2’-F-ANA-guanosine and 2’-F-guanosine as powerful tools for structural manipulation of G-quadruplexes

Christopher Jacques Lech, Zhe Li, Brahim Heddi, and Anh Tuân Phan*

School of Physical and Mathematical Sciences,

Nanyang Technological University, Singapore 637371

SUPPORTING INFORMATION

Contents

MATERIALS AND METHODS

DNA Sample Preparation

DNA oligonucleotides were chemically synthesized using a DNA synthesizer (model 394, Applied Biosystems, Foster City, CA). Oligonucleotides were then purified using a Poly-Pak cartridge (Glen Research) following standard protocol. Purified samples were dialyzed against successively against H_2O, ~25 mM KCl, and H_2O. Samples were then lyophilized and rehydrated in conditions described below. The 2'-deoxy-2'-fluoroarabino-guanosine phosphoramidites were purchased from Glen Research (Sterling, VA).

UV Spectroscopy

Thermal Stability measurements were made via UV melting experiments. Using a Varian 300 Cary Bio UV-visible (UV-Vis) spectrophotometer, the absorbance at 295 nm and 320 nm was monitored for DNA samples of 10 μM strand concentration, 20 mM potassium phosphate (pH 7) and 70 mM KCl. Absorbance at 295 nm was corrected by subtracting the absorbance at 320 nm to partially correct for noise. Samples were heated to 83°C (92°C for all syn samples) and then cooled to 19°C (37°C for all syn samples) at a rate of 0.2°C/min. Samples were allowed to equilibrate before being heated over the same range with the same heating rate. Hysteresis of less than 1°C was observed for all samples. Melting Temperature (T_m) is determined by the temperature in which half of the sample population is folded and half is unfolded. Tm values presented represent an average over the heating and cooling curves. UV Samples were obtained from NMR samples after recording NMR spectra.

CD Spectroscopy

CD experiments were performed using a JASCO-815 spectropolarimeter (JASCO, Tokyo, Japan). Spectra represent an average of 10 accumulations for samples containing 5 μM strand concentration, 20 mM potassium phosphate (pH 7) and 70 mM KCl. CD absorbance was recorded in the range of 220 nm to 320 nm at a temperature of 25°C. Spectra were baseline corrected by subtracting the absorbance at 320 nm, and then normalized by strand concentration using concurrent UV absorption measurements. CD Samples were obtained from NMR samples after recording of NMR spectra.

NMR Spectroscopy

^1H NMR spectra were recorded using a 600 MHz Bruker (Billerica, MA) NMR spectrometer. Spectra were taken at using a jump-and-return based water suppression pulse sequence at a temperature of 25°C. Samples had a DNA strand concentration in the range of 100-400 μM dissolved in 20 mM potassium phosphate (pH 7) and 70 mM KCl. Chemical shifts were referenced using 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). Prior to recording NMR spectra, samples were annealed by heating to 100°C for 5 min and allowed to cool to room temperature over several hours.

Supporting Text 1:

In our work, we employ a combination of NMR and CD spectroscopy as complimentary tools to determine the number of conformations present in solution as well as the folding topology of these sequences. NMR spectroscopy allows us to monitor the conformations present in solution. The presence of signal in the 10-13 ppm range is characteristic of imino proton resonance from Hoogsteen-bound guanines like those within guanine-tetrads (or G-tetrads) of G-quadruplex DNA. For a G-quadruplex with three G-tetrad layers, one would expect the presence of 12 distinct peaks in this region, one for each guanine base (e.g. see Fig 2A bottom). The presence of minor peaks in this region (e.g. see Fig 2A top) indicates the presence of other folding topologies present in solution. This equilibrium can range from a small population of competing structures to many equivalently occupied folding topologies.

The reduction of minor peaks in the spectra of ^{FANA}G modified sequences (Fig 2A-B) is a clear indication that these modifications are effective tools in decreasing the polymorphism of these G-quadruplex sequences. We use CD spectroscopy to determine whether or not the folding topology of the G-quadruplex formed in Fig 2A-B is the same as the major conformer in the unmodified sequence. NMR spectra of unmodified and modified sequences are strikingly similar for the "Decreasing Polymorphism" experiments, suggesting the same fold is adopted. As a rigorous determination of folding topology by NMR would require extensive NMR assignments,[1] we make use of CD spectroscopy as a convenient tool to investigate the folding topology of modified sequences.

In the "Decreasing Polymorphism" experiments, samples HT1 and HT2 are in a mixture of conformers as shown by NMR spectroscopy with the major component being a (3+1) hybrid folding topology (Fig 2A-B). CD spectra of unmodified HT1 and HT2 samples are characteristic of a major conformation being a (3+1) hybrid folding topology with a positive peak at 290 nm, a shoulder at 265 nm, and a negative peak at roughly 235 nm.[2] CD spectra of samples containing a single ^{FANA}G substitution are very similar to the unmodified spectra, with a small increase observed in the 265 nm shoulder in the CD spectra of the HT2-16 FANA sample. This increase is likely to be caused by a different spectral contribution from the residual minor form in this sample.

CD spectroscopy is a useful tool for studying chiral molecules such as nucleic acids. The CD signature of G-quadruplex DNA is found to be unique for different folding topologies.[3] Spectral contributions from multiples species in solution can lead to complicated or even un-interpretable CD spectra. Additionally, the presence of multiple species of similar folding is often unrealized, particularly when the total CD spectra is characteristic of a familiar folding. Therefore, monitoring polymorphism using NMR is crucial for correct interpretation of CD spectra.

Supporting Text 2:

In the "Conformational Switching" experiments, we again use NMR and CD spectroscopy to monitor the polymorphism and conformation of modified sequences. The NMR spectra of sequences containing ^{F}G and ^{FANA}G modifications to *syn* positions (Fig 3A-B and S5) show 12 distinct imino proton peaks with little to no minor conformations present. Additionally, the imino proton patterns of *syn*-modified sequences are much different from the imino proton patterns of unmodified HT1 and HT2 sequences suggesting a different conformation is adopted.

CD spectra of *syn*-modified sequences reveal drastic changes compared to the unmodified sequences (Fig 3C-D). Unmodified sequences, as discussed in Supporting Text 1, have CD spectra that are characteristic of a (3+1) hybrid folding topology. After undergoing modifications to *syn*-positions, CD spectra are observed to have no peak at 295 nm, but rather a positive at 265 nm and a negative peak at 245 nm. This CD spectra is characteristic of a parallel folding topology.[4] Similar CD profiles have been observed in previous studies and rigorously determined to correspond to the conformational switching of (3+1) hybrid to parallel folding topologies.[4, 5] The combination of NMR spectroscopy showing the formation of a single conformation upon sequence modification coupled with the CD spectral signature of a parallel G-quadruplex folding topology leads us to conclude that a single parallel folding topology is obtained after modification.

The intensity of the CD spectra observed for the parallel folding topology is observed to be much higher than the (3+1) hybrid spectra (Fig 3C-D). While the reasons for this are complex, this intensity difference has been previously observed between (3+1) hybrid and parallel folding topologies.[4, 5]

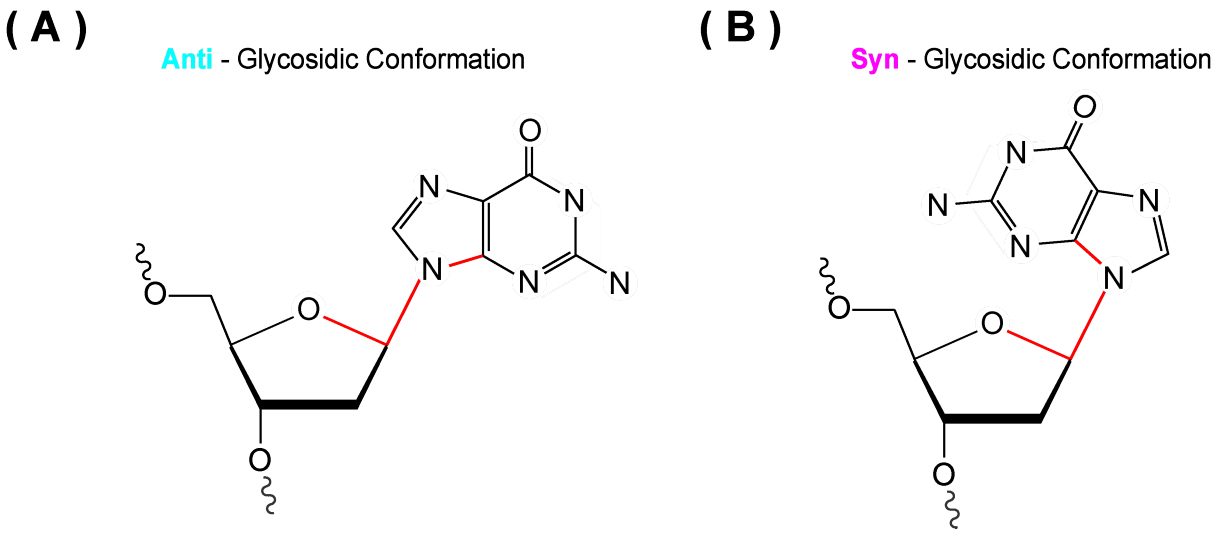

Figure S1: Schematic representation of guanine glycosidic conformation: **(A)** *anti* and **(B)** *syn* conformations are shown with the dihedral about the glycosidic bond highlighted in red. An *anti* base conformation is represented as cyan in Fig 1.C-D while *syn* base conformation is represented as magenta.

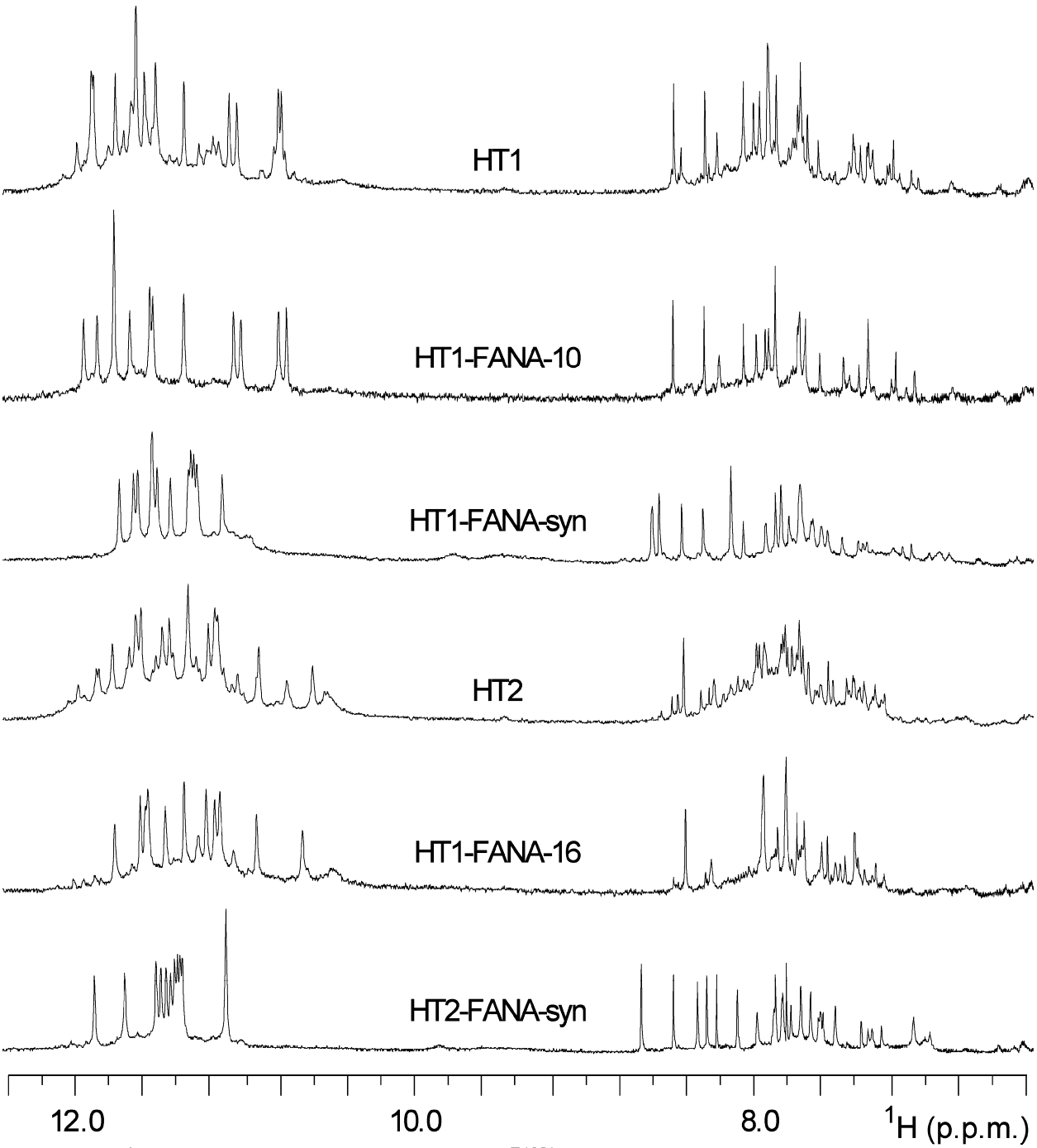

Figure S2: ^{1}H NMR spectra of unmodified and ^{FANA}G modified sequences explored in this study

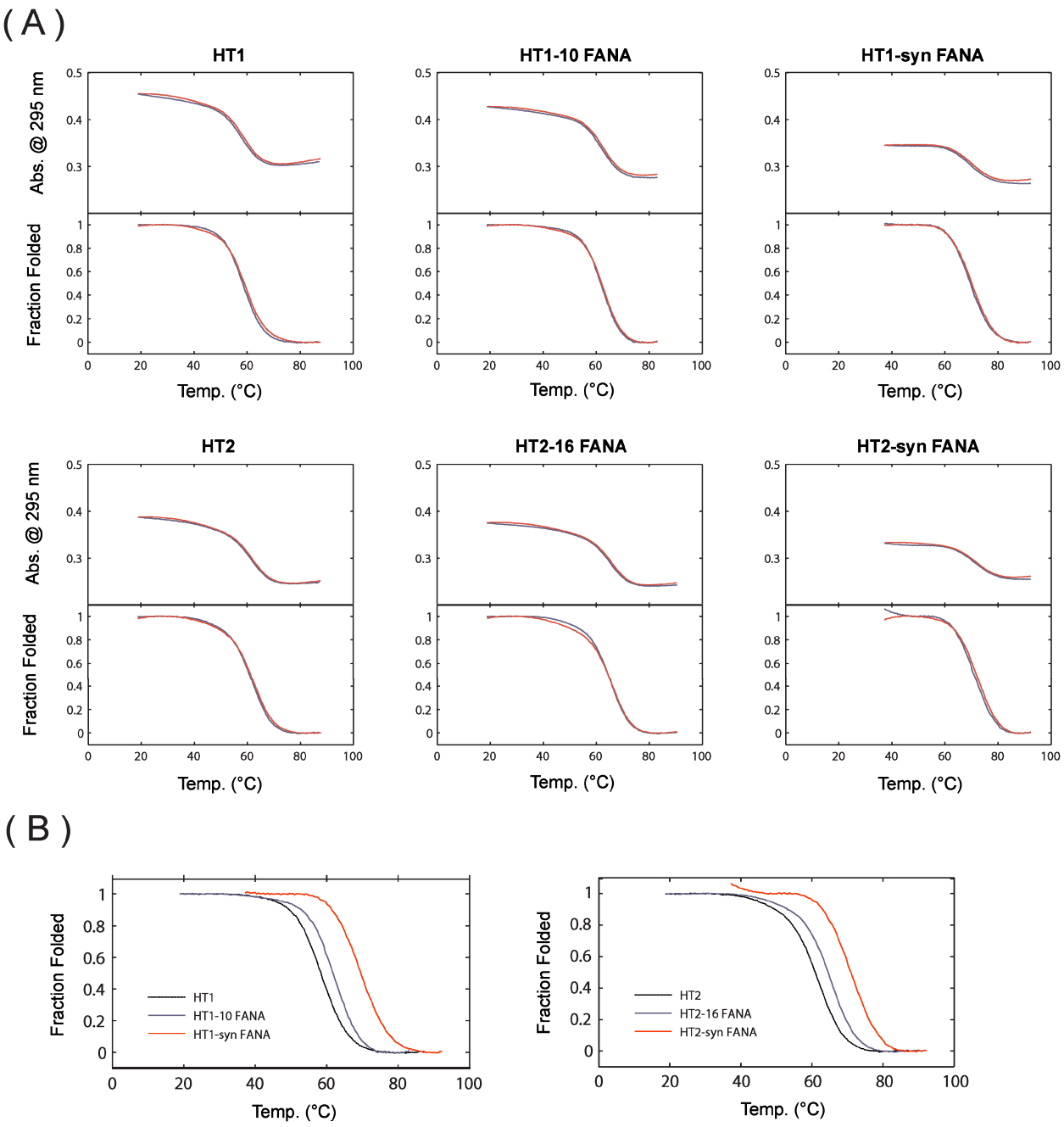

Figure S3: Melting Curves: **(A)** UV absorbance at 295nm is shown along with the normalized "Fraction Folded" curves for unmodified and ^{FANA}G modified sequences. Heating curve (red) and cooling curves (blue) are shown. **(B)** Superimposed cooling curves are shown for unmodified (black), single ^{FANA}G substituted (Blue) and syn-position ^{FANA}G substituted (red) HT1 (left) and HT2 (right) sequences.

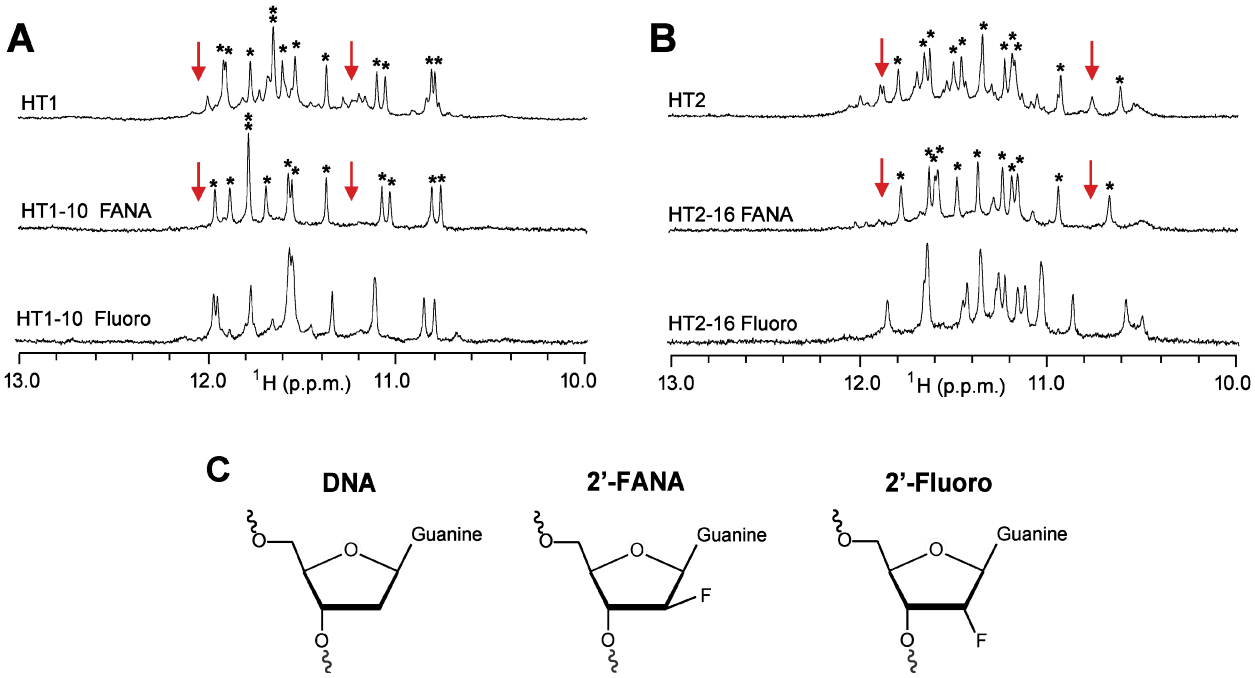

Figure S4: NMR Imino proton comparison of **(A)** HT1 and **(B)** HT2 sequences with single modifications of 2'F-ANA-guanine (^{FANA}G) and 2'-F-guanine (^{F}G). **(C)** Chemical schematics of modified nucleosides.

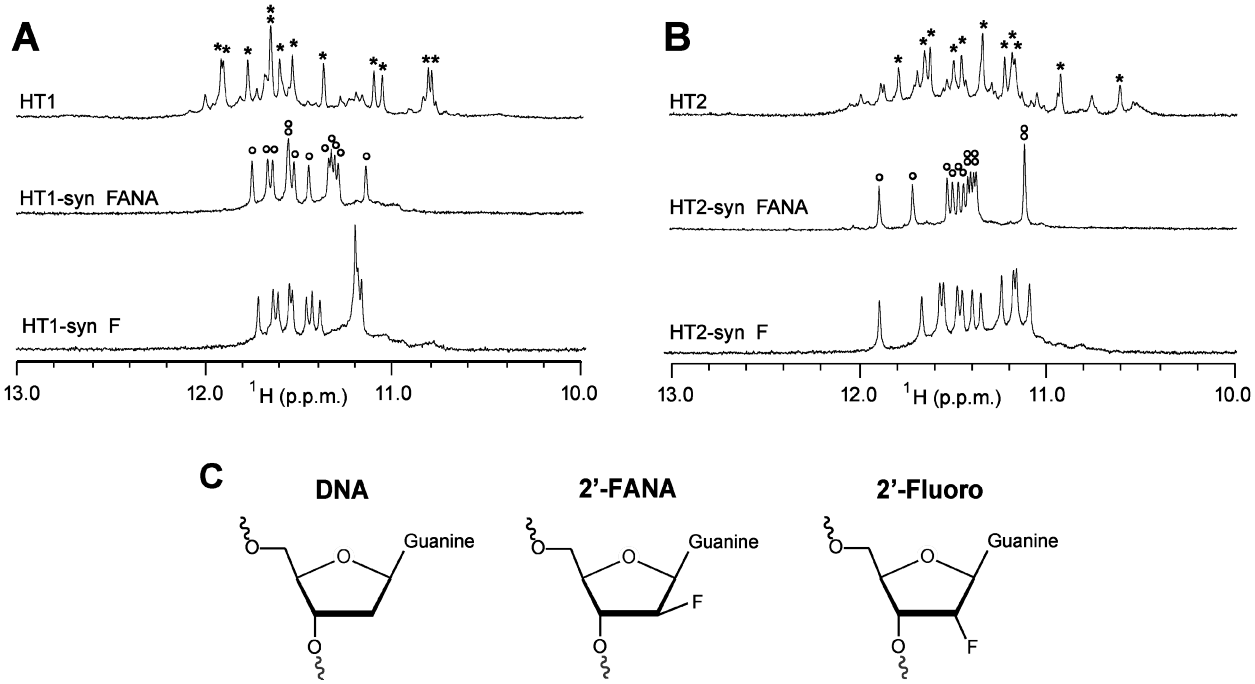

Figure S5: NMR Imino proton comparison of **(A)** HT1 and **(B)** HT2 sequences with *syn*-position modifications of 2'F-ANA-guanine (^{FANA}G) and 2'-F-guanine (^{F}G). **(C)** Chemical schematics of modified nucleosides.

SUPPORTING REFERENCES

1. M. Adrian, B. Heddi and A. T. Phan, *Methods*, 2012, **57**, 11-24.
2. C. J. Lech, J. K. C. Lim, J. M. W. Lim, S. Amrane, B. Heddi and A. T. Phan, *Biophys J*, 2011, **101**, 1987-1998.
3. P. Balagurumoorthy, S. K. Brahmachari, D. Mohanty, M. Bansal and V. Sasisekharan, *Nucleic Acids Res*, 1992, **20**, 4061-4067.
4. B. Heddi and A. T. Phan, *J Am Chem Soc*, 2011, **133**, 9824-9833.
5. Y. Xue, Z. Y. Kan, Q. Wang, Y. Yao, J. Liu, Y. H. Hao and Z. Tan, *J Am Chem Soc*, 2007, **129**, 11185-11191.