.9	176.7	135.4	133.9	179.7	168.5	176.7	153.3	186.4	179.0	166.1
G19	149.7	124.0	141.9	149.3	140.7	118.9	138.7	141.3	125.9	138.1	136.9
G20	135.1	134.1	159.5	146.7	155.9	179.4	146.4	164.0	138.1	136.2	149.5
T21	137.8	134.7	130.9	136.3	138.9	130.2	135.5	134.9	135.2	132.3	134.7

Table S3. Pseudoration phases for all sugars of c-kit2_py1 NMR ensemble refined with no restraints.

	Pseudorotation Phase [°]										
Nucleotide / Model	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Avg.**
$U^{py}1$	159.5	155.9	140.6	182.2	142.3	122.0	127.5	139.0	150.0	150.6	147.0
G2	179.2	159.5	204.8	128.9	200.7	184.3	165.2	161.4	172.5	151.4	170.8
G3	140.2	116.1	115.3	132.8	138.3	141.7	140.7	100.1	101.7	146.6	127.4
G4	155.4	162.2	175.1	152.1	149.9	153.1	152.3	159.7	164.8	165.0	159.0
C5	150.1	154.7	154.3	155.9	176.0	175.6	164.0	124.8	155.0	143.8	155.4
G6	177.7	170.2	168.8	158.5	196.6	190.4	144.7	180.7	160.0	192.7	174.0
G7	107.1	64.5	146.5	140.7	138.0	139.2	147.1	143.5	143.0	133.7	130.3
G8	125.2	171.2	153.7	10.9	142.3	146.5	160.1	165.7	156.7	140.0	137.2
C9	116.2	150.0	159.2	123.4	179.7	178.0	110.9	145.7	115.7	154.4	143.3
G10	144.3	153.2	138.6	224.3	157.5	156.7	129.9	149.9	148.5	146.5	154.9
C11	162.4	174.4	157.1	119.1	131.1	126.6	148.8	135.0	156.0	147.6	145.8
T12	147.7	121.0	130.6	142.2	162.0	171.0	132.3	122.9	120.1	145.1	139.5
A13	125.6	129.0	138.7	2.4	196.4	143.6	131.1	153.8	337.6	40.0	139.8
G14	179.4	176.9	159.7	186.2	136.3	122.0	189.4	86.5	131.9	139.0	150.7
G15	107.8	146.0	91.3	136.6	134.2	111.9	139.4	143.6	104.2	102.6	121.8
G16	147.5	169.5	152.8	145.1	157.1	112.1	149.6	165.7	136.6	152.2	148.8
A17	94.6	176.5	145.6	145.3	163.0	120.5	162.5	179.9	142.8	126.3	145.7
G18	124.8	182.0	170.1	190.3	186.6	188.2	178.7	175.4	76.7	185.4	165.8
G19	147.3	143.3	139.9	141.1	144.0	138.4	141.7	138.7	166.8	140.0	144.1
G20	150.6	158.7	152.3	155.1	145.9	156.1	155.0	149.1	171.4	165.4	156.0
T21	124.5	94.3	101.5	141.4	128.7	148.7	96.4	104.1	134.5	119.8	119.4

Table S4. Values of experimental difference in chemical shift between c-kit2 and c-kit2_py1 $\Delta\delta_{exp}$ compared to values obtained by fitting NMR ensemble of c-kit2_py1 refined with restraints $\Delta\delta_{calc,rst}$ and without any restraints $\Delta\delta_{calc,unrst}$ into calculated nucleus-independent chemical shift (NICS) of pyrene. NICS values for both ensembles suggest G6 as one of the most shielded imino protons. However, only unrestrained ensemble shows G14 shielded more than G2 as observed in NMR spectra. This suggests that pyrene is present above G14 during a transient state, which might not be captured by the time-averaged restraints. G18 was not shielded for any ensemble in agreement with NMR data. Absolute values of the shielding are not comparable to the experiment due to crude approximations in the NICS model (see text for discussion). Values are in ppm.

H1	**$\Delta\delta_{exp}$**	**$\Delta\delta_{calc,rst}$**	**$\Delta\delta_{calc,unrst}$**
G2	-0.05	-1.08 ± 0.36	-0.46 ± 0.52
G6	-0.88	-2.48 ± 0.31	-2.61 ± 0.28
G14	-1.04	-0.54 ±0.47	-1.45 ± 0.78
G18	-0.11	0.06 ± 0.08	-0.03 ± 0.14