G-Quadruplex Folding

DOI: 10.1002/anie.201206475

Decrease in RNA Folding Cooperativity by Deliberate Population of Intermediates in RNA G-Quadruplexes**

*Chun Kit Kwok, Madeline E. Sherlock, and Philip C. Bevilacqua**

RNA folding is central to diverse biological and chemical processes, including gene regulation and biosensing. Cooperativity is a key feature of RNA folding, and critical to the folding of small and large RNAs[1] alike and to the response of naturally occurring riboswitches.[2] Some biological RNAs fold with little or no cooperativity, whereas others exhibit high cooperativity (i.e., steep folding transitions).[2,3] Concerning biosensing, conventional or "classical" biosensors without cooperative behavior bind an analyte with a responsive range of approximately 100-fold in analyte concentration, using 10–90% of the maximal signal. When the analyte concentration is above or below this range, however, "dead zones" are found wherein the sensor will not work;[4] as such, biosensors with extended dynamic ranges are needed. Integrating negative (or "anti-") cooperativity into RNA folding provides one way of achieving this outcome. The goals of our study were to gain new fundamental insights into the folding of biological RNAs and to aid the development of biosensors with diverse response ranges through the development of a series of RNA scaffolds with tunable cooperativity.

Several types of biopolymers that show a broad response to their ligands have been reported. Broad-response detection of pH and ionic strength were attained by coupling signal output to a one-state downhill-folding protein scaffold.[5] Broad-response sensing was also achieved by mixing together three (or more) biosensors that differ in K_d of the analyte by approximately one to two orders of magnitude.[6] Using such an approach, broad-response sensors were prepared that detect changes in pH value[6b] or in DNA[6a] or protein[6c] concentration.

Herein, we developed a simpler approach to broad-range folding and sensing, which we link to anticooperative RNA folding. We use just one rationally designed RNA sequence as a scaffold for ligand detection. Our strategy exploits the deliberate population of intermediates along both the RNA-folding and analyte-binding pathways (Figure 1B). We used a G-quadruplex sequence (GQS) to explore the relationship between RNA sequence and anticooperative RNA folding. Analyte sensing and riboswitching were demonstrated by monitoring the concentration of K^+ ions with the fraction of folded GQS RNA as the read out, which was determined using either circular dichroism (CD) or fluorescence spectroscopy.

A G-quadruplex is a nucleic acid motif formed by guanine-rich RNA or DNA sequences of the pattern $G_xL_aG_xL_bG_xL_cG_x$, where x is the number of guanine residues that are involved in G-quartet formation, and L_a, L_b, and L_c are the loops between the four G-strands that can be of the same or different length and sequence;[7] such a sequence is referred to herein as a "GX" GQS. A GQS generally consists of two or more stacked quartet planes, with a K^+ ion between the planes, which stabilizes the quadruplex.[8] Extensive studies on DNA and RNA GQSs in different sequence contexts have revealed that GQSs bind K^+ ions selectively

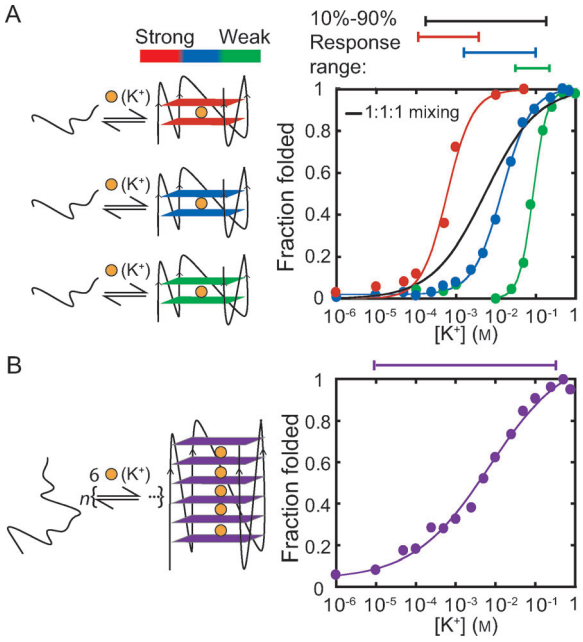

Figure 1. Two approaches to achieve broad-response sensing. A) Mixing of different sensors, as previously reported.[6a] Mixing together three different G2 GQS sequences with a ΔK_d of approximately 200-fold (G2w2, G2m1, and G2s2). The K^+ ion binding affinity (10–90% response range) towards individual GQS is represented by the color-coded bars above the plot: red = strong, blue = medium, and green = weak binders. The black curve simulates the response of a 1:1:1 mixture of the three G2 GQSs. B) Extending the length of the G-stretch of a GQS developed herein. Shown here is the G6 GQS.

[*] C. K. Kwok,[+] M. E. Sherlock,[+] Prof. Dr. P. C. Bevilacqua
Department of Chemistry and Center for RNA Molecular Biology
The Pennsylvania State University
University Park, PA 16802 (USA)
E-mail: pcb5@psu.edu
Homepage: http://research.chem.psu.edu/pcbgroup

[+] These authors contributed equally to this work.

[**] This study was supported by a Human Frontier Science Program (HFSP) Grant RGP0002/2009-C to P.C.B. and Penn State Summer Discovery Grant to M.E.S. We thank Prof. Sarah Assmann, Prof. Bratoljub Milosavljevic, and Dr. Melissa Mullen for helpful discussions.

Supporting information for this article (experimental details) is available on the WWW under http://dx.doi.org/10.1002/anie.201206475.

over other types of metal ions.[9] In addition, GQSs have been used as biosensors to determine K^+ ion concentration, in cells[10] and other samples.

RNA GQSs predominately form a parallel topology upon addition of K^+ ions.[11] This topology leads to an increase in ellipticity at approximately 260 nm. CD spectroscopy has been used extensively to study the folding of other G-quadruplexes.[9b,12] More recently, it has been reported that DNA G-quadruplexes are intrinsically fluorescent.[13] To detect GQS formation herein, we used both CD and fluorescence spectroscopy to monitor the change in folding upon K^+ ion addition (see Supporting Information). In addition, we show for the first time that RNA GQSs are intrinsically fluorescent and that the fluorescence signal responds to K^+ ion concentration in a fashion similar to the CD signal.

First, we investigated the folding of three G2 GQSs termed G2w1, G2m1, and G2s1, where w = weak, m = medium, and s = strong binding of K^+ ions, using CD signal.[14] These sequences have positive or neutral cooperativity and are related to those from our recent report.[12b] Binding strength was controlled by the sequence and length of the loops. $K^+_{1/2}$ values (the concentration of K^+ ions at which half the GQSs were folded) were determined as 410 ± 70 mM, 14 ± 1 mM, and 2.2 ± 0.5 mM, respectively (Table 1). We then identified a second set of G2 GQSs (G2w2, G2m2, and G2s2) having a higher A-content in the loops, which led to similar K^+ ion binding profiles, and somewhat higher Hill coefficients (n; Table 1).[12b] Folding transitions for all six G2 GQSs are provided in Figure 2. The G2 GQSs were then classified as having either a classical response (approximately two orders of magnitude) or digital response (less than two orders of magnitude) depending on whether their n values were approximately one or two, respectively.[15]

To determine the effective dynamic range of response possible by using the previously published approach of mixing

Table 1: GQS folding parameters.

GQS motif[a]	Fitting[b]	n[b,d]	$K^+_{1/2}$ [mM][b,d]
G2w1	2-state	2.2 ± 0.3	410 ± 70
G2m1	2-state	1.2 ± 0.1	14 ± 1
G2s1	2-state	1.2 ± 0.1	2.2 ± 0.5
G2w2[e]	2-state	2.7 ± 0.1	117 ± 28
G2m2[e]	2-state	2.5 ± 0.4	14 ± 1
G2s2[e]	2-state	1.7 ± 0.4	0.8 ± 0.2
G3m1	2-state	0.5 ± 0.1	0.7 ± 0.2
G3m1 (tr1)[c]	3-state	1.5 ± 0.1	0.06 ± 0.01
G3m1 (tr2)[c]	3-state	1.3 ± 0.2	6.4 ± 2.8
G4m1	2-state	0.5 ± 0.1	2.7 ± 2.2
G4m1 (tr1)[c]	3-state	1.2 ± 0.2	0.06 ± 0.06
G4m1 (tr2)[c]	3-state	1.2 ± 0.1	12 ± 1
G5m1	2-state	0.7 ± 0.1	3.7 ± 0.7
G6m1	2-state	0.5 ± 0.1	12 ± 2

[a] Full sequences are provided in Table S1. [b] GQS folding parameters were obtained from fitting CD titration experiments using a two-state model (Equation S1). [c] In several cases ("tr1" and "tr2" for transitions 1 and 2), two distinct transitions were observed and a three-state fitting was performed (Equation S2). [d] Values are an average ± standard deviation, obtained from three separate experiments. [e] Values are from Ref. [12b].

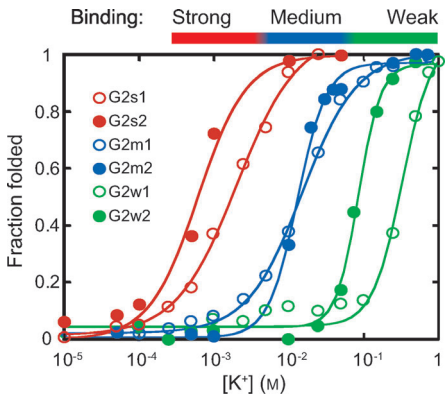

Figure 2. Graph of the fraction of G2 GQSs folded at different K^+ ion concentrations. Data at 262 nm were well fit by a two-state Hill equation (Equation S1). n and $K^+_{1/2}$ values are provided in Table 1. See Figures S1–S6 for full CD spectra and individual fits.

probes,[6] we selected G2w2, G2m1 and G2s2 GQSs, which vary by approximately two orders of magnitude in $K^+_{1/2}$ value. We simulated a linear broad-response signal (Supporting Information, Equation S4) and fit the resulting response with Equation S1 (Figure 1 A, black line). This gave an apparent $K^+_{1/2}$ value of 12 mM and an apparent Hill coefficient of 0.6. This mixing approach was thus expected to widen the 10–90 % response range to approximately three orders of magnitude in K^+ concentration, as represented by the black bar above Figure 1 A.

We next sought to attain broad-response sensing of K^+ ions by deliberately populating intermediates. To do so, we took advantage of our previous observation that a G2 GQS generally has a higher Hill coefficient and thus a steeper (i.e. more digital) response to K^+ concentration than an otherwise similar G3 GQS.[12b] We had attributed this trend to G2 sequences having poorly populated folding and binding intermediates—a phenomenon that generally enhances folding cooperativity.[16] Fewer populated intermediates in G2 sequences was attributed to fewer ways to assemble the quartets incorrectly, and to the presence of just one K^+ ion binding site, which precludes ion–ion interactions.[12b] We then reasoned that the inverse may hold: that is, increasing the length of the G-stretches in a single sequence might enhance the number and population of folding and binding intermediates, thereby lessening folding and binding cooperativity and broadening the response to analyte concentration (Figure 1 B).

To test this idea, we selected G2m1—which has the sequence $G_2UAG_2UAG_2CG_2$ and a conventional two orders of magnitude response—as our model GQS and systematically extended the length of the G-stretch, giving G3m1–G6m1, where G6m1 has the sequence $G_6UAG_6UAG_6CG_6$ (all sequences are shown in Table S1). Fitting of K^+ ion titrations determined by CD spectroscopy for G3m1 and G4m1 revealed a distinctly three-state folding behavior (Figure 3, green and magenta lines)—a conclusion that was further supported by plots of the residuals (Figure S8 and S10). Notably, each of the two transitions for G3m1 and G4m1 have a Hill coefficient of around one (Table 1), which gives well-separated transitions and a well-populated folding intermedi-

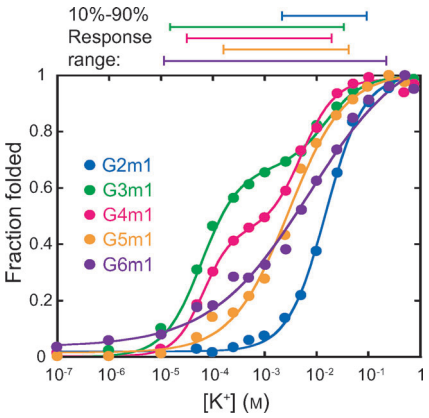

Figure 3. Graph of the fraction of G2m1–G6m1 GQSs folded at 262 nm by CD spectroscopy. G3m1 and G4m1 were fit to a three-state equation (Equation S2), while G2m1, G5m1, and G6m1 were fit to a two-state equation (Equation S1). The response range for each GQS is defined from 10–90% folded and is depicted with bars of matching color above the panel. The $K^+_{1/2}$ and *n* values are in Table 1. See Figure S2 and S7–S12 for full CD spectra and individual fits.

ate as well as agreeing with the three-state model shown in Equation S2. This intermediate may represent slipped quartets (i.e. with X-1 or X-2 quartets), and/or weakened binding of the second or third K^+ ion in G3m1 and G4m1, respectively, owing to ion–ion repulsion (see below). On one hand, population of the folding intermediate broadened the response, as anticipated; but, on the other hand, the response was no longer linear and instead had a prominent dead zone in the center of the graph (Figure 3, green and magenta lines).[17] Nonetheless, intermediates clearly became populated as the number of quartets increased, which motivated us to examine longer GQSs.

Potassium ion titrations of the G5m1 and G6m1 sequences are provided in Figure 3 (gold and purple lines, respectively) and are largely linear in the 10–90% transition region (i.e. there is no significant dead zone like in G3m1 and G4m1) and broad in their response. Given such behavior, these traces could be fit to Equation S1 for an apparent two-state transition (Figure S11, S12), which gave apparent Hill coefficients of just 0.7 ± 0.1 and 0.5 ± 0.1, respectively (Table 1). Reversion to apparent two-state behavior for G5m1 and G6m1 is due to the presence and population of many folding and binding intermediates, so many that the multiple transitions blur into one very broad transition. This idea is supported by the trends shown above and the fractional Hill coefficients. Such intermediates may represent an ensemble of states containing slipped quartet registers, and/or weakened binding of the fourth and fifth K^+ ions, owing to repulsion from the other bound K^+ ions (see below). Overall, G5m1 and G6m1 have astonishingly wide linear responses to K^+ ion concentration, spanning approximately 3–4.5 orders of magnitude as compared to just two for G2m1 (Figure 3). Indeed, these response ranges are similar to, or exceed, those extrapolated from the above G2 sequence mixtures (Figure 1A).

Next, we explored fluorescence as an alternative method of monitoring GQS folding. As mentioned above, it was recently shown that the intrinsic fluorescence of DNA G-quadruplexes increases upon quadruplex formation;[13] moreover, fluorescence is often used in developing biosensors. The same potassium titrations described above for CD detection were performed on unlabeled G2m1 and G6m1 and monitored with fluorescence detection. We found that these RNA G-quadruplexes, like their DNA counterparts, exhibit an increase in fluorescence intensity upon folding when excited at 270 nm (Figure S13, S14). The emission intensity increased upon addition of K^+ ions, with an emission maximum of approximately 360 nm. Titration curves showing the fraction of folded G-quadruplexes were produced using the data at 360 nm (Figure 4), and they were compared to the results from CD spectroscopy experiments. The G6m1 data were again fit to an apparent two-state model, owing to the presence of multiple intermediate states. The $K^+_{1/2}$ values and Hill coefficients between the two methods agree within the standard deviation of each other (Figure 4), making both methods equally useful in GQS detection. The intrinsic fluorescence of the G-quadruplexes is valuable not only because fluorescence is a convenient way to monitor folding and binding, but also because it can be used on unmodified oligonucleotides and thus avoids potentially perturbing the interaction.

To ensure the signal changes were due to intramolecular and not intermolecular GQS folding, concentration-dependent thermal denaturation experiments using UV spectroscopy (295 nm) were conducted, focusing on G2m1 and G6m1. The melting temperature (T_m) was found to be independent of the RNA concentration over the concentration range used in the CD and fluorescence measurments (Figure S15), indicating that intramolecular GQS folding was being monitored during the K^+ ion titrations.

Herein, we examined the cooperativity of RNA folding by deliberately populating folding intermediates. A broad-range response was achieved with G5 and G6 GQS RNAs whereas only a narrow-range was acheived with G2 GQS RNAs, which do not populate folding intermediates.[12b] K^+ ion binding was monitored by both CD and fluorescence spectroscopy. Although K^+ ion sensing has been shown previously, this was done with attached fluorophores and produced a steep-

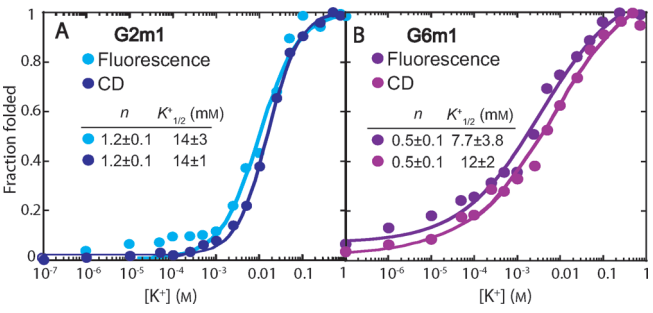

Figure 4. Comparison of K^+ ion titrations of GQSs monitored by either CD or fluorescence spectroscopy. CD detected at 262 nm and fluorescence emission detected at 360 nm. A) Graph of the fraction of G2m1 folded at different K^+ ion concentrations, fit to a two-state equation (Equation S1). B) Graph of the fraction of G6m1 folded at different K^+ ion concentrations, fit to a two-state equation (Equation S1). See Figures S13 and S14 for full fluorescence spectra and individual fits.

response curve. This is the first time that a broad response has been achieved by a single sequence rather than a mixture of sequences. While established approach of mixing biosensors can provide a broadened response range of three to four orders of magnitude,[6] this approach has severe limitations: multiple biosensors with well-controlled K_d values must be designed and prepared and accurate concentration determination and mixing of each probe is needed to ensure consistency in the response profile. Moreover, such sensors could interfere with each other through intermolecular base-pairing or by ion-induced aggregation.

As an alternative, we introduced the concept of intentionally populating folding and binding intermediates in a single RNA sequence as a means to broadening the linear response range. Extension of the G-stretch is a simple and robust strategy and should be applicable to DNA G-quadruplexes. It requires no mixing of multiple probes and achieves a similar or even better dynamic range of response to K^+ concentration. Moreover, it is much simpler to introduce single nucleic acid probes in vivo, rather than multiple probes. We clearly detected a folding intermediate for the G3 and G4 GQSs. Such an intermediate could include "slipped" GQS registers and/or through-space electrostatic repulsion of the adjacent K^+ ions. A similar phenomenon of weak association of the "later-binding ions" has been reported for K^+ ion binding in membrane ion channels.[18] The longer G-stretches of G5 and G6 sequences allow more G-quartets to form and therefore have more folding and binding intermediates. Broadening of the response with the number of quartets is most evident in a graph of apparent Hill coefficients versus number of quartets (Figure S16).

One unique feature of this study is the intrinsic fluorescence of the RNA GQSs. Although the fluorescence is somewhat weak, we found that it can be enhanced by altering the loop size, loop sequence, RNA topology, and ionic conditions (unpublished results). In addition, tandem GQSs could be introduced to multiply the fluorescence signal, as shown for related systems.[19] Coupling of GQS folding with ligand detection would then allow a unique read out of ligand binding, which circumvents the need for the introduction of synthetic fluorophores, such as FRET pairs or fluorophore-quencher pairs.

In summary, we have shown that the cooperativity of RNA folding can be tuned by adjusting the number of G-quartets present. Although the GQS biosensor developed in this study senses only K^+ ions, our approach to broad-response sensing is potentially applicable to other RNA or DNA aptamer systems[20] in which folding intermediates can be added to the system, for example through alternative folds or overlapping binding sites.

Received: August 10, 2012
Revised: September 9, 2012
Published online: November 20, 2012

Keywords: fluorescence · folding cooperativity · G-quadruplexes · RNA · RNA structures

[1] a) E. M. Moody, P. C. Bevilacqua, *J. Am. Chem. Soc.* **2003**, *125*, 16285 – 16293; b) E. M. Moody, J. C. Feerrar, P. C. Bevilacqua, *Biochemistry* **2004**, *43*, 7992 – 7998; c) B. D. Sattin, W. Zhao, K. Travers, S. Chu, D. Herschlag, *J. Am. Chem. Soc.* **2008**, *130*, 6085 – 6087; d) S. V. Solomatin, M. Greenfeld, D. Herschlag, *Nat. Struct. Mol. Biol.* **2011**, *18*, 732 – 734.

[2] M. Mandal, M. Lee, J. E. Barrick, Z. Weinberg, G. M. Emilsson, W. L. Ruzzo, R. R. Breaker, *Science* **2004**, *306*, 275 – 279.

[3] a) T. V. Erion, S. A. Strobel, *RNA* **2011**, *17*, 74 – 84; b) E. B. Butler, Y. Xiong, J. Wang, S. A. Strobel, *Chem. Biol.* **2011**, *18*, 293 – 298; c) J. Lipfert, A. Y. L. Sim, D. Herschlag, S. Doniach, *RNA* **2010**, *16*, 708 – 719.

[4] a) D. Koshland, A. Goldbeter, J. Stock, *Science* **1982**, *217*, 220 – 225; b) A. Vallee-Belisle, F. Ricci, K. W. Plaxco, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13802 – 13807.

[5] M. Cerminara, T. M. Desai, M. Sadqi, V. Munoz, *J. Am. Chem. Soc.* **2012**, *134*, 8010 – 8013.

[6] a) A. Vallée-Bélisle, F. Ricci, K. W. Plaxco, *J. Am. Chem. Soc.* **2012**, *134*, 2876 – 2879; b) J. Lin, D. Liu, *Anal. Chim. Acta* **2000**, *408*, 49 – 55; c) A. P. Drabovich, V. Okhonin, M. Berezovski, S. N. Krylov, *J. Am. Chem. Soc.* **2007**, *129*, 7260 – 7261.

[7] J. L. Huppert, *Biochimie* **2008**, *90*, 1140 – 1148.

[8] a) Y. Sannohe, H. Sugiyama, *Curr. Protoc. Nucleic Acid Chem.* **2010**, *Chapter 17*, Unit 17.12.11 – 17; b) K. Halder, J. S. Hartig, *Met. Ions Life Sci.* **2011**, *9*, 125 – 139.

[9] a) N. V. Hud, F. W. Smith, F. A. L. Anet, J. Feigon, *Biochemistry* **1996**, *35*, 15383 – 15390; b) M. A. Mullen, K. J. Olson, P. Dallaire, F. Major, S. M. Assmann, P. C. Bevilacqua, *Nucleic Acids Res.* **2010**, *38*, 8149 – 8163; c) A. Y. Q. Zhang, A. Bugaut, S. Balasubramanian, *Biochemistry* **2011**, *50*, 7251 – 7258; d) A. Bugaut, S. Balasubramanian, *Biochemistry* **2008**, *47*, 689 – 697.

[10] S. Takenaka, B. Juskowiak, *Anal. Sci.* **2011**, *27*, 1167 – 1167.

[11] a) K. Halder, M. Wieland, J. S. Hartig, *Nucleic Acids Res.* **2009**, *37*, 6811 – 6817; b) A. T. Phan, *FEBS J.* **2010**, *277*, 1107 – 1117; c) A. N. Lane, J. B. Chaires, R. D. Gray, J. O. Trent, *Nucleic Acids Res.* **2008**, *36*, 5482 – 5515.

[12] a) V. Víglaský, K. Tlučková, L. Bauer, *Eur. Biophys. J.* **2011**, *40*, 29 – 37; b) M. A. Mullen, S. M. Assmann, P. C. Bevilacqua, *J. Am. Chem. Soc.* **2012**, *134*, 812 – 815; c) M. Vorlíčková, I. Kejnovská, J. Sagi, D. Renčiuk, K. Bednářová, J. Motlová, J. Kypr, *Methods* **2012**, *57*, 64 – 75.

[13] a) M. A. Mendez, V. A. Szalai, *Biopolymers* **2009**, *91*, 841 – 850; b) N. T. Dao, R. Haselsberger, M.-E. Michel-Beyerle, A. T. Phan, *FEBS Lett.* **2011**, *585*, 3969 – 3977.

[14] The sequences chosen herein are either directly from a known gene, such as in the case of G2w2, G2m2, G2s2, and G4m1, or are closely related to these and representative of classes known to exist in genes.

[15] UA-rich loops led to Hill constants approaching unity, which may be because of the formation of intermediates involving base-pairing of the G residues with residues in the loops.

[16] a) S. A. Woodson, *Annu. Rev. Biophys.* **2010**, *39*, 61 – 77; b) T. R. Weikl, M. Palassini, K. A. Dill, *Protein Sci.* **2004**, *13*, 822 – 829.

[17] Notably, we previously reported folding intermediates in related G3 sequences, which were observed by CD spectroscopy of three-state titrations, the presence of intermediates on native gels, and RNase T1 protection.

[18] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, R. MacKinnon, *Science* **1998**, *280*, 69 – 77.

[19] D. Y. Vargas, K. Shah, M. Batish, M. Levandoski, S. Sinha, S. A. E. Marras, P. Schedl, S. Tyagi, *Cell* **2011**, *147*, 1054 – 1065.

[20] a) J. Liu, Z. Cao, Y. Lu, *Chem. Rev.* **2009**, *109*, 1948 – 1998; b) W. Mok, Y. Li, *Sensors* **2008**, *8*, 7050 – 7084.

