THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

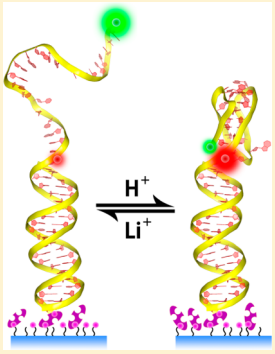
Destabilization of i-Motif by Submolar Concentrations of a Monovalent Cation

Sung Eun Kim,[†] Il-Buem Lee,[†] Changbong Hyeon,[‡] and Seok-Cheol Hong*[,†,‡]

[†]Department of Physics, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 136-713, Republic of Korea

[‡]School of Computational Sciences, Korea Institute for Advanced Study, 85 Hoegi-ro Dongdaemun-gu, Seoul 130-722, Republic of Korea

ABSTRACT: Counterions are crucial for self-assembly of nucleic acids. Submolar monovalent cations are generally deemed to stabilize various types of base pairs in nucleic acids such as Watson−Crick and Hoogsteen base pairs via screening of electrostatic repulsion. Besides monovalent cations, acidic pH is required for i-motif formation because protons facilitate pairing between cytosines. Here we report that Li^+ ions destabilize i-motif, whereas other monovalent cations, Na^+ and K^+, have the usual stabilizing effect. The thermodynamics data alone, however, cannot reveal which mechanism, enhanced unfolding or suppressed folding or both, is responsible for the Li^+-induced destabilization. To gain further insight, we examined the kinetics of i-motif. To deal with slow kinetics of i-motif, we developed a method dubbed HaRP to construct a long FRET time trace to observe a sufficient number of transitions. Our kinetics analysis shows clearly that Li^+ ions promote unfolding of i-motif but do not hinder its folding, lending strong support for our hypothesis on the origin of this unusual effect of Li^+. Although the subangstrom size of Li^+ ions allows them to infiltrate the space between cytosines in competition with protons, they cannot adequately fulfill the role of protons in mediating the hydrogen bonding of cytosine pairs.



INTRODUCTION

Monovalent cations are ubiquitous in biological systems.[1] They play critical roles in folding processes and stability of nucleic acids.[2,3] Since its first discovery in 1993,[4] i-motif has drawn much attention from chemists and structural biologists.[5] The i-motif is typically formed by four cytosine(C)-rich tracts intervened by loop-forming nucleotides, which results in a four-stranded folded structure with each strand at the vertex of its rectangular cross section (Scheme 1). The structure is maintained via intercalated hemiprotonated C−C^+ base pairs (or the bonds between N3 of the unprotonated cytosine and hydrogen at N3 of the protonated cytosine) between diagonally opposite DNA strands. A slightly acidic pH condition (pH = 6.0−6.5) facilitates protonation of cytosine. Demonstrated recently, the N3—H^+—N3 bonds can be described as hydrogen bonds with asymmetric double-well potentials.[6]

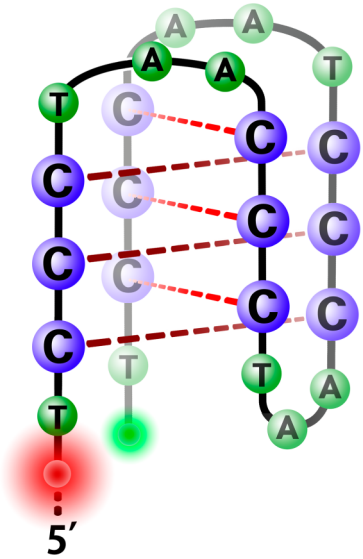
Biological significance of i-motif has been also suggested. Proteins have been identified to bind specifically to C-rich tracts in telomeric DNA.[7] A DNA sequence in the insulin-linked polymorphic region (ILPR) is eligible for i-motif formation, leading to replication inhibition.[8] Moreover, i-motif-induced repression of transcription has been described in the c-*myc* promoter region.[9] Recently, it was reported that stabilization of the i-motif structure in the *BCL2* promoter leads to significant up-regulation of the *BCL2* gene.[10,11] From the application viewpoint, i-motifs also hold promise for monitoring pH changes in vivo as demonstrated in recent studies in *Drosophila* cells[12] and in *C. elegans*.[13]

As the spatial dimension of base pairs in i-motif is in the range of a few Å, monovalent cations with different Å-scale sizes (i.e., Li^+, Na^+, and K^+) are expected to have distinct effects on the conformational dynamics of i-motif. Here we studied the effects of monovalent cations on the formation of i-motif. As a model system for i-motif, we used a C-rich human telomeric sequence.[14] The sequence has four triple repeats of cytosine (5′-$\underline{C}_3$TAA$\underline{C}_3$TAA$\underline{C}_3$TAA$\underline{C}_3$T-3′: see Materials and Methods for the full sequence of strands) that can accommodate up to six protons for complete folding (see Scheme 1).

To directly monitor and quantify the conformational transition between a single-stranded DNA coil and the folded i-motif structure, we employed the single-molecule FRET (fluorescence resonance energy transfer) technique.[15] This technique is ideal to study i-motif as the FRET efficiency between the donor and acceptor dyes flanking the C-rich sequence probes the formation of i-motif directly. The population of FRET values yields the thermodynamic information on the molecular states. We confirmed that acidic pH indeed facilitates i-motif formation. Interestingly, we discovered that increasing Li^+ concentration reduced the population of high-FRET states, whereas other cations (Na^+ and K^+) stabilized i-motif via screening as anticipated. By monitoring the conformation of each molecule in real time using a very long FRET time trace constructed by concatenating individual traces (see below), we could also gain the kinetic information on the transitions. From such

Received: January 6, 2014
Revised: April 7, 2014
Published: April 16, 2014

Scheme 1. Cartoon of the i-Motif by the Human Telomeric Sequence[a]



[a]Balls represent bases (C, T, and A for cytosine, thymine, and adenine, respectively) and dashed lines indicate the C−C$^+$ base pairs. Cytosines are exaggerated compared to other bases for clarity. Donor (Cy3) and acceptor (Cy5) dyes for FRET measurements are shown as green and red dots with halo, respectively. The folded i-motif state shown here yields a high-FRET signal (~0.9). The 5′-direction of the strand is shown next to Cy5.

measurements, we show clearly that Li^+ ions promote unfolding of i-motif but do not hinder folding of the structure.

MATERIALS AND METHODS

Materials. To make the i-motif construct, we purchased the following two strands from Integrated DNA Technologies (Coralville, Iowa, U.S.A.) and hybridized them: 5′-G G C G G C T G G C G A C G G C A G C G A G G C (Cy5) T C C C T A A C C C T A A C C C T A A C C C T (Cy3)-3′ and 5′-G C C T C G C T G C C G T C G C C A G C C G C C (Bio)-3′ where the underlined region is the telomeric sequence capable of i-motif formation. Cy3 and Cy5 are fluorescent dyes used for FRET; Bio denotes biotin, which enables surface immobilization.

Lithium compounds (LiCl, LiOH) and buffering reagent (MES) used in this study were purchased from Sigma-Aldrich. We prepared 50 mM MES buffers at three different pH values (pH 6.0, 6.2, and 6.4) supplemented with trolox, glucose, and gloxy as recommended for single-molecule fluorescence experiments.[15] We set the pH of each buffer by adding an appropriate amount of the base (LiOH, NaOH, or KOH) containing a cation of choice (Li^+, Na^+, or K^+). To prepare buffer solutions with different concentrations of the cation of choice, we continued to add the salts with the same cation (LiCl, NaCl, or KCl) until the total concentration of cation added up to designated values (30, 50, 100, 200, or 500 mM).

Single-Molecule FRET Experiment. To detect the folding and unfolding of the molecular construct in real time, we utilized the single-molecule FRET technique. The detail of the technique is given elsewhere.[15] To immobilize the C-rich molecules, we injected a buffer solution containing the C-rich DNA molecules at concentrations (~25 pM) much lower than those typically used in bulk assays (>10 nM) and washed the sample chamber with the same buffer solution after brief incubation. We acquired the FRET efficiency from a pair of dyes (donor and acceptor) on each of immobilized, properly dye-labeled molecules according to a typical FRET analysis scheme.[15]

Dependence of K_{eq} on the Concentration of Li^+. Considering the reactions described in Scheme 2, the equilibrium constants for the first reaction ($U + iH^+ + jLi^+ \leftrightarrow I$) and the second reaction ($I + lH^+ \leftrightarrow kLi^+ + F$) are given as $K_1 = [I]/([U][H^+]^i[Li^+]^j)$ and $K_2 = [F][Li^+]^k/([I][H^+]^l)$, respectively.

Experimentally, we measured the ratio of high-FRET and low-FRET states:

$$K_{eq} = \frac{[I] + [F]}{[U]} = K_1[H^+]^i[Li^+]^j\left(1 + \frac{K_2[H^+]^l}{[Li^+]^k}\right)$$

$$\ln K_{eq} = \ln K_1 + i\ln[H^+] + j\ln[Li^+] + \ln\left(1 + \frac{K_2[H^+]^l}{[Li^+]^k}\right)$$

The last term accounts for the fact that $\ln K_{eq}$ is a decreasing function of Li^+. By fitting the data in Figure 3b as shown in the text, we found $(K_2[H^+]^l/[Li^+]^k) \gg 1$, which further simplifies the above equation as

$$\ln K_{eq} = \ln(K_1K_2) + (i + l)\ln[H^+] + (j - k)\ln[Li^+] \equiv a + b\ln[H^+] + c\ln[Li^+] \quad (1)$$

Here, b and c indicate the net number of protons and lithium ions that are associated with the i-motif, respectively. Negative c implies that lithium ions destabilize the i-motif, and thus the net number of lithium ions participating in i-motif disruption is $-c$.

Kinetics Analysis: Handshaking Repeated Permutation (HaRP). We developed a method to build a very long FRET time trace from a large number of shorter traces, the duration of which was limited by dye lifetimes. We call this method Handshaking Repeated Permutation (HaRP), as we extract a trace from a pool of FRET traces a large number of times by *repeated permutation*, and we elongate the chain of traces by appending an extracted trace to the tail of the existing trace chain if their FRET states (as shown below, the FRET states of i-motif are binary and thus can be described by "handedness") are matched. This scheme is pictorially illustrated in Figure 1.

For the problem of i-motif, we only have two FRET states. Using Hidden Markov Modeling (HMM), we transform FRET time traces into binary (L or H) data arrays. Then, we classify each time trace into 4 groups according to the levels of its initial and final data points: H → H (high initial, high final: H_iH_f), H → L (H_iL_f), L → H (L_iH_f), and L → L (L_iL_f). Suppose that there are n independent time traces (in our case, n ~700), and we choose a first trace randomly. If the trace belongs to H_iH_f, for example, we then select the next trace from either H_iL_f or H_iH_f and append it to the previously selected trace. If the second trace belongs to H_iL_f (so far, we have H_iH_f -H_iL_f), we then select the third trace from either L_iL_f or L_iH_f. We can repeat this *handshaking* (wherein the binary handedness should be matched) process any number of times ($\gg n$). Once we have a very long trace, we can construct the dwell-time distributions and obtain the kinetic information.

dx.doi.org/10.1021/jp500120d | J. Phys. Chem. B 2014, 118, 4753−4760

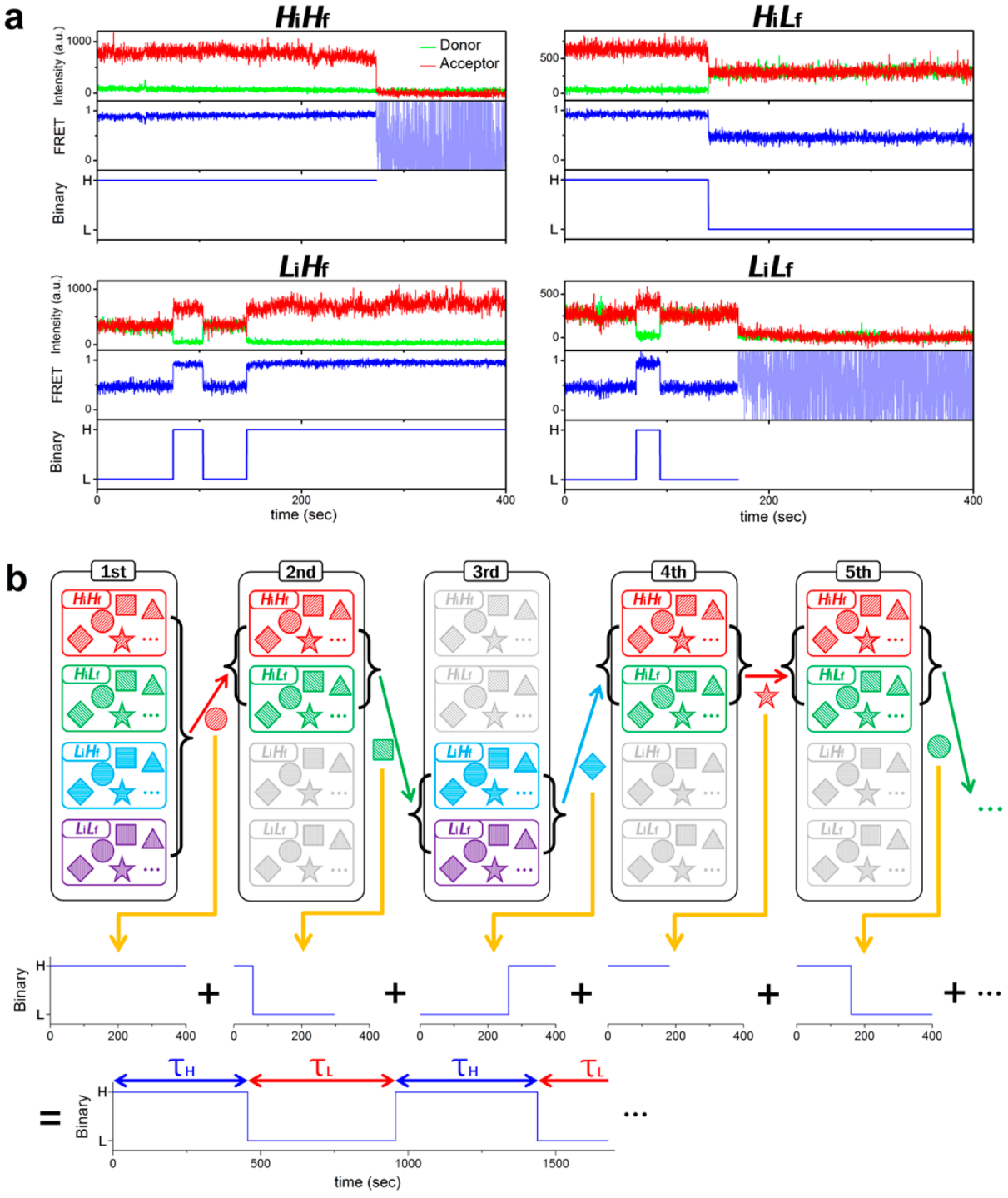


Figure 1. (a) Exemplary donor and acceptor fluorescence intensity time traces (top), FRET time traces (middle), and binary idealization of time traces by the HMM analysis (bottom) for four different groups of traces (H_iH_f, H_iL_f, L_iH_f, L_iL_f) as indicated. The acquisition time was typically 400 s and due to photobleaching of dyes the actual duration of a trace was often shorter than 400 s. Dim lines in FRET time traces indicate traces after photobleaching. (b) Schematics of the HaRP method. Each box corresponds to a round of selection. In each box, red/left-hatched, green/right-hatched, blue/horizontal-hatched, and purple/vertical-hatched symbols stand for H_iH_f, H_iL_f, L_iH_f and L_iL_f traces, respectively. Different symbols denote different traces in the same group. Curly brackets show pools to be considered for selection. Inactivated pools (not to be considered for the given round of selection) are displayed gray. For the first round of selection, we choose any trace randomly from all four pools. If a H_iH_f trace is selected (red/left-hatched circle after the 1st box), a trace is selected among the traces in H_iH_f and H_iL_f groups in the 2nd round. Selected (binary) traces are also shown below boxes. This process is continued for a large number of times (for the case of i-motif, it was 30 000 times). These binary traces are concatenated to form a very long trace. This trace can be used to yield the distribution of dwell times of high- and low-FRET states (see Figure 3).

dx.doi.org/10.1021/jp500120d | J. Phys. Chem. B 2014, 118, 4753−4760

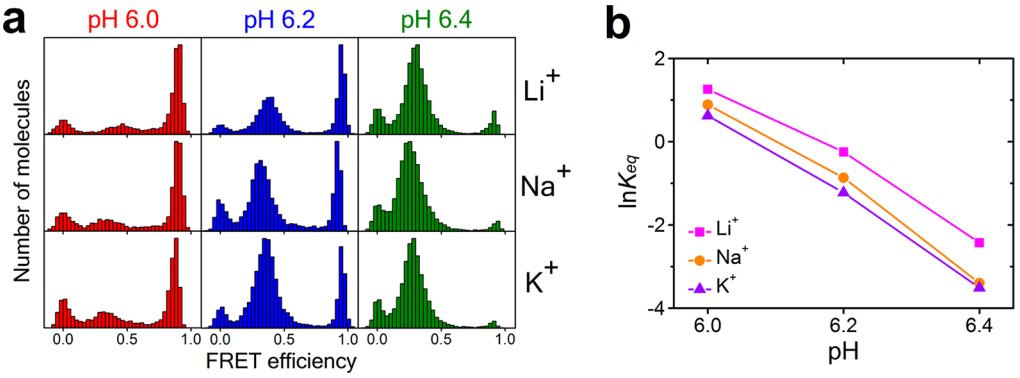


Figure 2. pH dependence of i-motif formation (a) Single-molecule FRET histograms from the C-rich human telomeric sequence under the conditions (pH values and types of cation (concentration: 50 mM)) specified in the figure. The FRET efficiency peaks at about 0.9, 0.3, and 0.0 correspond to folded (i-motif), unfolded states, and molecules with donor dye only, respectively. (b) pH-vs-lnK_{eq} relation for the i-motif DNA molecule drawn from the data in Figure 2a. Here, K_{eq} is the ratio of high-FRET to low-FRET populations, calculated by evaluating the area under the histogram. This graph emphasizes the linear relation between pH and ln K_{eq}.

RESULTS

pH Dependence of i-Motif Formation. The FRET histograms (Figure 2) show that the populations of high-FRET (~0.9) and low-FRET (~0.3) states are sensitively modulated over a narrow range of pH values. The C-rich single-stranded DNA is mainly folded (in view of high-FRET population) at pH = 6.0, whereas the same DNA molecule remains unfolded (in view of low-FRET population) at pH = 6.4. This was recently addressed at the single-molecule level.[16,17] Furthermore, as expected from the electrostatic screening effect, increasing the concentration of Na^+ or K^+ tipped the equilibrium toward the folded state, or at least, the change in salt concentration was insignificant to the transition of i-motif (Figure 3a).

Suppression of i-Motif Formation by Lithium Ions. Increasing the concentration of Li^+, however, suppressed the formation of i-motif significantly. As shown in Figure 3, the decrease in the equilibrium constant, K_{eq}, or destabilization of i-motif by Li^+ was found at all three pH values tested (6.0, 6.2, and 6.4) (Figure 3b). This effect is unique to Li^+. The destabilization observed here might be associated with the chaotropic nature of some lithium salts such as $LiClO_4$;[18] however, it should be noted that the concentration of Li^+ that induces i-motif unfolding is much lower than a typical Li^+ concentration (greater than a few molar) causing duplex denaturation. Furthermore, unlike $LiClO_4$, which has a salting-in (or denaturant-like) effect, LiCl used in our experiment is known to act as a salting-out agent. Thus, i-motif destabilization due to submolar Li^+ (or LiCl) levels is not straightforwardly explained from the perspective of chaotrope.

Kinetics Analysis of i-Motif Transition via the HaRP Method. To examine the effect of Li^+ on i-motif formation from the kinetics viewpoint, we carried out a dwell-time analysis on the FRET time traces. Although the single-molecule FRET technique is known to be suitable for acquiring kinetic information about conformational transitions, it has a limitation in the case of i-motif, because the transition is very slow and the measurement time is restricted by the finite lifetime of the fluorescent dye molecules. Over a typical measurement period of 400 s, we usually observed no transitions or just one or two transitions between the two states.[19] Thus, it is hard to obtain reliable statistics for the analysis. To this end, assuming ergodicity or time translation invariance, we developed a method dubbed HaRP and used it to build a very long time trace. The validity of this method can be understood as follows: If the number of traces is large enough and the traces are random (the moment of photobleaching of fluorescent dyes is also stochastic) and show no long-term evolution, the populations of initial-high (or final-high) and initial-low (or final-low) samples would represent the true probabilities of their being in high- or low-FRET states at any given moment during measurement. Traces in the pool are considered as mutually independent and randomly chosen excerpts from a very long time trace of a representative i-motif molecule. If one trace is ended with a low FRET, another trace beginning with a low FRET from the pool is selected and appended after the former, as the probability of having a low FRET at any moment is the same as the probability of finding initial-low samples. Thus, the selection is justified. By concatenating random ensemble traces, we can construct a very long single time trace.

In order to validate the method, we simulated the HaRP procedure with short FRET time traces acquired from duplex–triplex transition of DNA.[3] Typical dwell times for folding and unfolding transitions of triplex are of the order of 1 s, and time traces were collected for ~200 s. Thus, we could get reliable statistics for the dwell times from tens of independent traces. To test the validity of the method, we chopped ~20 traces (duration ~200 s each) into traces (hereafter called t-chips) with much shorter duration (hereafter called division time δt) and carried out HaRP for t-chips to construct a very long time trace. The length of t-chips was varied from δt = 30 s down to 0.3 s, the latter of which is just the three times the interval of data acquisition (0.1 s). As shown in Figure 4a-c, we plotted the survival probabilities of the high-FRET state of DNA triplex from the original traces and HaRP-generated traces. The survival probabilities exhibit double-exponential kinetics, $S(t) = \varphi_1 e^{-t/\tau_1} + \varphi_2 e^{-t/\tau_2}$. From this, the time constants of the transition (the average dwell time) can be determined ($\langle\tau\rangle = \int_0^\infty S(t)dt = \sum_i \varphi_i \tau_i$). Not surprisingly, when division time gets comparable to the interval of data acquisition, $\langle\tau\rangle$ obtained from HaRP deviated significantly from the one obtained from the original traces. This is because our HaRP scheme creates a bias in linking time traces, suppressing the frequency of transitions at the boundary, which leads to overestimation of the dwell times. As a figure of merit to judge the fidelity of linkage, we define R, linking error rate (see Appendix). For data in Figure 4, R varies from 0.0033 for δt = 30 s to 0.33 for δt =

dx.doi.org/10.1021/jp500120d | J. Phys. Chem. B 2014, 118, 4753–4760

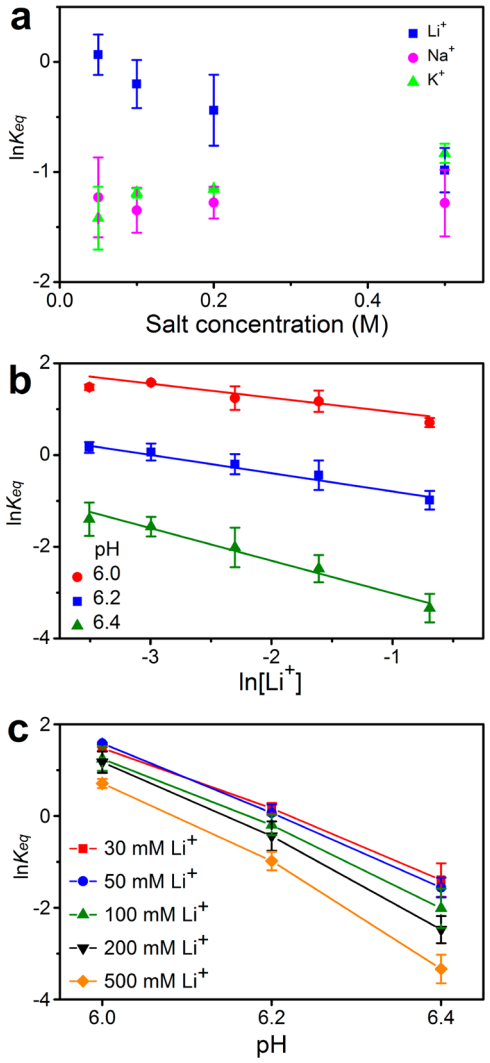


Figure 3. (a) ln K_{eq} vs the concentration of various monovalent cations at pH = 6.2. (b) ln K_{eq} vs ln[Li^+] at three different pH values. [Li^+] is measured relative to the 1 M standard state. Lines represent fits using eq 1: c = −0.31, −0.40, and −0.71 for pH = 6.0, 6.2, and 6.4, respectively. (c) ln K_{eq} vs pH for the i-motif DNA molecule in the presence of various concentrations of Li^+. Redrawn from the data in Figure 3b, this graph reveals the linear relation between pH and ln K_{eq}. Error bars represent the standard deviations of each data point.

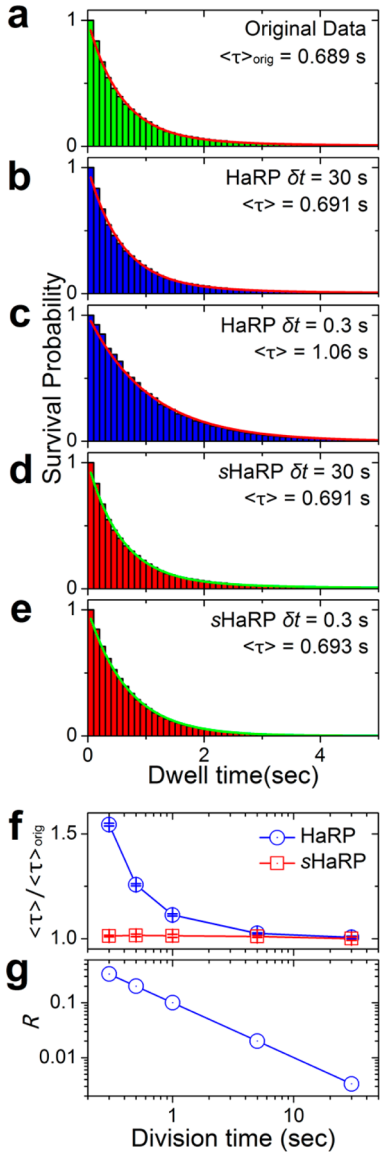


Figure 4. (a−e) Survival probabilities of the high-FRET state of DNA triplex from original traces (a), HaRP-generated long trace with δt = 30 s (b) and 0.3 s (c), and *s*HaRP-generated long trace with δt = 30 s (d) and 0.3 s (e). Average dwell time, <τ>, is also shown in each panel. <τ>$_{orig}$ is the dwell time obtained from original traces (Figure 4a). The survival probabilities are fitted using a double-exponential function. (f) Normalized average dwell times (with respect to <τ>$_{orig}$) for various division times. The total numbers (Q) of traces used in (*s*)HaRP are 1, 1, 2, 5, and 10 (× 10^4) for δt = 30, 5, 1, 0.5, and 0.3 s, respectively. <τ> by HaRP deviates from <τ>$_{orig}$ for small division times, whereas <τ> by *s*HaRP remains in good agreement with <τ>$_{orig}$ over the entire range of division time. Error bars represent the standard deviations of each data point obtained by five runs of (*s*)HaRP. (g) Linking error rate (R) vs division time.

0.3 s (Figure 4g). Evidently, a small R (<0.02) guarantees the validity and reliability of the HaRP scheme, although a large R (>0.1) correlates with a large deviation in <τ>. As a rule of thumb for this method, a 50% overestimate of τ at R = 0.33 sets a rough upper bound of $\langle\tau\rangle$.

In order to show that the overestimation of <τ> indeed arises from linking error, we applied the HaRP method in compliance with the knowledge of what level of data point should follow the last point of each t-chip. We call this method *sample* HaRP (*s*HaRP). *s*HaRP is devised by the idea that as long as we avoid making a mistake in linking two consecutive t-chips, we can minimize the error rate, R, and thus minimize the discrepancy in the value of <τ>. In *s*HaRP, the final data point of each t-chip is classified into four kinds: H_fh, High and remain at high in the next data point; L_fl, Low and remain at low in the next data point; L_fh, Low and rise to high in the next data point; and H_fl, High and fall to low in the next data point. T-chips are then categorized into eight groups (L_iL_fl, L_iL_fh, L_iH_fl, L_iH_fh, H_iL_fl, H_iL_fh, H_iH_fl, and H_iH_fh). In this scheme, t-chips such as L_iL_fh, L_iH_fh, H_iL_fh, and H_iH_fh can be linked to "high-initial" t-chips and the rest can be linked to "low-initial" t-chips. When we applied *s*HaRP on the triplex data, we obtained the same survival probabilities and <τ> as those obtained from direct kinetics analysis of original traces (Figure 4d−f). This asserts that low linking error is critical for HaRP (Figure 4g). Parenthetically speaking, *s*HaRP could be practically useful for constructing a long time trace in case only a limited number

dx.doi.org/10.1021/jp500120d | J. Phys. Chem. B 2014, 118, 4753−4760

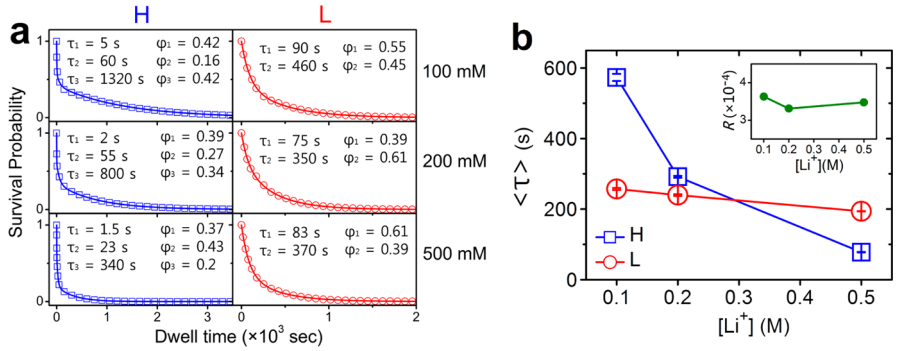


Figure 5. (a) Survival probabilities of the high-FRET (H) and low-FRET (L) states obtained by HaRP are presented in left and right columns, respectively, for [Li^+] = 100, 200, and 500 mM (top to bottom). The total number of traces used in HaRP is 30 000. Survival probabilities of H (blue square) and L (red circle) states are fitted using triple- and double-exponential functions (solid lines), respectively. Li^+ concentrations and relevant time constants (τ_i: larger i for longer-lived ensemble) and normalized populations of each ensemble (φ_i;$\sum_i\varphi_i = 1$) are shown in each graph. (b) The average dwell time ($\langle\tau\rangle$) of H decreased with Li^+ concentration while that of L remained unchanged. Error bars represent the standard deviations of each data point obtained by three runs of HaRP. Inset: linking error rate, R, for i-motif.

of time traces are available and each of the traces is long enough to exhibit many transitions. Taken together, the application of HaRP (Figure 5a,b) to our i-motif data is valid and not subject to errors in kinetics, because R for i-motif is very small (R < 0.0004) in our measurements (see inset in Figure 5b).

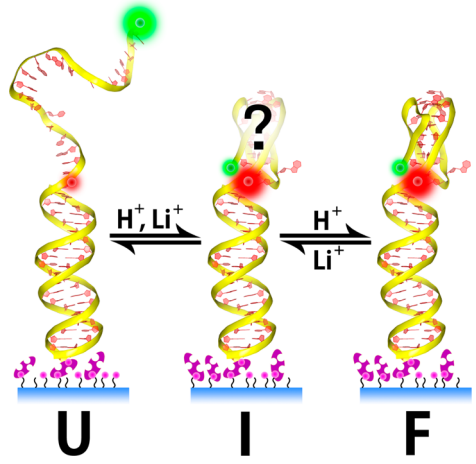
Using the long FRET trace of i-motif constructed by HaRP, we plotted survival probabilities $S(t)$ of the high-FRET and low-FRET states for three different Li^+ concentrations (100, 200, and 500 mM) and obtained the time constants for the dwell times of those states (Figure 5a).

The survival probabilities of both high-FRET and low-FRET states exhibit multiexponential kinetics, which suggests that folding and unfolding of i-motif, starting from a heterogeneous structural ensemble,[16,20,21] occur via a kinetic partitioning mechanism on a rugged folding landscape.[22] The survival probabilities for the high-FRET and low-FRET states were fitted using triple- and double-exponential functions ($S(t) = \sum_i\varphi_i e^{-t/\tau_i}$ with $\sum_i\varphi_i = 1$), respectively. With increasing Li^+ concentration, the time constants of unfolding (or the average dwell time of the high-FRET states) decreased (among τ_i, τ_3 changed most clearly), whereas the time constants for folding (or the average dwell time of the low-FRET states) remained nearly unchanged. This lends strong support to the idea that Li^+ destabilizes the folded state while having little effect on the folding process. The single-molecule approach used here was pivotal in acquiring previously inaccessible kinetic information and in unveiling the role of Li^+ in i-motif formation.

Thermodynamics of i-Motif Transition and a Three-State Destabilization Model. Given that i-motif is formed in acidic conditions that promote the protonation of cytosine, it could be argued that Li^+ interferes with protonation of cytosine. It is tempting to consider an inhibitor-based model where Li^+ competes with proton and inhibits folding of i-motif. In the inhibitor-based model, high Li^+ concentration sequesters the molecule to a hypothetical third state that holds an unfolded conformation. As a consequence, the folding rate should decrease with increasing Li^+ concentration. The kinetic measurement described above however rules out this model. Instead, we propose a minimal three-state destabilization model for Li^+-dependent dynamics of i-motif (Scheme 2).

In Scheme 2, the states U and F are the ensembles of unfolded (low-FRET) and folded (high-FRET) states, respectively, and the state I denotes an ensemble of short-lived intermediates that is conceived as displaying the same

Scheme 2. Three-State Destabilization Model of i-Motif along with Detection of i-Motif Formation via Single-Molecule FRET[a]



[a]The model presents the dual role of lithium ions in i-motif formation. F and I represent the ensembles of folded i-motif states and putative (thus, marked with "?") high-FRET intermediate states, respectively. U represents an unfolded ensemble. Green and red dots represent donor and acceptor dyes.

high-FRET value, thus indistinguishable from the state F. Electrostatic screening by Li^+ favors compact, high-FRET conformations (I and F). On the other hand, Li^+ can infiltrate the space between two cytosines and disrupts hydrogen bonding between the bases in the state of F by acting as an incomplete surrogate of proton. As a result from the model depicted in Scheme 2, the unfolding rate of F, thus the populations of I and U, would be increased with an increasing Li^+ concentration while the folding rate need not be affected. This is consistent with the kinetic measurement described above. Applying the law of mass action to the model, we obtained eq 1. From the slope of the graphs in Figure 3c, we found that the average number of protonation b was ≈3.6. Other independent set of data gave a similar value. Note that 3.6 is less than the maximum number of protonation sites, 6. By fitting the data in Figure 3b, we found that the net number of lithium ions participating in i-motif disruption ($-c$) varies from

dx.doi.org/10.1021/jp500120d | J. Phys. Chem. B 2014, 118, 4753−4760

0.31 (pH = 6.0) to 0.71 (pH = 6.4). It is of note that at high pH more lithium ions are associated with i-motif disruption.

DISCUSSION

We discovered that Li^+ ions reduce the population of i-motif in contrast with other cations (Na^+ and K^+) and that Li^+ ions promote unfolding of i-motif, not hindering folding of i-motif. The latter observation was made possible by the kinetic analysis method developed in this work. What is the origin of the difference between Li^+ and other monovalent cations? Most likely, it is ion size. The radius of Li^+ is around 0.7 Å.[23] In contrast, the radii of Na^+ and K^+ are larger than 1 Å. From the NMR data, the distance between the two paired cytosines in i-motif is about 1.8 Å.[24] Thus, one Li^+ ion can fit in the pocket between the two cytosines, but it cannot adequately reproduce the quantum mechanical effects of a proton on the hydrogen bonding;[6] this leads to disruption of cytosine pairs.

Ion size makes a critical contribution to the function and conformational dynamics of biological molecules. Ion channels are a well-known example that shows the molecular mechanism of size selectivity, operating with nanometer precision. More recently, the size of divalent cations was shown to be a key determinant of the stability and foldability of ribozymes.[25,26] In the nanoscopic world, a minor difference in ion size can lead to a notable difference.

The structure of the state I is not clear yet except that the state exhibits the same high-FRET value as F. A recent work reported an intriguing piece of new information that three C-rich tracts could form a folded structure linchpinned by a stack of hemiprotonated cytosine pairs between two terminal cytosine tracts.[27] This folded structure is evolved to more stable intermolecular i-motif via association with a single C-rich tract. This suggests that the state I could be related to a partly folded conformation similar to the linchpinned structure. This partly folded conformation is expected to yield a high-FRET value with less structural stability. Or alternatively, for molecules with four C-rich tracts, the linchpinned conformation could serve as a transient state (possibly, folded state with dwell time <5 s in Figure 5a), branching into two i-motif isomers with different dwell times, namely, 5′E and 3′E.[21] In the linchpinned conformation, only two C-rich tracts are locked and thus it can lead to either 5′E or 3′E. In order to clarify this issue, high-resolution structural and kinetic study should be conducted in future.

The physiological concentration of Li^+ is very low (~μM) compared to that of Na^+ and K^+ (~100 mM). Evolutionarily, this is perhaps because Na^+ and K^+ better facilitate the formation of folded structures or normal biomolecule functioning. Although a low Li^+ environment facilitates formation of G-quadruplex[28] and i-motif, it is also possible to impede their formation by simply perturbing the surrounding ionic balance. Intriguingly, a number of lithium salts are used at high doses as mood-stabilizing drugs, primarily for the treatment of bipolar disorder;[29] Li^+ may alter the gene expression of affected individuals.

The dependence of i-motif's foldability and kinetics on pH and concentration of monovalent cation, revealed in this study, is of potential use to improve i-motif-based pH sensors by allowing one to tailor the dynamic range of the i-motif population in each conformational state. Our study sheds lights on the microscopic understanding for how lithium-based drugs act on the conformational dynamics of DNA.

APPENDIX

Definition and Calculation of *R*

Here we define linking error rate (R), which represents the frequency of linking errors by the HaRP scheme, as the ratio of the number of potential linking errors (E) to the total number of transition events (T). We also define N as the total number of data points: $N = N_L + N_H$ where N_L and N_H are the numbers of data points exhibiting states L and H, respectively.

We shall consider two different situations: in one case, we have a sufficiently long trace available so that we know the transition probabilities (p_L and p_H) as well as the proportions ($\omega_L = N_L/N$ and $\omega_H = N_H/N = 1 - \omega_L$) of state L and state H. $p_{H\ (or\ L)}$ is defined as the number of transition events from H to L (or from L to H) divided by the number of data points at state H (or L): $p_H = T_{H\to L}/N_H = T/2N_H$ and $p_L = T_{L\to H}/N_L = T/2N_L$ where $T = T_{L\to H} + T_{H\to L} = 2T_{L\to H} = 2T_{H\to L}$ and $T_{i\to j}$ is the number of transition events from state i to j; if the initial and final states of a trace are the same, $T_{L\to H}$ and $T_{H\to L}$ are the same, but otherwise, they just differ by 1. In the other case, when each trace is too short, we do not know the transition probabilities beforehand due to lack of information on transition events and need to determine their intrinsic values from ensemble of many time traces.

Let us consider the former case first. For simplicity, we assume that every trace has the same duration δt. (This is true for our triplex test, as we chopped long traces into t-chips with equal division time δt. In real samples, for which we get FRET traces until a dye is photo-bleached, the length of a trace varies from trace to trace.)

First, let us find the expression for the number of potential linking errors. If the total number of traces is Q and the numbers of traces ending with states H and L are Q_H and Q_L ($Q_H + Q_L = Q$), respectively, E is given as follows: $E = Q_H p_H + Q_L p_L$. R becomes

$$R = \frac{E}{T} = \frac{Q_H p_H + Q_L p_L}{T} = \frac{Q_H T/2N_H + Q_L T/2N_L}{T} = \frac{\omega_H Q}{2N_H} + \frac{\omega_L Q}{2N_L} = \frac{Q}{N} = \frac{Q}{r\sum_{i=1}^{Q}\delta t_i}$$

where δt_i is the duration of the i^{th} trace in a HaRP-generated long trace and r is the repetition rate (10 Hz: reciprocal of the interval of data acquisition). If the length of each trace is all the same as δt, R reduces to $1/r\delta t$. For triplex data with δt from 0.3 s to 30 s, R varies from 1/3 to 1/300. From the result shown in main text, HaRP carried out with $R = 1/50$ ($\delta t = 5$ s) or smaller yields the same $\langle\tau\rangle$ as $\langle\tau\rangle_{orig}$.

Next, we consider the case the intrinsic transition rate is not known in advance. For i-motif data, R (= Q/N) is very small (< 0.0004), as shown in Figure 5b (inset). Thus, the number of linking errors is negligible and therefore $\langle\tau\rangle$ from HaRP should properly represent the true intrinsic average lifetime.

AUTHOR INFORMATION

Corresponding Author

*Email: hongsc@korea.ac.kr.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea

(NRF) funded by the Ministry of Science, ICT & Future Planning (2010-00-10594 and 2012R1A1A2021736).

REFERENCES

(1) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*; Garland Science: New York, 2002.
(2) Nakano, S.-i.; Fujimoto, M.; Hara, H.; Sugimoto, N. Nucleic Acid Duplex Stability: Influence of Base Composition on Cation Effects. *Nucleic Acids Res.* **1999**, *27*, 2957−2965.
(3) Lee, I.-B.; Hong, S.-C.; Lee, N.-K.; Johner, A. Kinetics of the Triplex-Duplex Transition in DNA. *Biophys. J.* **2012**, *103*, 2492−2501.
(4) Gehring, K.; Leroy, J.; Guéron, M. A Tetrameric DNA Structure with Protonated Cytosine. Cytosine Base Pairs. *Nature* **1993**, *363*, 561−565.
(5) Guéron, M.; Leroy, J.-L. The i-Motif in Nucleic Acids. *Curr. Opin. Struct. Biol.* **2000**, *10*, 326−331.
(6) Lieblein, A. L.; Krämer, M.; Dreuw, A.; Fürtig, B.; Schwalbe, H. The Nature of Hydrogen Bonds in Cytidine···H^+···Cytidine DNA Base Pairs. *Angew. Chem., Int. Ed.* **2012**, *51*, 4067−4070.
(7) Marsich, E.; Piccini, A.; Xodo, L. E.; Manzini, G. Evidence for a Hela Nuclear Protein That Binds Specifically to the Single-Stranded d(CCCTAA)n Telomeric Motif. *Nucleic Acids Res.* **1996**, *24*, 4029−4033.
(8) Catasti, P.; Chen, X.; Deaven, L. L.; Moyzis, R. K.; Bradbury, E. M.; Gupta, G. Cystosine-Rich Strands of the Insulin Minisatellite Adopt Hairpins with Intercalated $Cytosine^{+}$· Cytosine Pairs. *J. Mol. Biol.* **1997**, *272*, 369−382.
(9) Dai, J.; Hatzakis, E.; Hurley, L. H.; Yang, D. I-Motif Structures Formed in the Human c-MYC Promoter Are Highly Dynamic−Insights into Sequence Redundancy and I-Motif Stability. *PLoS One* **2010**, *5*, e11647.
(10) Kendrick, S.; Kang, H.-J.; Alam, M. P.; Madathil, M. M.; Agrawal, P.; Gokhale, V.; Yang, D.; Hecht, S. M.; Hurley, L. H. The Dynamic Character of the BCL2 Promoter i-Motif Provides a Mechanism for Modulation of Gene Expression by Compounds That Bind Selectively to the Alternative DNA Hairpin Structure. *J. Am. Chem. Soc.* **2014**, *136*, 4161−4171.
(11) Cui, Y.; Koirala, D.; Kang, H.; Dhakal, S.; Yangyuoru, P.; Hurley, L. H.; Mao, H. Molecular Population Dynamics of DNA Structures in a Bcl-2 Promoter Sequence Is Regulated by Small Molecules and the Transcription Factor hnRNP LL. *Nucleic Acids Res.* **2014**, DOI: 10.1093/nar/gku185.
(12) Modi, S.; Swetha, M.; Goswami, D.; Gupta, G. D.; Mayor, S.; Krishnan, Y. A DNA Nanomachine That Maps Spatial and Temporal pH Changes inside Living Cells. *Nat. Nanotechnol.* **2009**, *4*, 325−330.
(13) Surana, S.; Bhat, J. M.; Koushika, S. P.; Krishnan, Y. An Autonomous DNA Nanomachine Maps Spatiotemporal pH Changes in a Multicellular Living Organism. *Nat. Commun.* **2011**, *2*, 340−346.
(14) Leroy, J.-L.; Guéron, M.; Mergny, J.-L.; Hélène, C. Intramolecular Folding of a Fragment of the Cytosine-Rich Strand of Telomeric DNA into an i-Motif. *Nucleic Acids Res.* **1994**, *22*, 1600−1606.
(15) Selvin, P. R.; Ha, T. *Single-Molecule Techniques*; Cold Spring Harbor Laboratory Press: New York, U.S., 2008.
(16) Dhakal, S.; Schonhoft, J. D.; Koirala, D.; Yu, Z.; Basu, S.; Mao, H. Coexistence of an ILPR i-Motif and a Partially Folded Structure with Comparable Mechanical Stability Revealed at the Single-Molecule Level. *J. Am. Chem. Soc.* **2010**, *132*, 8991−8997.
(17) Choi, J.; Kim, S.; Tachikawa, T.; Fujitsuka, M.; Majima, T. pH-Induced Intramolecular Folding Dynamics of i-Motif DNA. *J. Am. Chem. Soc.* **2011**, *133*, 16146−16153.
(18) Breslow, R.; Guo, T. Surface Tension Measurements Show That Chaotropic Salting-in Denaturants Are Not Just Water-Structure Breakers. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 167−169.
(19) Zhao, Y.; Zeng, Z. x.; Kan, Z. y.; Hao, Y. h.; Tan, Z. The Folding and Unfolding Kinetics of the i-Motif Structure Formed by the C-Rich Strand of Human Telomere DNA. *ChemBioChem* **2005**, *6*, 1957−1960.
(20) Jin, K. S.; Shin, S. R.; Ahn, B.; Rho, Y.; Kim, S. J.; Ree, M. pH-Dependent Structures of an i-Motif DNA in Solution. *J. Phys. Chem. B* **2009**, *113*, 1852−1856.
(21) Lieblein, A. L.; Buck, J.; Schlepckow, K.; Fürtig, B.; Schwalbe, H. Time-Resolved NMR Spectroscopic Studies of DNA i-Motif Folding Reveal Kinetic Partitioning. *Angew. Chem. Int. Ed.* **2012**, *51*, 250−253.
(22) Thirumalai, D.; Hyeon, C. RNA and Protein Folding: Common Themes and Variations. *Biochemistry* **2005**, *44*, 4957−4970.
(23) Marcus, Y. Ionic Radii in Aqueous Solutions. *Chem. Rev.* **1988**, *88*, 1475−1498.
(24) Phan, A. T.; Guéron, M.; Leroy, J.-L. The Solution Structure and Internal Motions of a Fragment of the Cytidine-Rich Strand of the Human Telomere. *J. Mol. Biol.* **2000**, *299*, 123−144.
(25) Koculi, E.; Hyeon, C.; Thirumalai, D.; Woodson, S. A. Charge Density of Divalent Metal Cations Determines RNA Stability. *J. Am. Chem. Soc.* **2007**, *129*, 2676−2682.
(26) Moghaddam, S.; Caliskan, G.; Chauhan, S.; Hyeon, C.; Briber, R.; Thirumalai, D.; Woodson, S. A. Metal Ion Dependence of Cooperative Collapse Transitions in RNA. *J. Mol. Biol.* **2009**, *393*, 753−764.
(27) Dhakal, S.; Lafontaine, J. L.; Yu, Z.; Koirala, D.; Mao, H. Intramolecular Folding in Human ILPR Fragment with Three C-Rich Repeats. *PLoS One* **2012**, *7*, e39271.
(28) Risitano, A.; Fox, K. R. Influence of Loop Size on the Stability of Intramolecular DNA Quadruplexes. *Nucleic Acids Res.* **2004**, *32*, 2598−2606.
(29) Halford, B.; BOSTON, C. Limits of Lithium. *Chem. Eng. News* **2013**, *91*, 15−20.