Supplemental Information

Structure of Two G-quadruplexes in equilibrium in the KRAS promoter.

Julien MARQUEVIELLE[1], Coralie ROBERT[1], Olivier LAGRABETTE[1], Mona WAHID[1], Anne BOURDONCLE[1], Luigi XODO[2], Jean-Louis MERGNY[1], Gilmar F. SALGADO[1]

1. European Institute of Chemistry and Biology (IECB), INSERM U1212 - CNRS UMR 5320, University of Bordeaux, France.

2. Department of Medical and Biological Sciences, School of Medicine, 33100 Udine, Italy.

NMR structure calculation

NOE distance Restraints

Distance between the protons of KRAS32R G9T and G25T were obtained from NOESY experiments at 37°C with a different mixing time (250, 500 ms). In the final structure calculations, only data from the 250 ms mixing time was used. Peaks assignment was performed with Sparky and peaks volumes were extracted from the spectra with manual correction depending on the reliability of the integration. NOE distance restraints were then calculated using CCPN Analysis software using chemical shifts and peaks list from Sparky with Thymine H6-H7 distance as a reference.

Hydrogen Bonds

Hoogsteen hydrogen bonds between guanines were restrained using H21-N7 distance (2.0 ± 0.3) and N1-O6 distance (2.0 ± 0.3). Hydrogen bond between guanine and adenine in KRAS32R G9T triad used H21-N3 distance (2.0 ± 0.3)

Planarity restraints

Planarity restraints were used for G2→G6→G11→G25, G3→G7→G12→G26, G4→G32→G13→G27 tetrads and for G28→A30→G31 triad in the case of KRAS32R G9T conformer.

Planarity restraints were used for G2→G6→G11→G26, G3→G7→G12→G27, G4→G9→G13→G28 tetrads in the case of KRAS32R G25T conformer.

ARIA2/CNS structure calculation

All restraints files were included in ARIA2.3/CNS1.2 for structure calculation. Eight iterations of calculations were performed with 100 to 750 structures per iteration with mixed Cartesian and torsion angle dynamics during the simulated annealing runs. The protocol contains several stages. First, an initial high-temperature torsion angle simulated annealing of 10,000 steps at 10,000 K followed a torsion angle dynamic cooling stage of 10,000 steps from 10,000 to 2000 K. Then, a Cartesian dynamics cooling stage of 5,000 steps from 2000 to 1000 K, and finally a Cartesian dynamics cooling stage of 4,000 steps from 1000 to 50 K with 3 fs per step. For distances and hydrogen bonds, 10 kcal $mol^{-1}Å^{-2}$ was applied during the initial stage of dynamics and was increased up to 50 kcal $mol^{-1}Å^{-2}$ for the remaining steps of the dynamics. For dihedral restraints, different energy constants were applied depending in the calculation phase: 5 kcal $mol^{-1}Å^{-2}$ for the high temperature steps and then 25 and 200 kcal $mol^{-1}Å^{-2}$ respectively for the first and second cooling step. An energy constant of 25 kcal

mol^{-1}Å^{-2} was applied for planarity restraints. In the final run, the best twenty structures were extracted and analysed to choose structures for AMBER refinement.

Structure refinement in Explicit Solvent (AMBER12)

Before neutralizing the system by using K^+ cations, two K^+ cations were included in the G-tetrad core of the two best structures obtained from ARIA calculation. Then, the system was solvated with water molecules (TIP3P) in a truncated octahedral box. Distance restrains and planarity restrains were imposed during molecular dynamics refinement. The system was first minimized with harmonic potential position restraints at 25 kcal mol^{-1}Å^{-2} over 1000 steps of steepest descent minimization. Then it was heated from 100 K to 310 K during 20 ps with constant volume and the same potential position restraints as before. Then a several equilibrations were performed using gradually reduced positional restraints: 22, 20, 17, 15 and 12 kcal mol^{-1}Å^{-2}. Finally, the system was equilibrated without positional restraints for 5 ns at 310K and the ten lowest-energy structures were chosen.

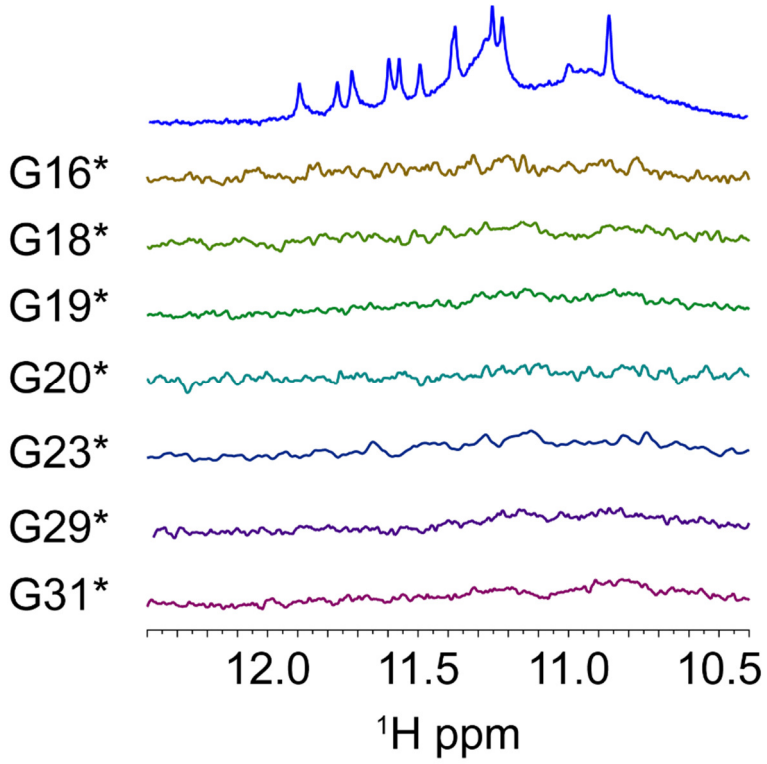


Figure S1. 15N-filtered NMR spectra at 37°C of samples containing 5% of 15N-enriched isotope of residues not implicated in KRAS32R G-quadruplex. All experiments were performed at 37°C in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

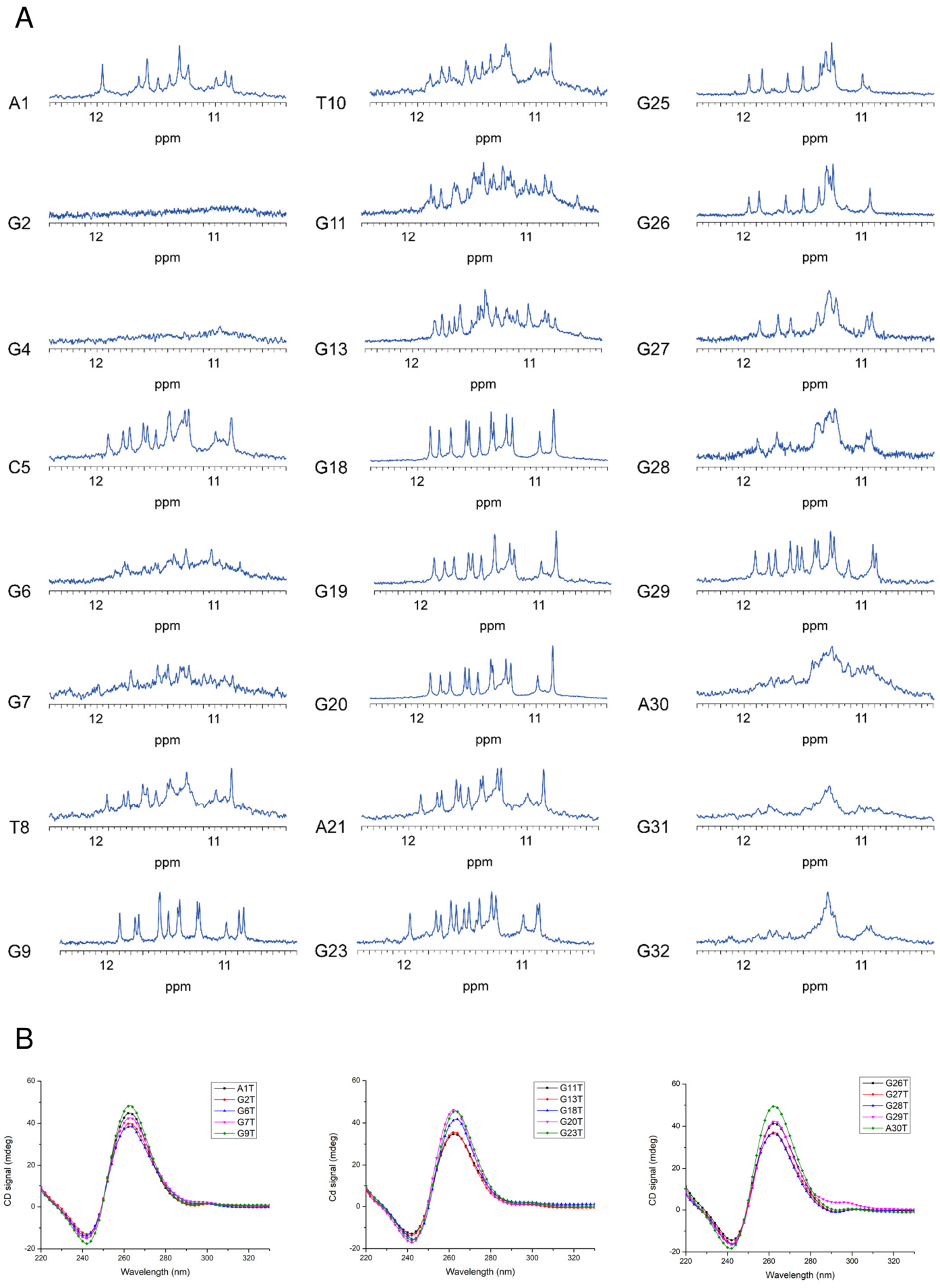
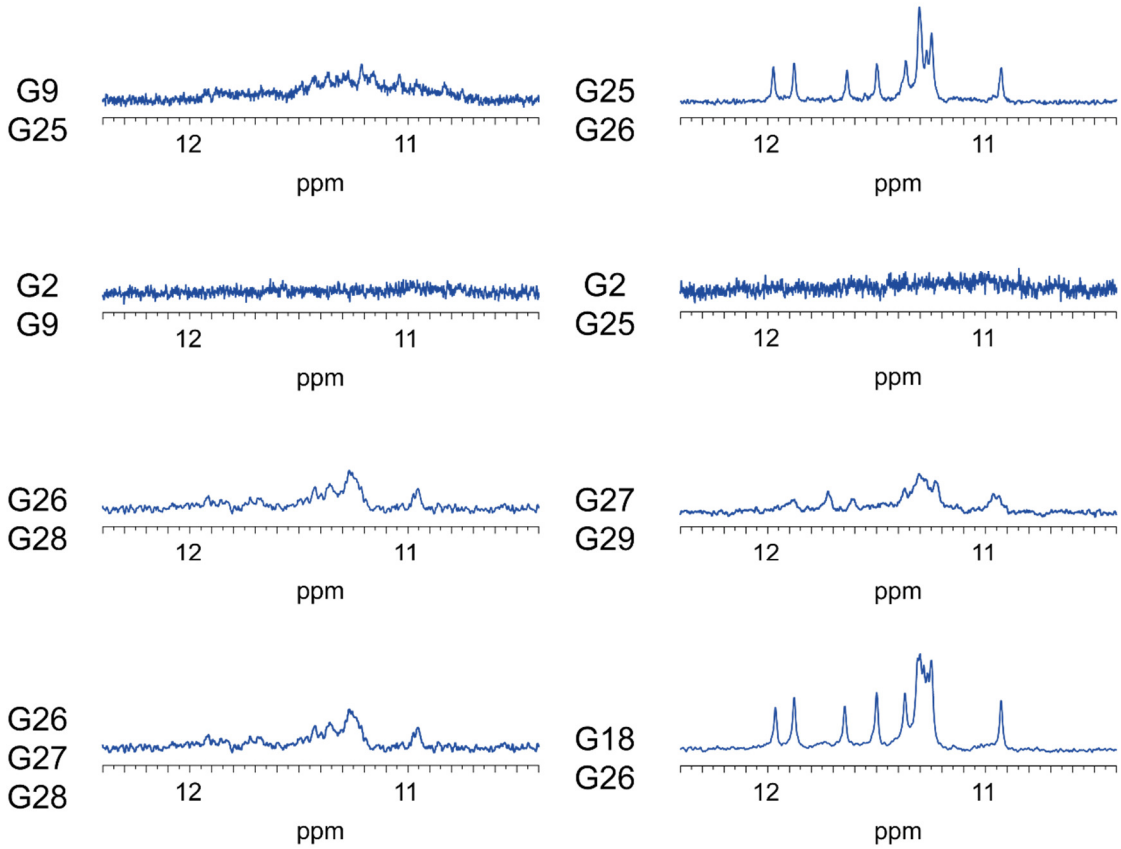


Figure S2. A) All NMR spectra imino regions of KRAS32R simple mutants at 37°C used in this study to better understand the KRAS32R G4 formation and the implication of the different residues with in **B**) the corresponding CD spectra at 37°C. All experiments were performed in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

A



B

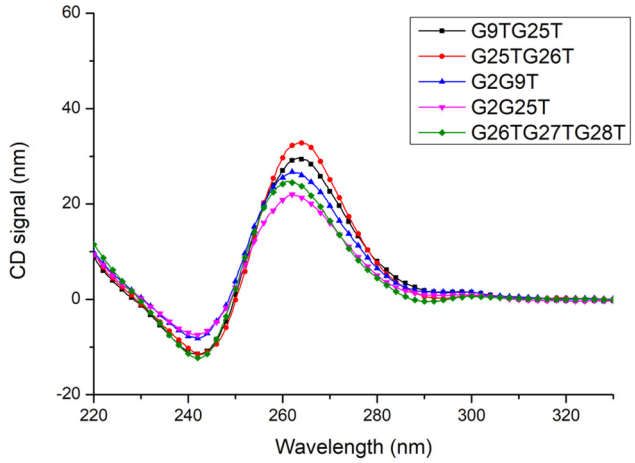


Figure S3. A) ^{1}H spectra imino regions of double mutants performed within KRAS32R WT at 37°C to identify exchanges among bases implicated in tetrads with in **B**) the corresponding CD spectra. These results confirm the models for G9T and G25T and give more information on the role of G2 and the guanines in the last tract. All experiments were performed at 37°C in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

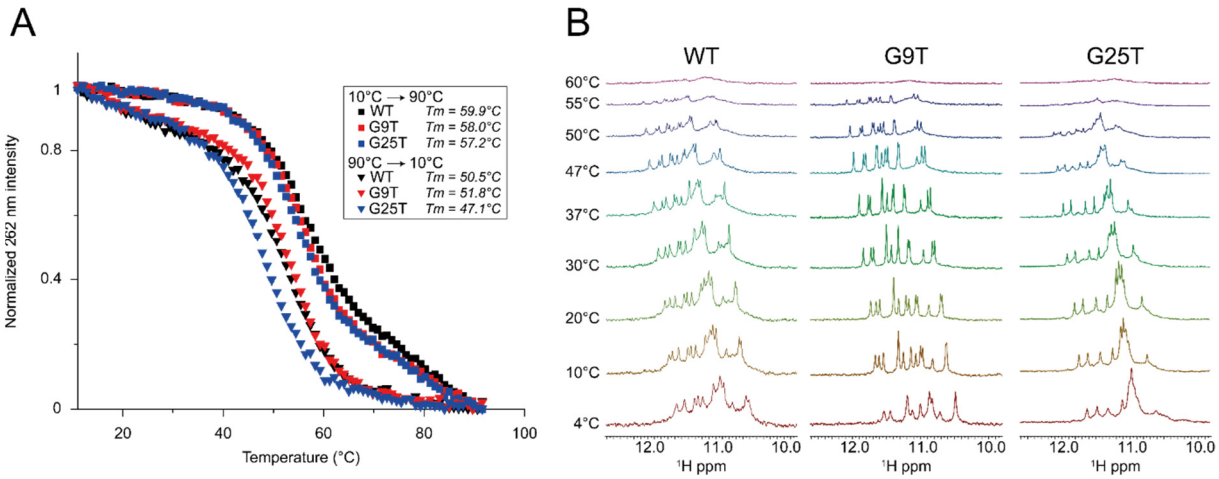


Figure S4. In **A**, CD spectra of wild type (black), G9T (red) and G25T (blue) from 10°C to 90°C (square) show very close melting temperatures but from 90°C to 10°C stabilization is the same compared to NMR observed in **B**) NMR melting experiment of wild type, G9T and G25T conformers in A from 4°C to 60°C showing that G9T conformer seems to be more stable than wild type which is itself more stable than G25T, especially when looking at the 1D NMR spectra obtained at 55°C. All experiments were performed in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

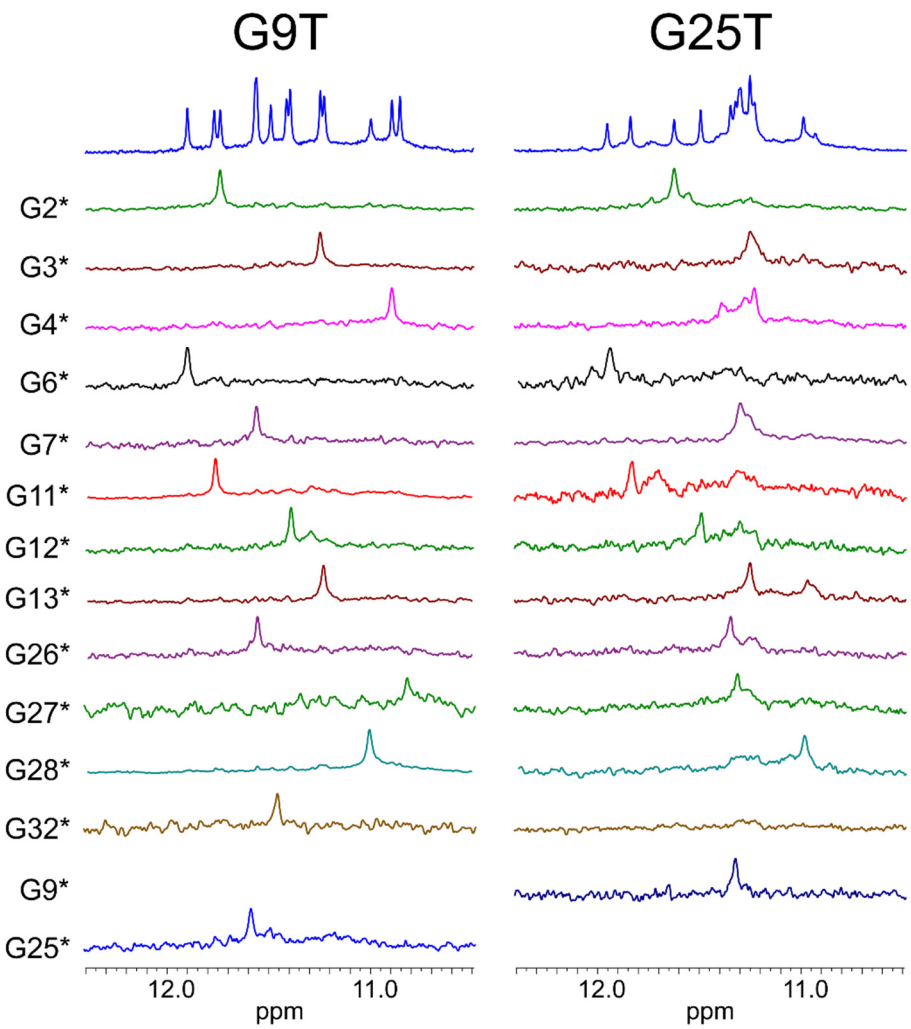


Figure S5. Determination of the guanines implicated in KRAS32R G9T and G25T tetrads using 15N-filtered NMR spectra at 37°C of samples containing ≈5% of 15N-enriched isotope. G9T imino protons give only one peak meaning that each of them is implicated in only one conformation whereas in G25T, several imino protons such as G4 or G7 are implicated in more than one conformation. All experiments were performed at 37°C in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

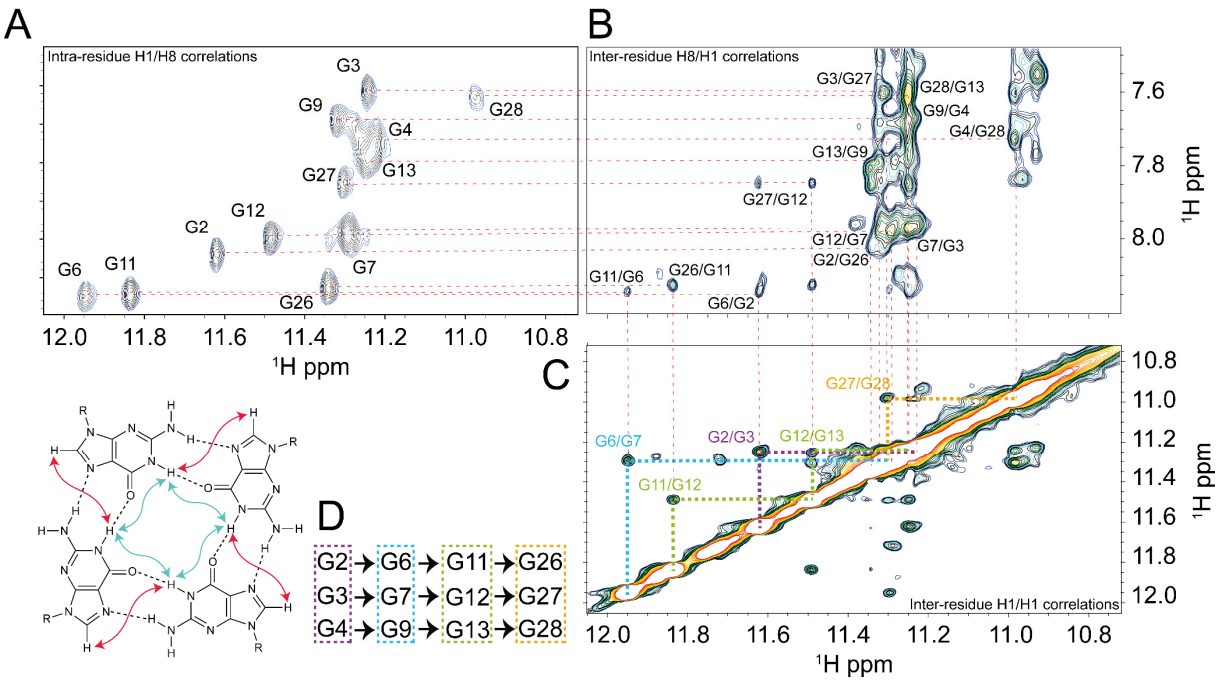


Figure S6. **A**) NMR HCCNH-TOCSY experiment of KRAS32R G25T showing intra-residue H1/H8 correlations in order to identify H8 of guanines implicated in G9T G-quadruplex and NOESY inter-residue H8/H1 correlations in **B**) and in **C**) NOESY inter-residue H1/H1 correlations were used to determine tetrads pattern in **D**).

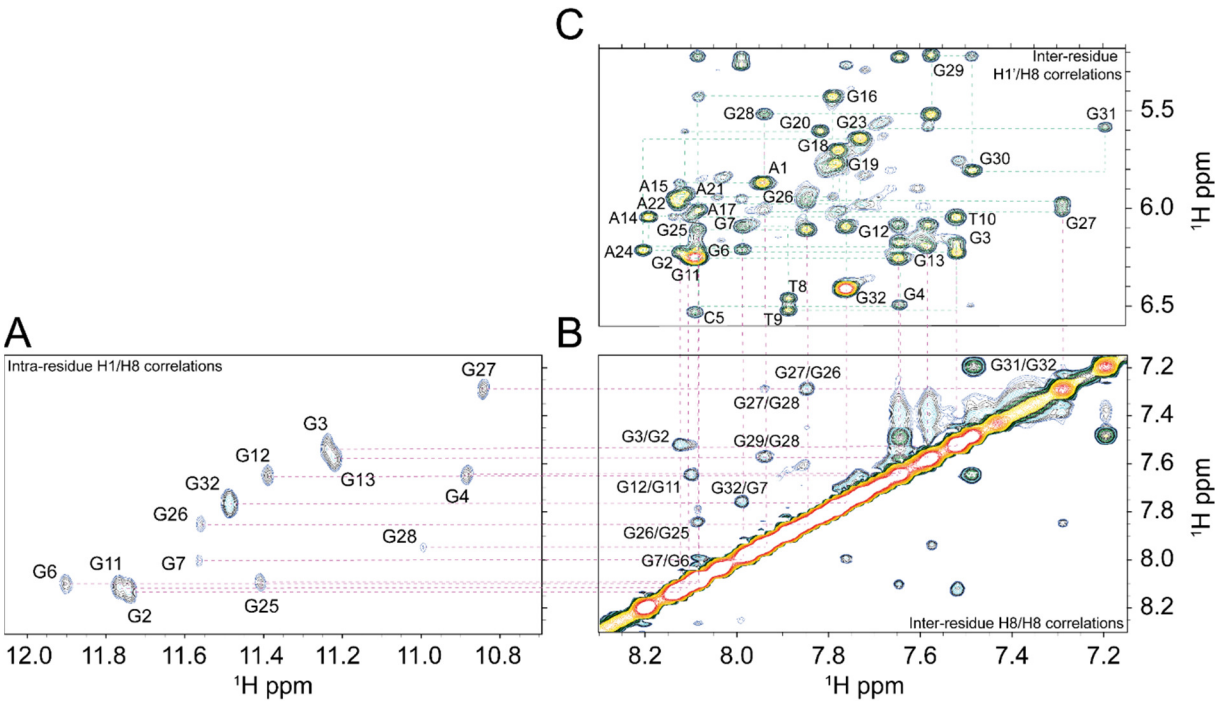


Figure S7. **A**) NMR HCCNH-TOCSY experiment of KRAS32R G9T showing intra-residue H1/H8 correlations used to identify H8 from guanines implicated in the G4. Then in **B**) Inter-residue H8/H8 correlations were identified between these guanines. All these results together help to determine KRAS32R G9T "walk" identifying inter-residue correlations H1'/H8 in **C**) between H1' from a residue and H8 from the next residue. Some of these correlations are missing, especially for residues implicated in triad formation due to its particular geometry and cross-peaks are also missing in bulges. All experiments were performed at 37°C in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

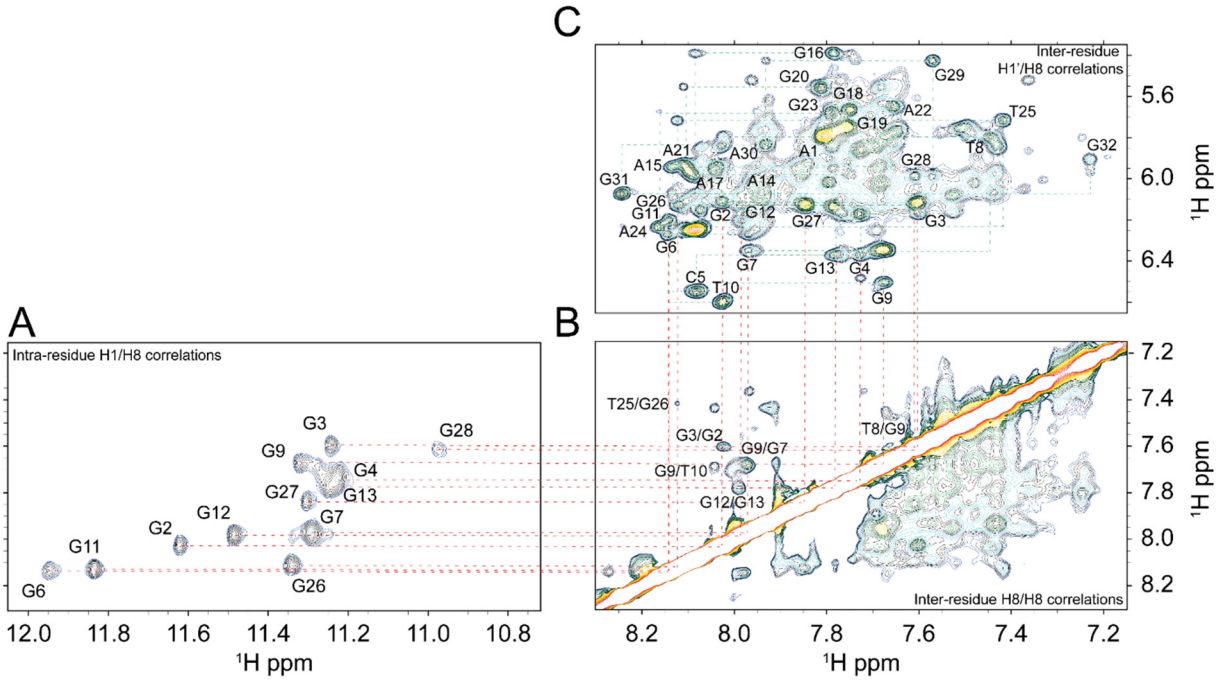


Figure S8. **A**) NMR HCCNH-TOCSY experiment of KRAS32R G25T showing intra-residue H1/H8 correlations used to identify H8 from guanines implicated in the G4. Then in **B**) Inter-residue H8/H8 correlations were identified between these guanines. All these results together help to determine KRAS32R G25T "walk" identifying inter-residue correlations H1'/H8 in **C**) between H1' from a residue and H8 from the next residue. Cross-peaks in bulges are mostly missing. All experiments were performed at 37°C in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

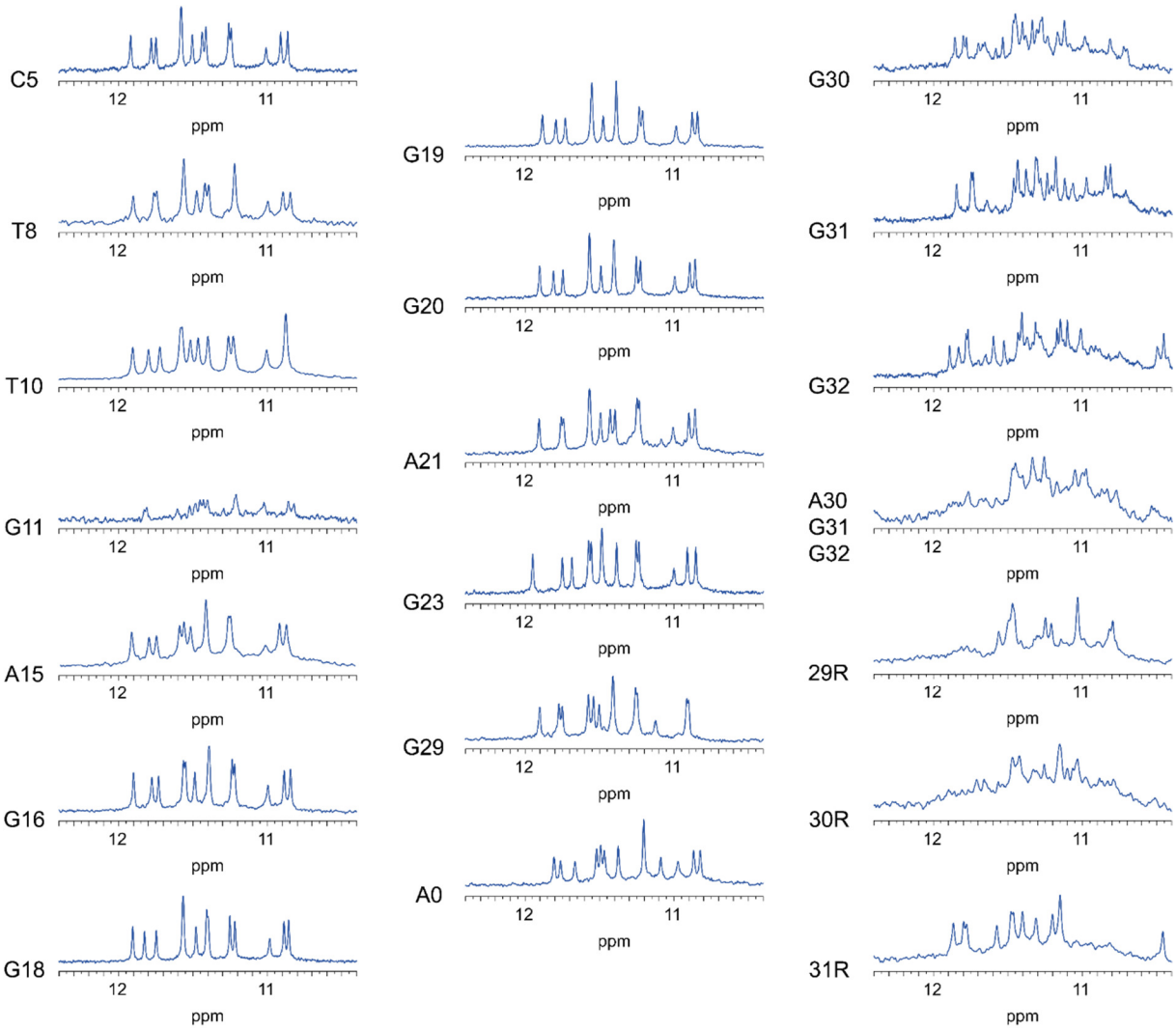


Figure S9. All NMR spectra imino regions of KRAS32R G9T mutants at 37°C used in this study. A0 corresponds to the addition of an adenine before A1 in the sequence. 29R, 30R and 31R correspond to the corresponding length of sequence. These results confirm the implication of G28, A30 and G31 in the structure and the participation of G32 in a tetrad. No residue from the loop has been identified to participate in any structural element. All experiments were performed in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

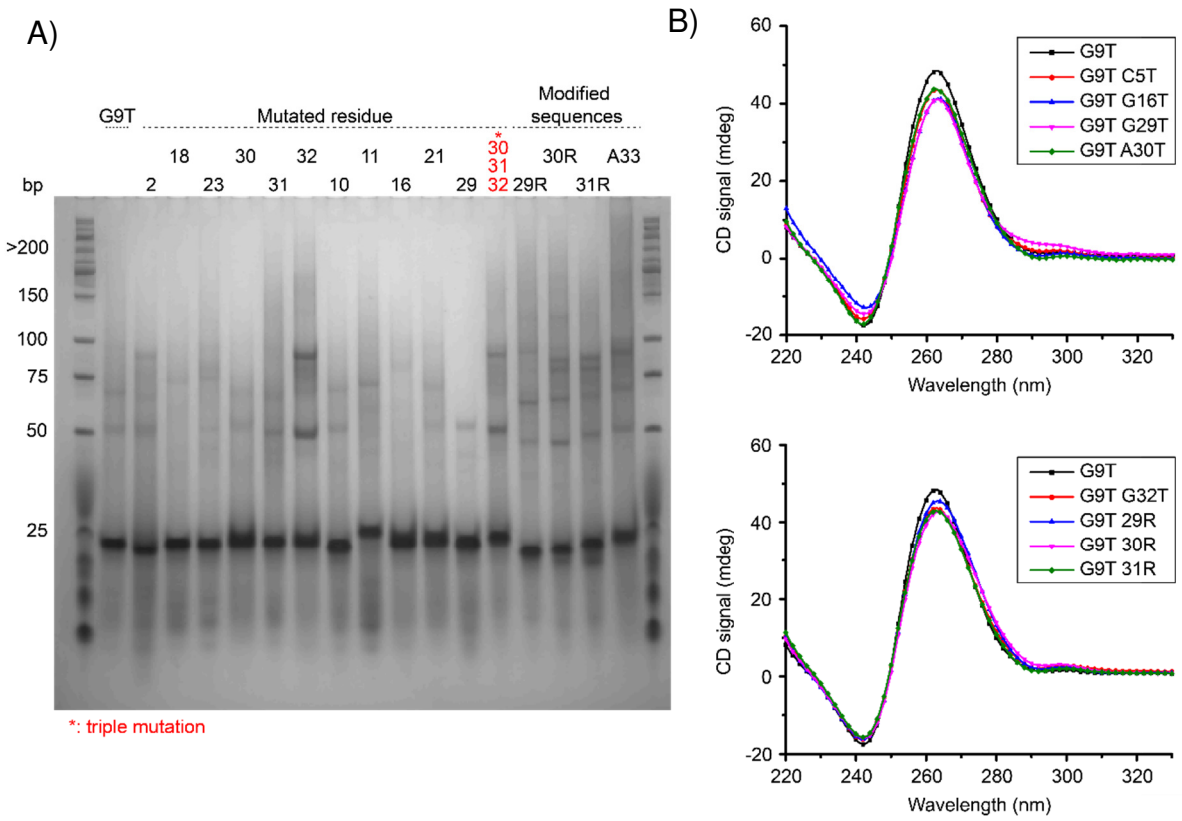


Figure S10. **A**) Native gel experiment of KRAS32R G9T and several mutants with also modified sequences. Most of the mutations do not affect migration but several of them especially mutations of the last four residues lead to higher weigh complexes. **B**) KRAS32R G9T mutants' CD spectra at 37°C show that most of the mutations do not affect G4 conformation showing all parallel conformation with characteristic peaks around 260 nm (positive) and 240 nm (negative). All experiments were performed in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

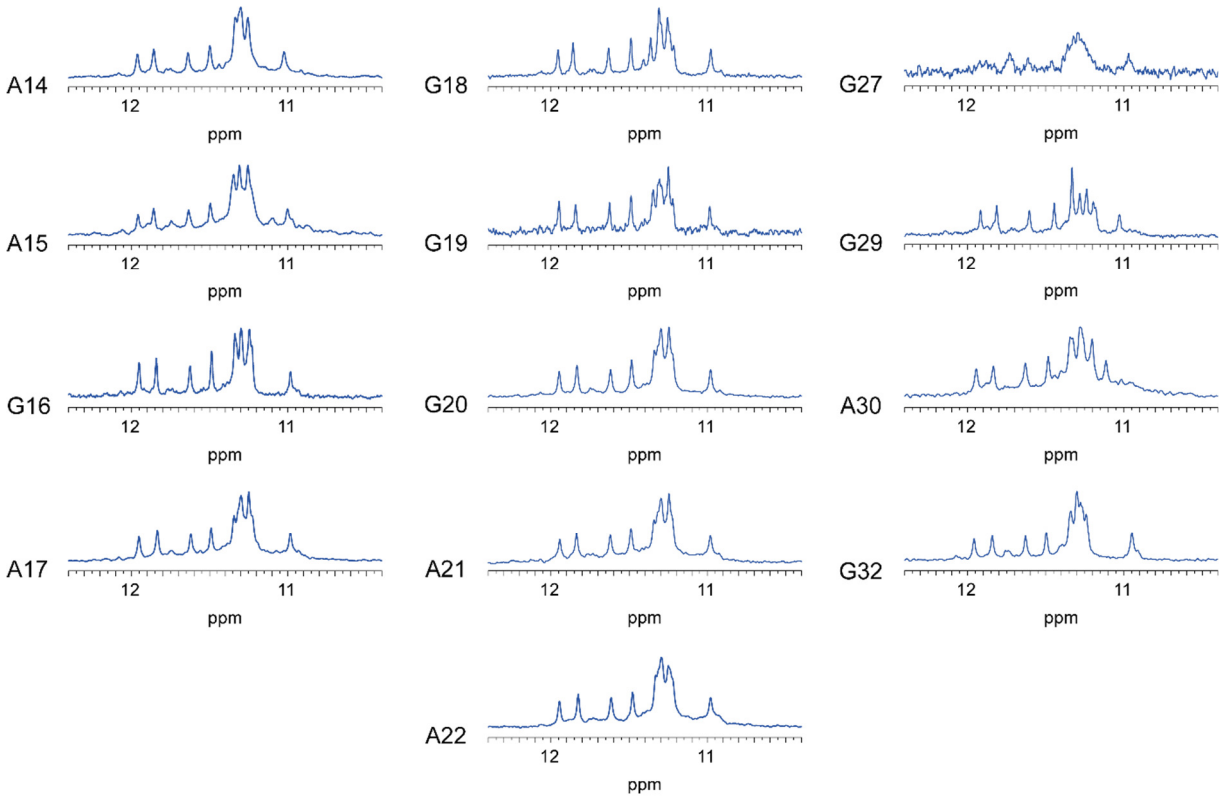


Figure S11. All NMR spectra imino regions of KRAS32R G25T mutants at 37°C used in this study. No residue from the loop neither the last four residues have been identified to participate in any structural element. All experiments were performed in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

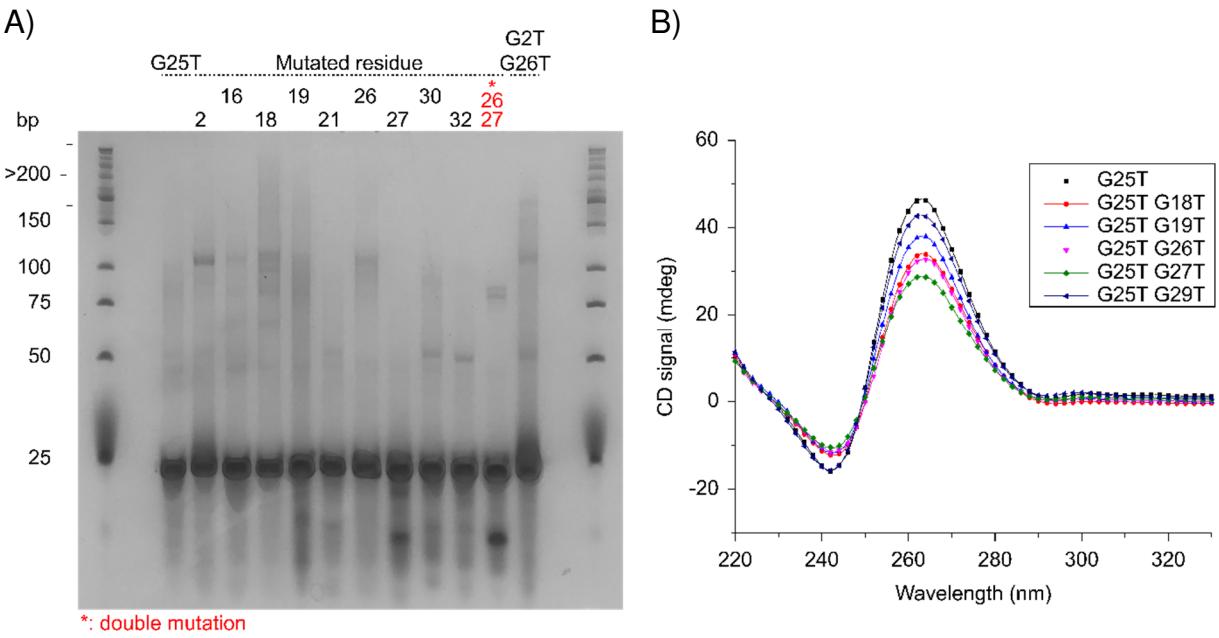


Figure S12. **A**) Native gel experiment of KRAS32R G25T and several. Mutations do not affect migration except mutations implicating G27 which is known to be part of G4 tetrad. **B**) KRAS32R G25T mutants' CD spectra at 37°C show that most of the mutations do not affect G4 conformation showing all parallel conformation with characteristic peaks around 260 nm (positive) and 240 nm (negative). All experiments were performed at 37°C in buffer 1X (50 mM KCl; 10 mM KPi; pH 6.66).

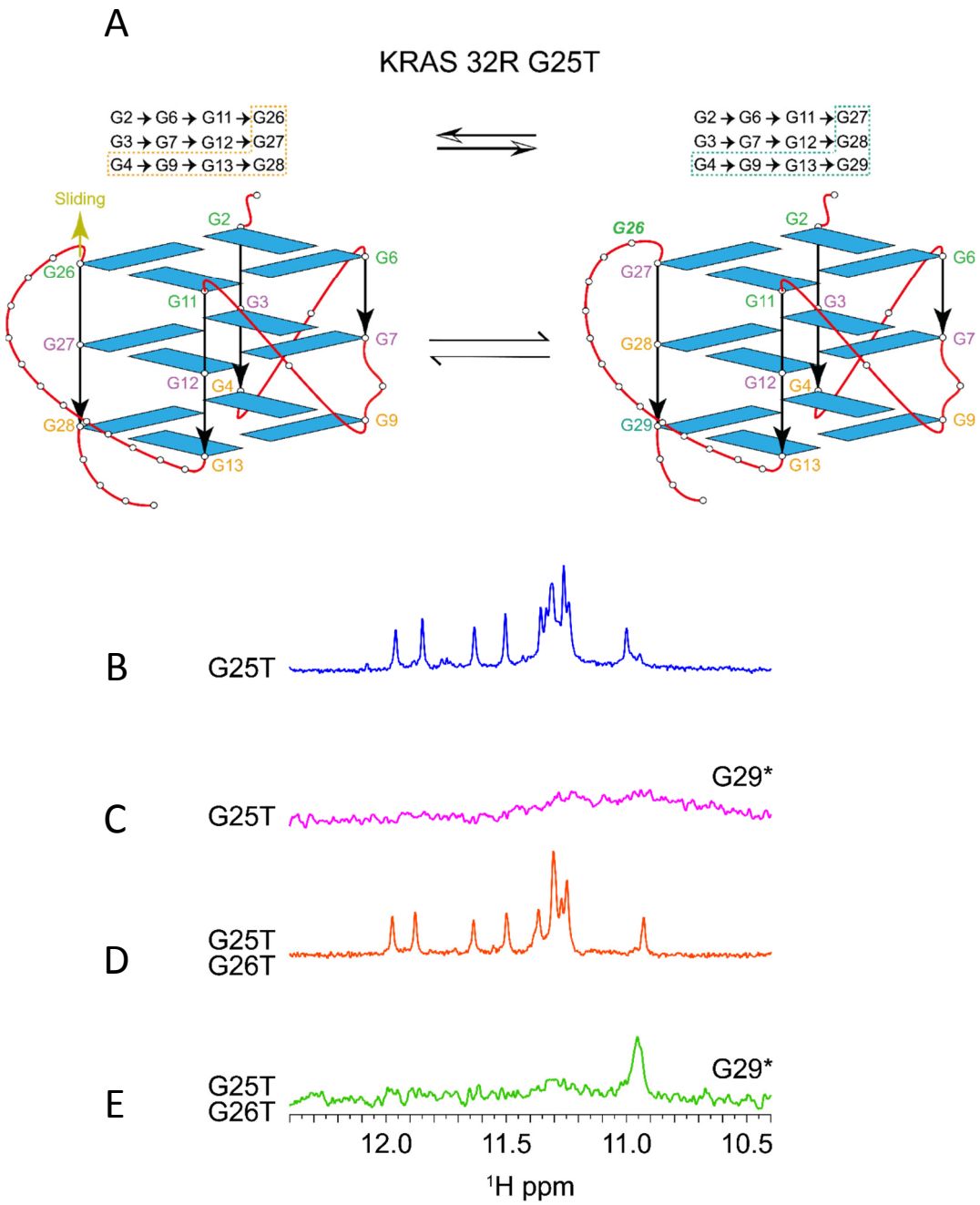


Figure S13 (**A**) Proposed model for G25T sliding into the tetrad when both G25 and G26 are simultaneously mutated into T. 1D NMR spectra comparing KRAS32R G25T simple mutant (**B**) with G25TG26T (**D**) double mutant imino region at 37°C. Results show that G29 is implicated in G-quadruplex formed within double mutant (**E**) but not in KRAS32R G25T (**C**).

NMR Restrains (H_2O)	
distance restrains	873
intraresidue distance restrains	503
sequencial (i, i+1) distance restrains	199
long-range (i, i+2) distance restrains	116
short-range non sequencial distance restrains	55
dihedral restrains	15
H bonds restrains	28
Structural statistics	
structure calculation	
total calculated structures	750
NOE violations	
number (>0.3 Å)	0.13
RMSD of violations (Å)	0.26 ± 0.03
molecular dynamics	
simulation time (ns)	5.0
extracted structures (lowest energy)	10
RMSD	
All heavy atoms (Å)	1.5

Table S1. Table containing NMR restrains for KRAS32R G9T G-quadruplex structure calculation with structural statistics after structure calculation and refinement by molecular dynamics

NMR Restrains (H_2O)	
distance restrains	564
intraresidue distance restrains	381
sequencial (i, i+1) distance restrains	109
long-range (i, i+2) distance restrains	37
short-range non sequencial distance restrains	37
dihedral restrains	12
H bonds restrains	24
Structural statistics	
structure calculation	
total calculated structures	750
NOE violations	
number (>0.3 Å)	0.09
RMSD of violations (Å)	0.33 ± 0.03
molecular dynamics	
simulation time (ns)	5.0
extracted structures (lowest energy)	10
RMSD	
All heavy atoms (Å)	1.9

Table S2. Table containing NMR restrains for KRAS32R G25T G-quadruplex structure calculation with structural statistics after structure calculation and refinement by molecular dynamics.

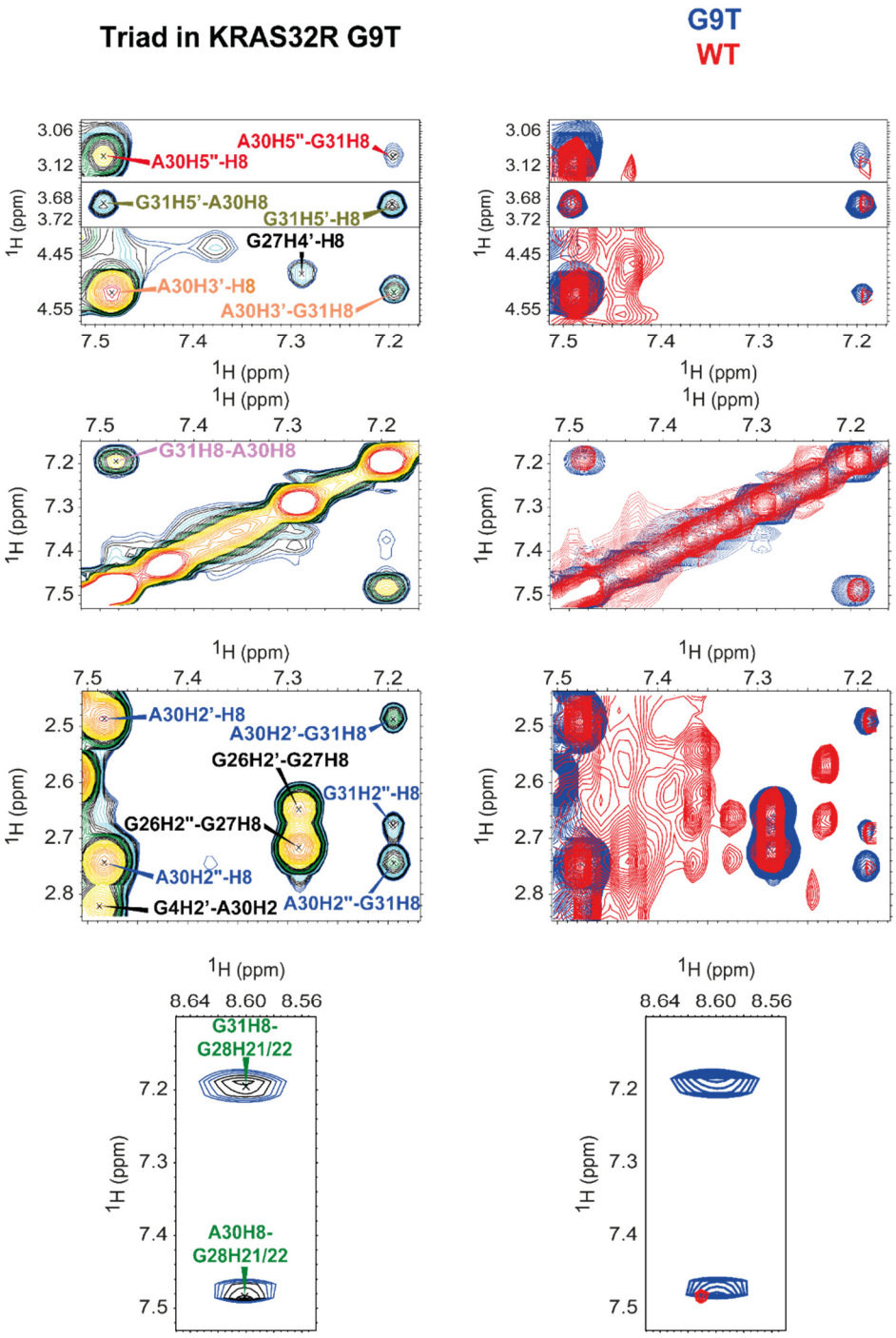


Figure S14. Cross-peaks indicating the presence of a triad in Kras32RWT sample. Peaks proving the existence of the triad at 37°C in KRAS32R G9T {^{1}H-^{1}H} NOESY spectrum are also found in KRAS32R WTT ^{1}H-^{1}H NOESY spectrum.

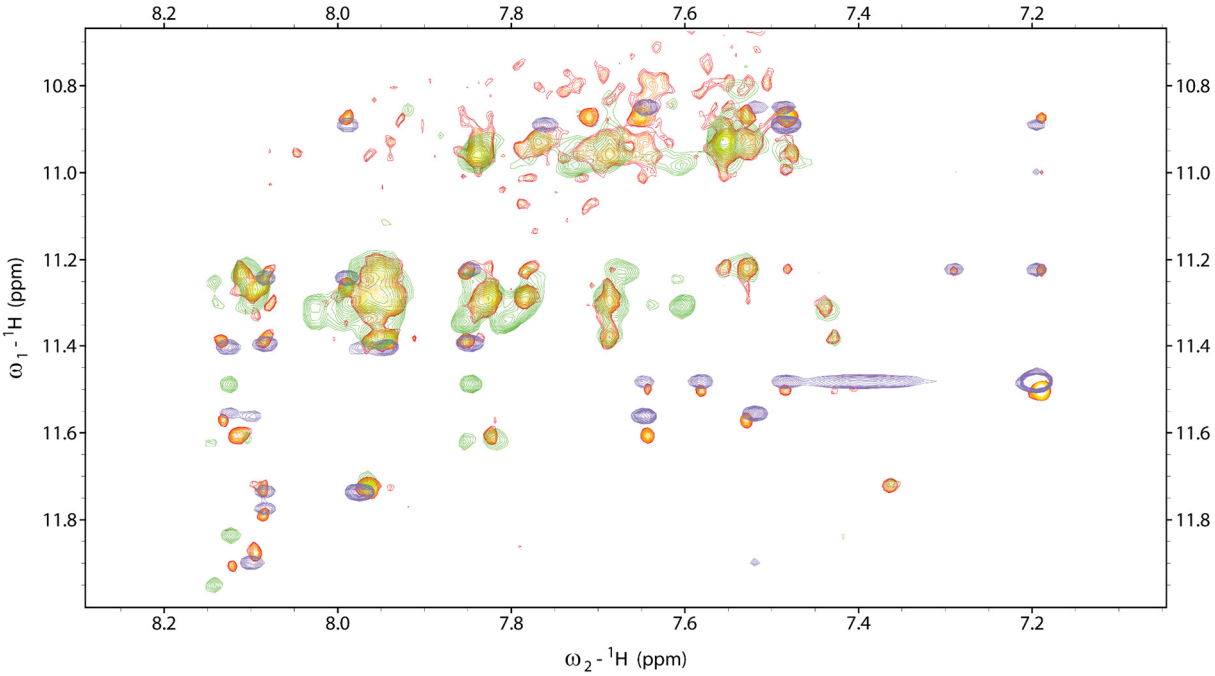


Figure S15. Comparison of {^{1}H-^{1}H} NOESY spectra of KRAS32R WT, G9T and G25T in red-orange, purple and green respectively. The image depicts the imino-aromatic cross-peaks region. With a few exceptions, the strong correlation between overlapped peaks both conformers G9T and G25T with WT 32R strongly supports the existence of both conformers in the WT sample.

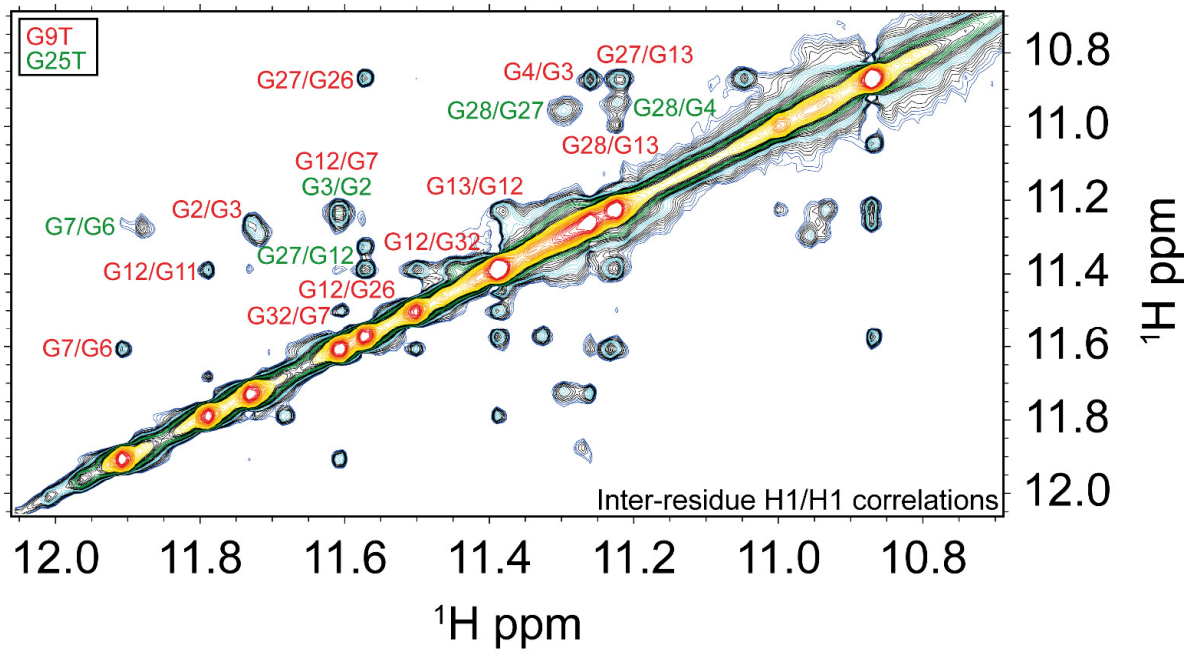


Figure S16. 2D {^{1}H-^{1}H} NOESY spectra depicting the imino region of KRAS32R WT, with some unambiguous assignments of G9T and G25T conformers obtained with from 1D ^{15}N-filtered HMQC experiments as shown in figure 1D and S5.

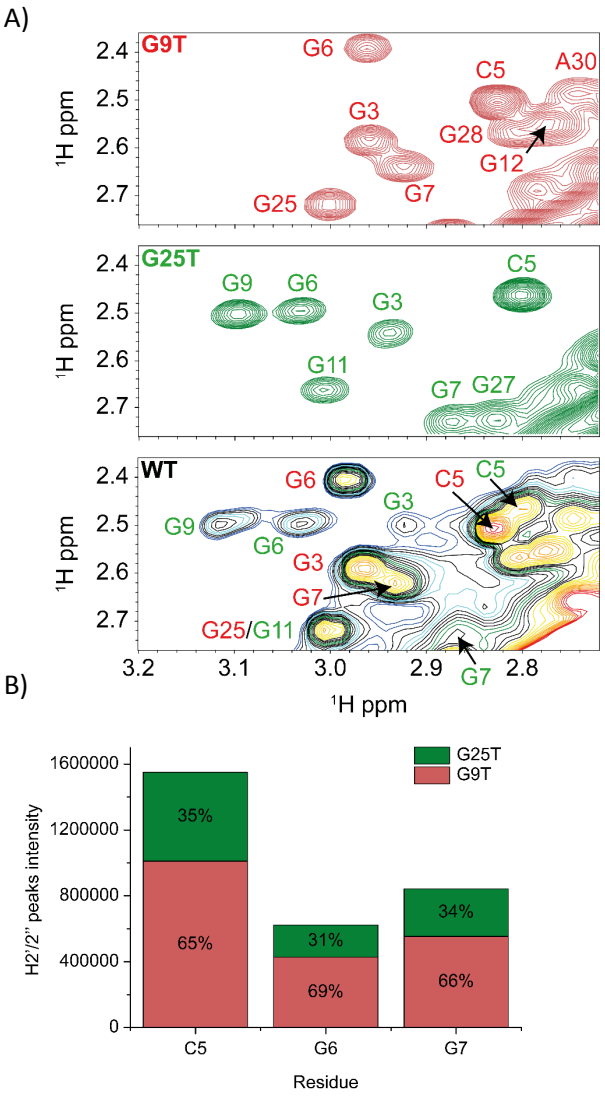


Figure S17. A) Inspection of the H2'-H2" cross peaks from isolated bases for G9T top (red), G25T middle (green) and 32RWT bottom, from each respective 2D {^{1}H-^{1}H} Noesy spectra. The H2'-H2" cross peaks from bases C5, G6 and G7 were used to estimate the distribution of conformers in 32RWT (B) with an average population of ≈66 and 33 % for G9T and G25T respectively of folded sample. For the final values in the NMR tube we need to take into account 5-10% of unfolded sample.