An RNA Hairpin to G-Quadruplex Conformational Transition

Anthony Bugaut,*[†,§] Pierre Murat,[†,§] and Shankar Balasubramanian*[†,‡]

[†]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.
[‡]Cambridge Research Institute, Cancer Research UK, Li Ka Shing Center, Cambridge CB2 0RE, U.K.

S Supporting Information

ABSTRACT: RNA molecules can fold into noncanonical structures such as the four-stranded structures known as G-quadruplexes. G-quadruplexes in the transcriptome have recently emerged as relevant regulatory elements of gene expression. Conformational transitions in RNA molecules offer an important way to regulate their biological functions. Here we report on the competition between a canonical hairpin structure and a G-quadruplex structure within an RNA molecule. We show that the conformational preference strongly depends on the relative amounts of mono- and divalent metal ions present in solution. In our system, the G-quadruplex, whose formation is not predicted by available predictive RNA folding programs, is the major conformer at physiologically relevant K^+ and Mg^{2+} concentrations. Furthermore, we show that a synthetic small molecule can displace the structural dynamic equilibrium in favor of the hairpin conformer. This work highlights a new and important level of complexity in RNA folding that could be relevant to the biological functions and targeting of RNAs comprising G-quadruplex motifs.

The biological functions of RNA molecules generally depend on their folding properties.[1] Some RNAs can adopt several stable folds that coexist in an intramolecular conformational equilibrium.[2,3] Changes in RNA conformation in response to fluctuations in intracellular conditions provide a mechanism that can regulate gene expression.[3,4] This is exemplified by naturally occurring RNA regulatory elements that respond to extrinsic stimuli such as temperature, pH, and ionic conditions or to interacting trans-acting factors such as proteins, nucleic acids, and small-molecule metabolites.[3,4]

Besides canonical purine (Pu)−pyrimidine (Py) helix-based structural elements, some RNA sequences may also adopt noncanonical base-pairing schemes leading to alternative folds that are fundamentally different from helix-based RNA structures and would be unexpected if only Pu/Py base pairs are considered. This includes guanine (G)-rich four-stranded structures, called G-quadruplexes, which arise from Hoogsteen-type hydrogen bonds between Gs.[5] RNA G-quadruplexes have recently emerged as motifs that can regulate gene expression and have been implicated in several aspects of RNA metabolism such as transcription termination, splicing, and translation.[6] Furthermore, they have also been used as switchable modules that allow artificial control of the function of attached RNA sequences.[7] Therefore, it is becoming increasingly important to obtain a greater understanding of the properties of such RNA folds.

Biophysical studies on RNA G-quadruplexes published to date have focused on small pieces of G-rich RNAs that unambiguously fold into G-quadruplex structures[8] but have not addressed the question of RNA G-quadruplex formation in competition with helix-based structures, which could occur in longer stretches of RNA. Herein we report the first description of the competitive formation of a helix-based hairpin structure and a G-quadruplex structure within an RNA molecule.

We designed an RNA molecule, HpQd, that has the potential to fold into two mutually exclusive intramolecular structures: a hairpin structure (Hp) and a G-quadruplex structure (Qd) (Figure 1). The hairpin-forming sequence was derived from a
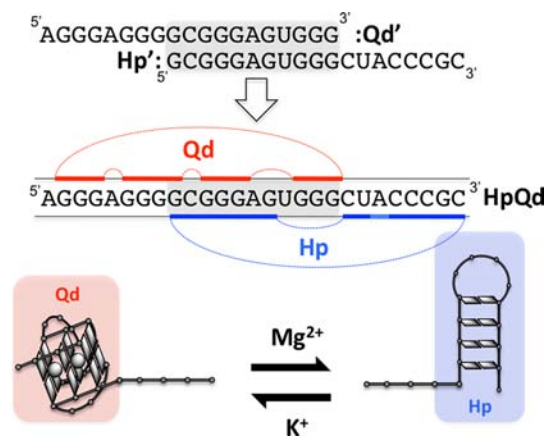
Figure 1. A bistable RNA conformational switch (HpQd) under cation-mediated control.

previously documented hairpin structure belonging to the VPK pseudoknot structure in the mouse mammary tumor virus, whose the thermodynamic stability strongly depends on the Mg^{2+} concentration.[9] The G-quadruplex-forming sequence was derived from a G-quadruplex motif previously identified within the 5′-UTR of human *N-RAS* messenger RNA.[10] To design HpQd, we introduced (i) three G/C base permutations within the stem-forming sequence of the wild-type VPK hairpin sequence and (ii) two mutations (A to U and G to C) within the third loop of the wild-type NRAS G-quadruplex (Figure S1 in the Supporting Information). These modifications adapted the wild-type sequences in such a way that the mutated

sequences (denoted as Hp′ and Qd′, respectively) share an identical 11 nucleotide (nt) sequence element (rGCGGGA-GUGGG), leading to structural ambivalence between the two conformers within HpQd (Figure 1 and Figure S1). Circular dichroism (CD)[12] and thermal difference[13] spectra of the individual Hp′ and Qd′ reference sequences (Figure 1) confirmed that they retained the hairpin and the G-quadruplex structures, respectively, adopted by the parent RNA sequences (Figure S2). It is notable that even though HpQd can possibly fold into a G-quadruplex structure, available RNA folding prediction programs such as the popular MFold[11] do not yet include a consideration of Hoogsteen base pairs and thus cannot predict G-quadruplex folds (Figure S3).

We then subjected the Hp′ and Qd′ references to UV melting experiments. The melting temperatures obtained in 10 mM sodium cacodylate (pH 7.0) containing 1 mM KCl were 71 ± 1 and 62 ± 1 °C, respectively. Van't Hoff analysis of the melting profiles[14] indicated that the thermodymamic stabilities of the two structures differed by ~ 5 kJ mol^{-1}, suggesting an 85:15 equilibrium in favor of Hp formation over Qd formation within the HpQd sequence. Under the same conditions, the melting profiles of HpQd recorded at 260, 280, and 295 nm all displayed a reversible hyperchromic sigmoidal curve with a transition point at 70 ± 1 °C (Figure S4), in agreement with the hairpin structure being the major conformer.[15]

However, the preferential formation of an RNA structure can strongly depend on the environmental conditions, in particular the relative abundance of metal ions in solution.[16] For example, the thermodynamic stability of RNA hairpin structures has been reported to be dependent on the concentration of divalent cations, especially Mg^{2+} (the most abundant divalent cation in the cytoplasm of a cell), whereas G-quadruplex formation generally depends on the concentration of monovalent cations, especially K$^+$ (the most abundant monovalent cation in the cytoplasm of a cell). For DNA, there was an early report by Hardin et al.[17] on a cation-dependent equilibrium between a DNA hairpin and a tetramolecular G-quadruplex from a d(CGCG$_3$GCG) oligonucleotide showing that K$^+$ ions strongly favor the G-quadruplex fold. In addition, short oligonucleotides mimicking the G-rich DNA strands of the fragile X chromosome repeats have been shown to form either hairpins or tetraplexes depending on the concentrations of Na$^+$ and K$^+$ ions and the temperature.[18]

UV melting studies of the Qd′ reference revealed that the melting temperature of the G-quadruplex was strongly influenced by the concentration of K$^+$ in solution, whereas the Mg^{2+} concentration had no or little effect (Figure 2). On the other hand, as expected, the concentration of Mg^{2+} had a strong effect on the thermal stability of the Hp′ hairpin structure, but it had only a moderate effect on the Qd′ G-quadruplex (Figure 2). These initial data suggested a possible cation-dependent competitive hairpin−quadruplex equilibrium within the HpQd RNA molecule.

We then used 1D ^{1}H NMR spectroscopy to investigate the propensity of HpQd to adopt either of its two possible conformers (Hp or Qd) preferentially under different ionic conditions at thermodynamic equilibrium. The imino proton NMR resonances of nucleic acid structures generally reflect their base-pairing arrangement, and indeed, 1D NMR of imino protons has previously been used to study hairpin to hairpin transitions within small RNA molecules.[19] To start, we recorded the ^{1}H NMR spectrum of HpQd in 10 mM phosphate-buffered saline (PBS) (pH 7.0) in the absence of
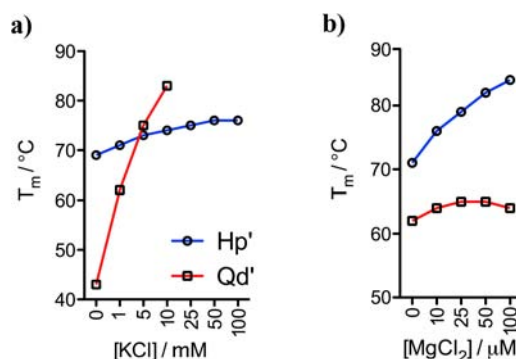
Figure 2. Influence of the K$^+$ and Mg^{2+} concentrations on the thermal stabilities of the hairpin (Hp′) and G-quadruplex (Qd′) RNA structures. Denaturation studies were performed in 10 mM sodium cacodylate (pH 7.0) supplemented with (a) increasing amounts of KCl (K$^+$ dependence) and (b) 1 mM KCl and increasing amounts of MgCl$_2$ (Mg^{2+} dependence).

any added salt. The spectrum revealed two patterns of imino proton signals encompassing two distinct spectral regions (10.5−11.8 and 12−13 ppm), suggesting the coexistence of two structural RNA conformers in solution (Figure 3). On the basis
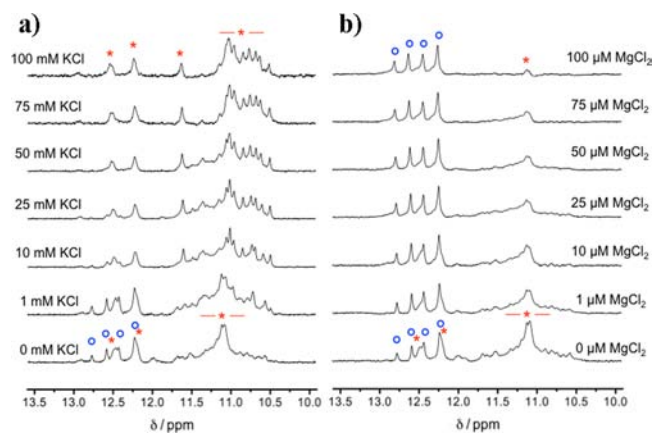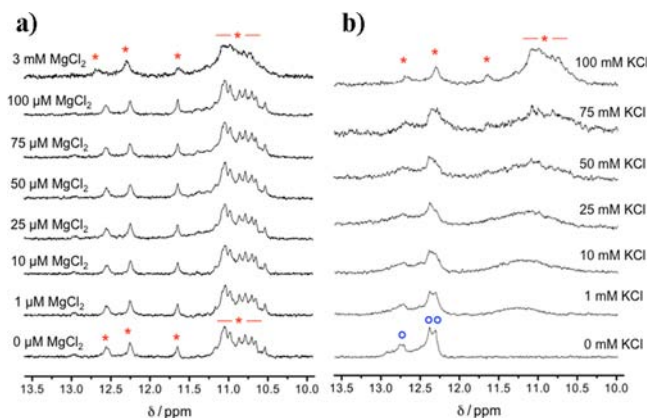

Figure 3. ^{1}H NMR titrations of the HpQd sequence with increasing amounts of (a) KCl and (b) MgCl$_2$. The initial spectra were acquired in 10 mM PBS (pH 7.0) in the absence of added KCl or MgCl$_2$. Signals labeled with red stars and blue circles correspond to the Qd and Hp structures, respectively.

of the ^{1}H NMR spectra recorded for the Hp′ and Qd′ reference sequences (Figure S6) these imino fingerprints were attributed to the Qd (below 12 ppm) and Hp (above 12 ppm) structures. Indeed, imino proton peaks from 12 to 14 ppm are characteristic of Watson−Crick base pairs, whereas those from 10 to 12 ppm are characteristic of Hoogsteen base pairs and indicate G-tetrad formation.[20] Next, we titrated the HpQd sequence with increasing concentrations of KCl (0−100 mM). As shown in Figure 3a, this led to the disappearance of the imino peaks characteristic of the hairpin motif along with remodelling and sharpening of the peaks below 12 ppm, indicating a displacement of the equilibrium in favor of a unique stable G-quadruplex motif. For KCl concentrations above 25 mM, the characteristic signals of the hairpin structure were no longer detectable.[21] On the other hand, titration of HpQd with increasing amounts of MgCl$_2$ (0−100 μM) led to the disappearance of the Hoogsteen imino proton resonances and sharpening of the ^{1}H signals between 12 and 13 ppm

(Figure 3b), indicating a displacement of the structural equilibrium in favor of the hairpin. At 100 μM MgCl$_2$ concentration, the signals from the quadruplex conformer almost disappeared, and only the hairpin signals remained. These results illustrated that the structural preference of HpQd is controlled by the cations present in solution, with K$^+$ promoting the formation the G-quadruplex motif and Mg^{2+} promoting the hairpin form.

We next assessed the competitive formation of the two RNA structures in an environment that better mimics a cellular context where both K$^+$ and Mg^{2+} ions were present at high concentration. The ^{1}H NMR spectrum of HpQd recorded in 10 mM PBS (pH 7.0) and 100 mM KCl indicated the preferential formation of the Qd G-quadruplex structure (Figures 3a and 4a). Upon addition of increasing amounts of
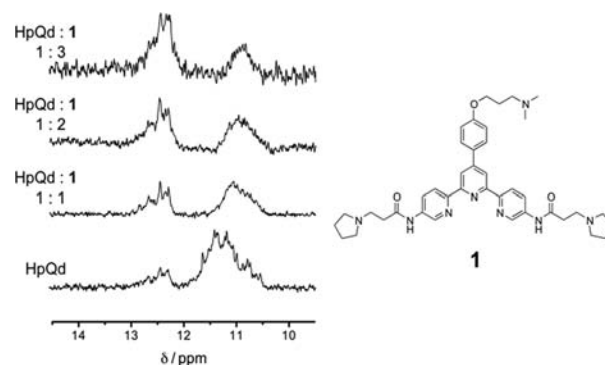

Figure 4. ^{1}H NMR titrations of the HpQd sequence with increasing amounts of (a) MgCl$_2$ and (b) KCl. The initial spectra were acquired in 10 mM PBS (pH 7.0) in the presence of (a) 100 mM KCl and no MgCl$_2$ and (b) 3 mM MgCl$_2$ and no KCl. Signals labeled with red stars and blue circles correspond to the Qd and Hp structures, respectively.

MgCl$_2$ (0−3 mM), no significant modification of the imino proton NMR spectrum of HpQd was observed (Figure 4a), indicating that the structural equilibrium was not displaced and that, once formed in the presence of 100 mM KCl, the G-quadruplex motif is the more stable conformer even at near-physiological millimolar Mg^{2+} concentrations. When HpQd was pre-equilibrated in 10 mM PBS (pH 7.0) and 3 mM MgCl$_2$, titration with increasing concentrations of KCl (0−100 mM) revealed the disappearance of the imino peak signals attributed to the hairpin motif and the emergence of a imino proton envelope below 12 ppm corresponding the Qd G-tetrad resonance signals (Figure 4b). It is noteworthy that the NMR spectra obtained at the ends of the two titrations were identical. In summary, these data demonstrated that HpQd preferentially folds into a G-quadruplex conformation under near-physiological conditions in presence of 100 mM KCl and 3 mM MgCl$_2$, regardless of its initial state (Figure S7).

RNA folds can also be manipulated by external triggers such as synthetic small molecules that can act as regulators for chemical biology studies.[22] We next investigated the use of a synthetic molecule to control the equilibrium between the two secondary structures. While the G-quadruplex conformation is the favored conformation of HpQd under near-physiological ionic conditions, we sought to determine whether a small-molecule nucleic acid binder could influence this structural

preference. We selected a triarylpyridine (TAP) derivative, **1** (Figure 5), that we previously showed can disrupt DNA G-

Figure 5. ^{1}H NMR titration of the HpQd sequence with increasing amounts of **1**. The initial spectrum was recorded in 10 mM PBS (pH 7.0) containing 10 mM KCl and 200 μM MgCl$_2$.

quadruplex formation.[23] We first performed a titration of the Qd′ reference sequence with **1**. Increasing concentrations of **1** led to a progressive decrease in the imino proton resonance envelope between 10 and 12 ppm, which almost completely disappeared after the addition of 3 equiv of **1** (Figure S8a), indicating small-molecule-mediated disruption of the quadruplex motif. In a control experiment, titration of the Hp′ reference revealed only minor alterations of the imino signals, and importantly, the imino resonances corresponding to the hairpin structure still persisted even after the addition of 3 equiv of **1**, indicating that the hairpin fold was not disrupted (Figure S8b). Next, we titrated HpQd with **1** under ionic conditions favoring the formation of the G-quadruplex motif. We observed that increasing the concentration of **1** led to a decrease in the imino signal envelope below 12 ppm with a concomitant increase of the signal above 12 ppm (Figure 5). Structural interconversion was confirmed by integrating the changes in the imino resonance envelope areas triggered by the addition of **1** (Figure S9). Collectively, these results demonstrated the ability of the TAP derivative to displace the structural equilibrium in favor of the hairpin motif by unfolding the G-quadruplex motif.

In conclusion, we have shown that a reversible hairpin G-quadruplex structural equilibrium within an RNA molecule is controlled by cations and can be manipulated by a synthetic small molecule. In view of the fact that the regulatory functions of RNA molecules are often linked to conformational transitions,[3] these observations have important implications for the inevitable competition between such structures and their functions in nature as well as their targeting. Thus, a bioinformatic search for naturally occurring hairpin/G-quadruplex switchable elements in the transcriptome now represents a challenging but exciting perspective (Figure S10).

■ ASSOCIATED CONTENT

Ⓢ Supporting Information
Materials and methods and Figures S1−S10. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
ab605@cam.ac.uk; sb10031@cam.ac.uk

Author Contributions

§A.B. and P.M. contributed equally.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) (a) Brion, P.; Westhof, E. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113. (b) Batey, R. T.; Rambo, R. P.; Doudna, J. A. *Angew. Chem., Int. Ed.* **1999**, *38*, 2326.

(2) Micura, R.; Hobartner, C. *ChemBioChem* **2003**, *4*, 984.

(3) Dethoff, E. A.; Chugh, J.; Mustoe, A. M.; Al-Hashimi, H. M. *Nature* **2012**, *482*, 322.

(4) (a) Nagel, J. H. A.; Pleij, C. W. A. *Biochimie* **2002**, *84*, 913. (b) Huthoff, H.; Berkout, B. *Biochemistry* **2002**, *41*, 10439. (c) Tucker, B. J.; Breaker, R. R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 342. (d) Schwalbe, H.; Buck, J.; Fürtig, B.; Noeske, J.; Wöhnert, J. *Angew. Chem., Int. Ed.* **2007**, *46*, 1212. (e) Kortmann, J.; Narberhaus, F. *Nat. Rev. Microbiol.* **2012**, *10*, 255.

(5) *Quadruplex Nucleic Acids*; Neidle, S., Balasubramanian, S., Eds.; RSC Biomolecular Sciences: Cambridge, U.K., 2006.

(6) (a) Patel, D. J.; Phan, A. T.; Kuryavyi, V. *Nucleic Acids Res.* **2007**, *35*, 7429. (b) Collie, G. W.; Parkinson, G. N. *Chem. Soc. Rev.* **2011**, *40*, 5857. (c) Halder, K.; Hartig, J. S. *Met. Ions Life Sci.* **2011**, *9*, 125. (d) Bugaut, A.; Balasubramanian, S. *Nucleic Acids Res.* **2012**, *40*, 4727. (d) Millevoi, S.; Moine, H.; Vagner, S. *Wiley Interdiscip. Rev.: RNA* **2012**, *3*, 495.

(7) (a) Wieland, M.; Hartig, J. S. *Angew. Chem., Int. Ed.* **2006**, *45*, 5875. (b) Henn, A.; Joachimi, A.; Gonçalves, D. P.; Monchaud, D.; Teulade-Fichou, M.-P.; Sanders, J. K. M.; Hartig, J. S. *ChemBioChem* **2008**, *9*, 2722.

(8) (a) Saccà, B.; Lacroix, L.; Mergny, J.-L. *Nucleic Acids Res.* **2005**, *33*, 1182. (b) Arora, A.; Maiti, S. *J. Phys. Chem. B* **2009**, *113*, 10515. (c) Joachimi, A.; Benz, A.; Hartig, J. S. *Bioorg. Med. Chem.* **2009**, *17*, 6811. (d) Zhang, D. H.; Fujimoto, T.; Saxena, S.; Yu, H. Q.; Miyoshi, D.; Sugimoto, N. *Biochemistry* **2010**, *49*, 4554. (e) Zhang, A. Y.; Bugaut, A.; Balasubramanian, S. *Biochemistry* **2011**, *50*, 7251.

(9) Gonzalez, R. L.; Tinoco, I., Jr. *J. Mol. Biol.* **1999**, *289*, 1267.

(10) Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. *Nat. Chem. Biol.* **2007**, *4*, 218.

(11) Zucker, M. *Nucleic Acids Res.* **2003**, *31*, 3406.

(12) Paramasivan, S.; Rujan, I.; Bolton, P. H. *Methods* **2007**, *43*, 324.

(13) Mergny, J.-L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. *Nucleic Acids Res.* **2005**, *33*, No. e138.

(14) Mergny, J.-L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515.

(15) CD spectroscopy confirmed that changing the RNA concentration had no effect on the HpQd conformation (Figure S5).

(16) Pyle, A. M. *J. Biol. Inorg. Chem.* **2002**, *7*, 679.

(17) Hardin, C. C.; Watson, T.; Corregan, M.; Bailey, C. *Biochemistry* **1992**, *31*, 833.

(18) Fotjik, P.; Kejnovska, I.; Vorlickova, M. *Nucleic Acids Res.* **2004**, *32*, 298.

(19) (a) Höbartner, C.; Micura, R. *J. Mol. Biol.* **2003**, *325*, 421. (b) Hörbarter, C.; Ebert, M.-O.; Jaun, B.; Micura, R. *Angew. Chem., Int. Ed.* **2002**, *41*, 605.

(20) (a) Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. *Nat. Chem. Biol.* **2005**, *3*, 167. (b) Fürtig, B.; Richter, C.; Wöhnert, J.; Schwalbe, H. *ChemBioChem* **2003**, *4*, 936.

(21) The two left peaks above 12 ppm did not correspond to the Hp structure but may correspond to the formation of Watson−Crick base pairs between nucleobases of the quadruplex flanking sequence and free nucleobases in the loop of the quadruplex.

(22) Thomas, J. R.; Hergenrother, P. J. *Chem. Rev.* **2008**, *108*, 1171.

(23) Waller, Z. A. E.; Sewitz, S. A.; Hsu, S.-T. D.; Balasubramanian, S. *J. Am. Chem. Soc.* **2009**, *131*, 12628.