JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

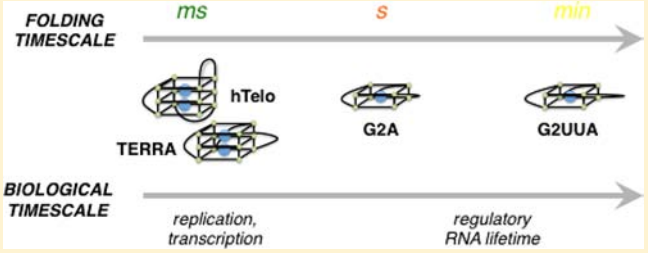
The Kinetics and Folding Pathways of Intramolecular G-Quadruplex Nucleic Acids

Amy Y. Q. Zhang[‡] and Shankar Balasubramanian*[,‡,§,∥]

[‡]Department of Chemistry, University of Cambridge, Lensfield Rd, Cambridge CB2 1EW, United Kingdom
[§]Cancer Research U.K., Cambridge Research Institute, Li Ka Shing Center, Cambridge, CB2 0RE, United Kingdom
[∥]School of Clinical Medicine, University of Cambridge, Cambridge, CB2 0SP, United Kingdom

Supporting Information

ABSTRACT: The folding kinetics of G-quadruplex forming sequences is critical to their capacity to influence biological function. While G-quadruplex structure and stability have been relatively well studied, little is known about the kinetics of their folding. We employed a stopped-flow mixing technique to systematically investigate the potassium-dependent folding kinetics of telomeric RNA and DNA G-quadruplexes and RNA G-quadruplexes containing only two G-quartets formed from sequences r[$(GGA)_3GG$] and r[$(GGUUA)_3GG$]. Our findings suggest a folding mechanism that involves two kinetic steps with initial binding of a single K^+, irrespective of the number of G-quartets involved or whether the G-quadruplex is formed from RNA or DNA. The folding rates for telomeric RNA and DNA G-quadruplexes are comparable at near physiological [K^+] (90 mM) (τ = ~60 ms). The folding of a 2-quartet RNA G-quadruplex with single nucleotide A loops is considerably slower (τ = ~700 ms), and we found that the time required to fold a UUA looped variant (τ > 100 s, 500 mM K^+) exceeds the lifetimes of some regulatory RNAs. We discuss the implications of these findings with respect to the fundamental properties of G-quadruplexes and their potential functions in biology.



INTRODUCTION

G-quadruplex forming sequences have been identified throughout both the genome[1,2] and transcriptome,[3,4] and various studies have been conducted to elucidate the biological functions of the G-quadruplex structures formed from these sequences. The investigations have collectively led to the description of G-quadruplexes as structural motifs that perform a wide range of putative physiological functions. They have been ascribed roles in the regulation of gene expression,[5] DNA recombination,[6] and splicing[7] through to neurite localization signaling[8] and mitochondrial transcription termination.[9] The effective use of small molecule tools to both identify cellular G-quadruplexes[10] and modulate their function[11] have further contributed to our understanding of the biology of G-quadruplexes.

The human telomeres, which are involved in maintaining genome stability and cell growth by protecting chromosome ends,[12] offer one of the best characterized and validated examples of G-quadruplex involvement in a cellular process.[13] The recent discovery that telomeric DNA is transcriptionally active has further emphasized the importance of G-rich telomeric sequences. The r[$(GGGUUA)_n$] repeat transcripts localize at the telomere[14] and inhibit telomerase[15] and also have a propensity to form RNA G-quadruplex structures.[16,17]

The biophysical characterization of G-quadruplexes is requisite for underpinning our understanding of their biological relevance. Such measurements can be considered in relation to the physical properties of well-understood cellular processes that G-quadruplexes might partake in or compete with. While the thermodynamic and structural properties of G-quadruplexes have been well studied, there is considerably more to be understood about the kinetic aspects of their folding and unfolding. In particular, determining a kinetic mechanism for G-quadruplex folding can answer critical questions about whether the time scale of folding is biologically relevant and about what molecular events take place along the folding pathway.

Investigations into the cation-driven kinetics of intramolecular G-quadruplex folding have to date focused on DNA. While telomeric DNA G-quadruplexes fold on the millisecond to second time scale,[18,19] the observed folding rates and number of folding phases appear to depend on whether the quadruplex-forming sequence has flanking nucleotides, the cation and buffer conditions, and the probe that is used to measure folding. Nonetheless, the observed multiphasic kinetics has suggested that folding of DNA G-quadruplexes is likely to involve intermediates.[18−21] In contrast, little is known about the folding process for RNA G-quadruplexes. Under equilibrium conditions, it is well-known that the thermodynamic and structural properties of RNA and DNA G-quadruplexes differ markedly.[22−24] RNA G-quadruplexes tend to be more stable

Received: October 5, 2012
Published: October 31, 2012

than DNA G-quadruplexes of the same sequence. Consequently, RNA G-quadruplexes formed from only two G-quartets are still very stable[25−27] and may possess a functional role.[27,28] RNA G-quadruplexes also show a preference for parallel topologies in their final folded state, whereas DNA G-quadruplexes are known to adopt multiple topologies.[29] We might therefore anticipate a difference in the kinetic folding properties between RNA and DNA G-quadruplexes, and such differences may provide insight into their utility as recognition motifs and regulatory structures.

A stopped-flow device allows rapid mixing of reaction components followed by direct observation of the kinetics of the reaction at millisecond resolution. In this paper we describe the use of stopped-flow mixing to determine the kinetics of the folding reaction between G-quadruplex forming oligonucleotides and K^+. We used the RNA G-quadruplex formed from the human telomeric repeat as a basis for comparison to its DNA equivalent and to two RNA G-quadruplexes that each comprise only two G-quartets. The measurements revealed comparable time scales of folding for the telomeric RNA and DNA G-quadruplexes, which were also within the time scales for eukaryotic transcription and translation. While the 2-quartet RNA G-quadruplex with single A loops folded with a similar sequential kinetic mechanism to the 3-quartet G-quadruplexes, the kinetic behavior of the UUA looped variant suggested parallel folding pathways leading to a G-quadruplex and a competing structure. Our findings suggest that thermodynamically less stable versions of 2-quartet RNA G-quadruplexes (e.g., those with longer loops)[22] may serve a better role as tunable regulatory elements as long as they can still form within a biologically relevant time scale.

MATERIALS AND METHODS

Oligonucleotides. DNA and RNA oligonucleotides used in this study were purchased from Biomers (Germany) and used as supplied (HPLC purified and lyophilized). The abbreviations and sequences of the oligonucleotides are as follows: TERRA, (r[$(GGGUUA)_3GGG$]); hTelo, (d[$(GGGTTA)_3GGG$]); G2A, (r[$(GGA)_3GG$]); and G2UUA (r[$(GGUUA)_3GG$]). Stock solutions were prepared with nuclease free water (Ambion Biosciences). Concentrations of single stranded oligonucleotide were determined by UV absorption at 260 nm after heating at 95 °C for 5 min, and molar extinction coefficients were calculated using OligoAnalyzer version 3.1 (http://eu.idtdna.com/analyzer/applications/oligoanalyzer/default.aspx). Molar extinction coefficients used for the calculations were: hTelo, 215 000; TERRA, 223 400; G2A, 121 800; G2UUA, 183 000 L mol^{-1} cm^{-1}.

UV Melting. UV melting curves were collected using a Cary 100 UV−vis spectrophotometer (Agilent) by measuring absorbance at 295 nm as a function of the temperature. RNA oligonucleotide solutions were prepared to a final concentration of 5 μM in 10 mM lithium cacodylate buffer (pH 7.0) containing potassium chloride at specified concentration. The melting experiments were performed in 10 mm path length quartz cuvettes with 1 mL of buffered oligonucleotide solution. A steady stream of nitrogen was applied to prevent condensation at low temperatures, and the solutions were covered with a layer of mineral oil to minimize evaporation. Temperature ramps consisted of low → high → low → high cycles with a temperature range of 10−90 °C. A ramp rate of 0.25 °C/min was maintained for all experiments with data collection every °C. Melting temperatures were determined as where fraction folded = 0.5.[30]

Circular Dichroism. Circular dichroism spectra were recorded on a Chirascan spectropolarimeter (Applied Photophysics, Leatherhead, U.K.) using a 1 mm path length quartz cuvette at an oligonucleotide concentration of 10 μM in 10 mM Tris buffer (pH 7.0) containing concentration of potassium chloride specified. Sample preparation involved heating at 95 °C for 5 min followed by slow cooling to room temperature. Scans were performed at 25 °C over a range of 220−340 nm, and each final spectrum is the average of four scans taken with a step size of 1 nm, a time per point of 1 s, and a bandwidth of 0.5 nm. A buffer only blank was subtracted from each spectrum, and data were zero corrected at 340 nm.

Equilibrium Titration. UV absorption spectra were measured from 220−340 nm after serial additions of concentrated potassium chloride stock solution to 5 μM oligonucleotide solution in 10 mM Tris (pH 7.0) at 25 °C on a Cary 100 UV−vis spectrophotometer (Agilent). After each addition of potassium, the solution was mixed by inversion and allowed to equilibrate for a minimum of 1 min before recording the spectrum. Each spectrum was corrected for buffer and salt absorbance and dilution. Difference spectra were generated by subtracting the spectrum at $[K^+]$ = 0 mM from the spectrum at each $[K^+]$ ($\Delta A_i = A_i - A_0$), and the absorbance was converted to a difference in molar extinction coefficient $\Delta\varepsilon_i$. To assess the dependence of G-quadruplex folding on $[K^+]$, titration curves were generated by plotting $\Delta A_i/\Delta A_f$ at 295 nm for TERRA and hTelo (or 297 nm for G2A and G2UUA) against $[K^+]$. ΔA_f is the difference in absorbance between $[K^+]$ = 0 mM and [K+] where 295/297 nm is maximum, i.e., where G-quadruplex folding is complete. $K_{0.5}$ and the Hill coefficient (n_H) were obtained by fitting the curve to eq 1, which assumes two state folding:

$$\Delta A_i/\Delta A_f = \frac{[K^+]^{n_H}}{K_{0.5} + [K^+]^{n_H}} \quad (1)$$

Kinetic Measurements. The kinetics of oligonucleotide folding upon rapid mixing with potassium chloride was measured using a SX20 stopped-flow device (Applied Photophysics, Leatherhead, U.K.). G-quadruplex folding was monitored via absorbance changes at 295 nm as a function of time (path length 1 cm, bandwidth 4 nm). Oligonucleotide solutions were prepared at a concentration of 10 μM in 10 mM Tris (pH 7.0) and mixed 1:1 (v/v) with potassium chloride in the same buffer, to give a final reacting oligonucleotide concentration of 5 μM. Prior to each mixing experiment, the buffered oligonucleotide solutions were heated at 95 °C for 5 min and slow cooled to rt to ensure an equilibrium mixture of minimally structured oligonucleotide (i.e., the starting state). A circulating water bath was used to maintain constant temperature. Control mixing experiments of buffer/buffer, buffer/buffered oligonucleotide, and buffer/buffered potassium chloride did not show any significant changes in absorbance at 295 nm over time. A minimum of five successive mixing reactions was averaged for analysis (10 000 points, logarithmic scale), and each experiment was performed in duplicate. Since [oligonucleotide] ≪ $[K^+]$, the conditions are pseudo-first order so the data were fitted to an equation describing a double exponential using Origin 8.0 (OriginLab) (eq 2):

$$Abs_{295nm} = A_1 \exp^{-k_1 t} + A_2 \exp^{-k_2 t} + c \quad (2)$$

The data were fitted from 15 ms onward to eliminate mixing artifacts. Reactions with double the oligonucleotide concentration (10 μM) showed no change in rate (within error), demonstrating that the folding process is intramolecular (TERRA k_1 = 121.7 s^{-1}, k_2 = 20.4 s^{-1}; hTelo k_1 = 159.5 s^{-1}, k_2 = 22.5 s^{-1}).

$\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ were determined from the linear form of the Eyring equation:[31]

$$\ln k_{obs} = \frac{\Delta S^{\ddagger}}{R} - \frac{\Delta H^{\ddagger}}{RT} + \ln \frac{k_B T}{h} \quad (3)$$

where $\Delta H^{\ddagger}$ is the enthalpy of activation and $\Delta S^{\ddagger}$ is the entropy of activation. An Eyring plot of $\ln(k_{obs}/T)$ vs $1/T$ gives slope = $-\Delta H^{\ddagger}/RT$ and intercept = $\ln(k_B/h) + \Delta S^{\ddagger}/RT$.

Equation 4 was used to fit the data from plots of k_{obs} vs $[K^+]$:

$$k_1 = \frac{k_{1,max}[K^+]^n}{[K^+]^n + K_{eq}} \quad (4)$$

where $k_{1,max}$ is the value of the observed rate constant k_1 at saturating $[K^+]$ and $K_{eq} = (k_{rv} + k_1)/k_{fw}$ (k_{fw} and k_{rv} are the forward and reverse

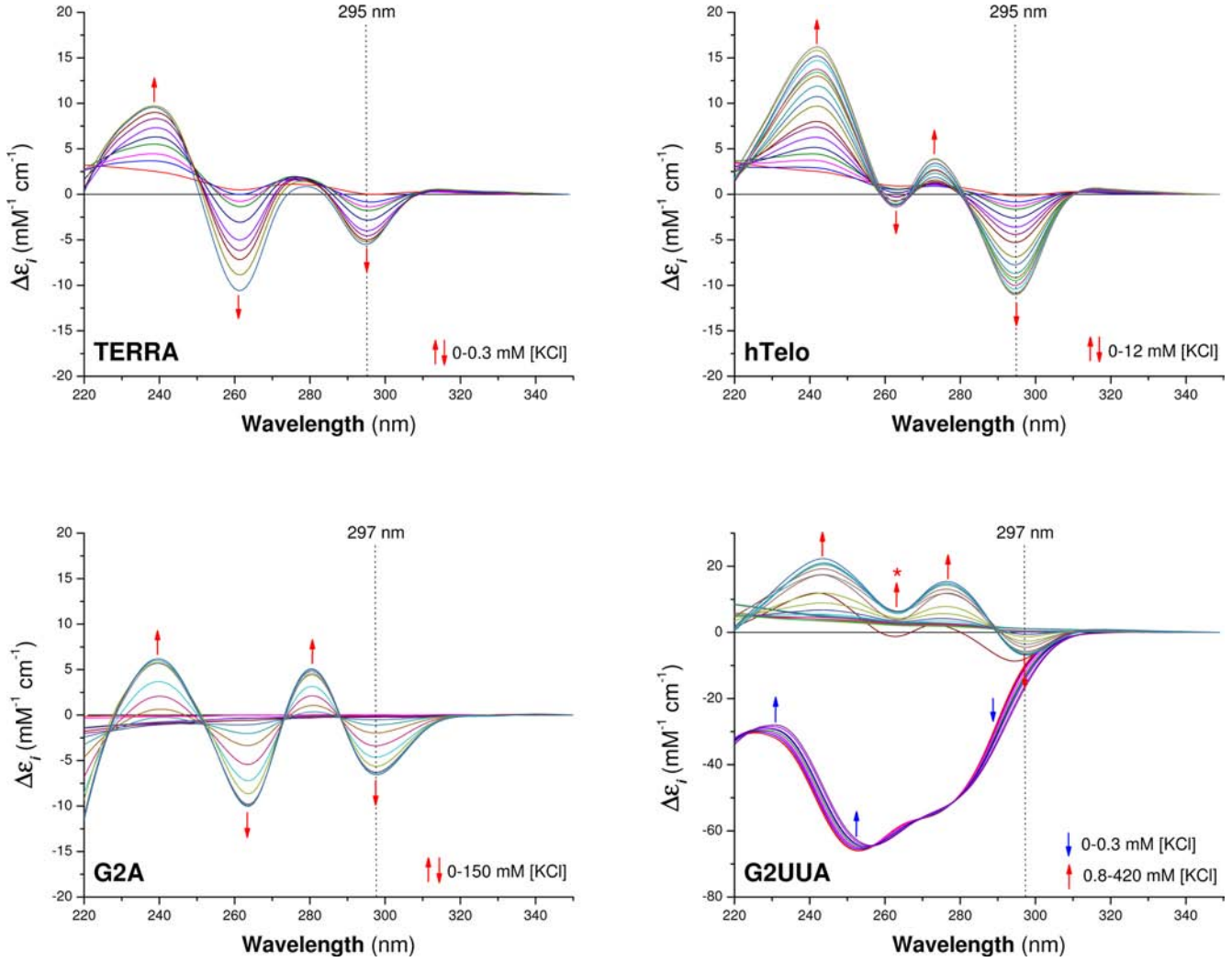


Figure 1. UV difference spectra for the quadruplex-forming oligonucleotides used in this study. The spectra were generated from serial titrations of buffered KCl into 5 μM oligonucleotide in 10 mM Tris buffer (pH 7.0) at 20 °C. The arrows indicate the direction of peak growth. The asterisked arrow on G2UUA plot denotes the region of unique spectral behavior among the oligonucleotides.

rate constants for the pre-equilibrium, respectively). K_{eq} is the $[K^+]$ at which k_1 is at half-maximum and is an inverse measure of the affinity of K^+ for the starting state ensemble. Fixing n = 1 produced a better fit than when the data was allowed to fit to a sigmoidal curve with variable n. For reference, fitting to a sigmoidal dependence generated the following values for n: TERRA, 1.1; hTelo, 1.1; and G2A, 1.2. In all cases, the adjusted R^2 value was higher for the hyperbolic fit than for the sigmoidal fit, where a higher adjusted R^2 indicates a better model. The value is calculated based on the R^2, which denotes the percentage of variation in the dependent variable that can be explained by the independent variables. The adjusted R^2 is a better comparison between models with different numbers of variables and different sample sizes since it adjusts the R^2 for the sample size and the number of variables in the model.

RESULTS AND DISCUSSION

Experimental Design. A panel of four G-quadruplex forming oligonucleotides were selected for this study: TERRA (r[(GGGUUA)$_3$GGG]), hTelo (d[(GGGTTA)$_3$GGG]), G2A (r[(GGA)$_3$GG]) and G2UUA (r[(GGUUA)3GG]).. TERRA and hTelo were chosen to enable a direct comparison between RNA and DNA G-quadruplexes of the same sequence, while G2A and G2UUA were chosen to represent RNA G-quadruplexes that contain only two G-quartets. It was recently reported that 2-quartet RNA G-quadruplexes fold with a stronger positive cooperativity than 3-quartet RNA G-quadruplexes, suggesting a lack of populated intermediates during folding.[27] G2UUA offers a direct comparison to TERRA since they have the same loop nucleotides. G2A was originally identified as an RNA aptamer against bovine prion protein[26] and was chosen for this study because it is already well characterized.[32,33] The folding of each quadruplex-forming oligonucleotide was initially characterized under equilibrium conditions in order to establish the expected changes in absorption upon folding and to determine the minimal K^+ concentrations necessary to induce complete folding in the kinetics experiments. Folding reactions of each quadruplex-forming oligonucleotide were then initiated upon rapid mixing with K^+ using a stopped-flow instrument, and the progress of each folding reaction was followed by monitoring absorbance changes at 295 nm.[34] The design of these experiments is such that they primarily report on G-quartet formation and stacking as the G-quadruplex folds.[34]

Folding Equilibria. *Oligonucleotides Demonstrate Distinct Spectral Responses During K^+ Titration.* UV difference spectra (Figure 1) report on the extent of oligonucleotide folding at each potassium concentration during a titration and

dx.doi.org/10.1021/ja309851t | J. Am. Chem. Soc. 2012, 134, 19297−19308

can offer clues about the level of complexity of the folding process. Each oligonucleotide produced a distinct set of difference spectra, indicative of variations in structural sensitivity to a given [K^+]. The spectral changes were consistent with G-quadruplex folding,[18] whereby the difference spectra gradually adopt sets of increasingly more intense minima near 260 and 295 nm and maxima near 240 and 275 nm during the titration. For TERRA, hTelo, and G2A, these changes were complete in the presence of 0.3, 12, and 100 mM KCl, respectively. In the case of G2UUA, the titration revealed that G-quadruplex formation occurred over a much larger range of [K^+] (10−300 mM KCl) and was preceded by an additional set of unique spectral changes at low [K^+] (<0.5 mM) that was not observed with the other oligonucleotides.

G2A was the only oligonucleotide to give rise to difference spectra with isosbestic wavelengths (226, 252, 273, 288 nm, Figure S1) in the G-quadruplex forming range of ∼0.3−12 mM KCl, supporting the notion that some 2-quartet RNA G-quadruplexes may be two-state folders.[27] In contrast, the folding of TERRA, hTelo, and G2UUA proceeded without isosbestic points. This is consistent with the presence of spectroscopically distinct folding intermediates.[18]

Ability of K^+ to Induce G-quadruplex Folding Does Not Correlate with the Thermal Stability of the K^+-bound G-quadruplex Structure. Absorbance changes at 295 nm (for hTelo and TERRA) and 297 nm (for G2A and G2UUA) were then used to extract titration curves for each oligonucleotide (Figure 2) by fitting the data to eq 1. The best-fit parameters

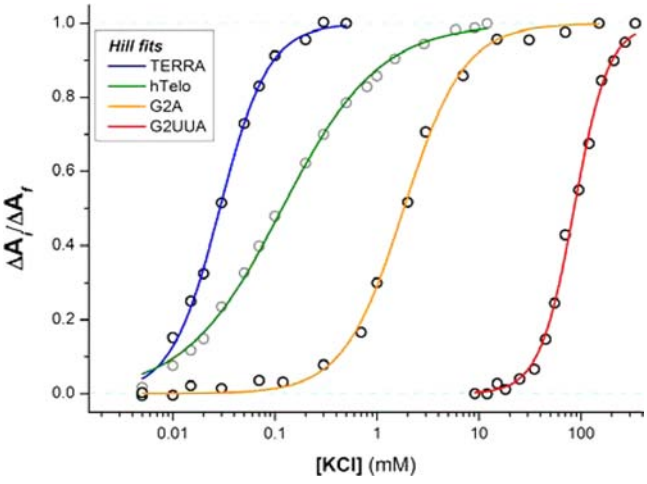


Figure 2. Titration curves derived from absorbance changes at wavelength used to monitor G-quadruplex folding. The slope of the curve indicates the degree of cooperativity with respect to [K^+].

for $K_{0.5}$ ([K^+] needed to fold half the oligonucleotide when measured at 295 nm) and n_H are presented in Table 1. The values obtained from the fitted curves are empirical descriptions of the coupling between K^+ binding and oligonucleotide folding, rather than pure binding constants or equilibrium folding constants.[35]

The midpoint of K^+ induced folding ($K_{0.5}$) of TERRA (0.03 mM K^+) and hTelo (0.12 mM K^+) occurred in the micromolar [K^+] range, while folding of G2A (1.8 mM K^+) and G2UUA (85 mM K^+) was only achieved much later in the titration, approaching physiological [K^+] in the case of G2UUA. It is interesting to note that TERRA and G2A both have a melting

Table 1. Melting Temperatures and Equilibrium Binding Parameters[a]

oligo	T_m (°C)	$K_{0.5}$ (mM)	n_H
TERRA	75	0.03 ± 0.005	1.8 ± 0.1
hTelo	70	0.12 ± 0.02	0.9 ± 0.1
G2A	75	1.8 ± 0.2	1.5 ± 0.1
G2UUA	42	85 ± 7	2.6 ± 0.3

[a] T_m were determined as where fraction folded = 0.5, and are accurate to ±1 °C. For the case of G2UUA, the quoted T_m is for the melting transition. During annealing, the transition occurred at 32 °C. Buffer = 10 mM lithium cacodylate (pH 7.0) containing 100 mM KCl. Buffer =10 mM Tris (pH 7.0).

temperature of 75 °C (100 mM KCl), but there is a >80-fold difference in their sensitivity to K^+.

Hill Coefficients Reveal Higher Folding Cooperativity for RNA G-quadruplexes and Infer Folding Intermediates for TERRA and hTelo. The Hill coefficient (n_H) has traditionally served two main purposes: as a convenient measure of the extent of cooperativity between multiple binding sites and as an estimate of the stoichiometry of ligation (n). While the specific conditions under which n_H gives an accurate n are considered physically impossible (i.e., all-or-nothing binding arising from absolute cooperativity),[36] it is still useful to compare n_H to the known binding stoichiometry since non-theoretical values of n_H can indicate departures from the two-state folding that is assumed by the Hill equation.

The values for n_H were in the order hTelo (0.9) < G2A (1.5), TERRA (1.8) < G2UUA (2.9). Consistent with previous reports,[27] the 2-quartet RNA G-quadruplexes G2A and G2UUA demonstrated strong positive folding cooperativity with respect to [K^+], and the non-unity n_H values possibly suggest a higher stoichiometry of binding between K^+ and the oligonucleotide than is accounted for by assumed interstitial binding of one K^+ alone. While the Hill coefficient for TERRA also indicated positive folding cooperativity, both of the 3-quartet G-quadruplexes demonstrated lower values for n_H than the theoretical value expected from binding of two interstitial K^+. Herschlag and colleagues have shown that heterogeneity can distort the cooperativity measured in bulk from the actual cooperativity of individual molecules,[37] whereby the observed cooperativity derived from the ensemble curve becomes smaller as heterogeneity increases. Therefore a possible interpretation of the departure from theoretical values of n_H for TERRA and hTelo folding is that stable folding intermediates contribute to sample heterogeneity leading to a shallower slope in the titration curve and smaller n_H. Additionally, the lower value for hTelo may reflect multiple final folded conformations that include a subpopulation of basket-type topologies that contain only two G-quartets and therefore only one interstitial binding site.[38]

G2UUA Behaves Differently. The difference spectra for G2UUA and the derived titration curves revealed that this oligonucleotide has a different response to K^+ compared to the other oligonucleotides. In addition to the unique spectral changes at low [K^+] that may account for initial secondary structure formation, the minimum peak near 260 nm also evolved differently. Specifically, the peak became more positive with increasing [K^+]; the opposite of what was observed for TERRA, hTelo, and G2A (Figure 1). This may suggest the presence of non-quadruplex structure that is also spectrally

dx.doi.org/10.1021/ja309851t | J. Am. Chem. Soc. 2012, 134, 19297−19308

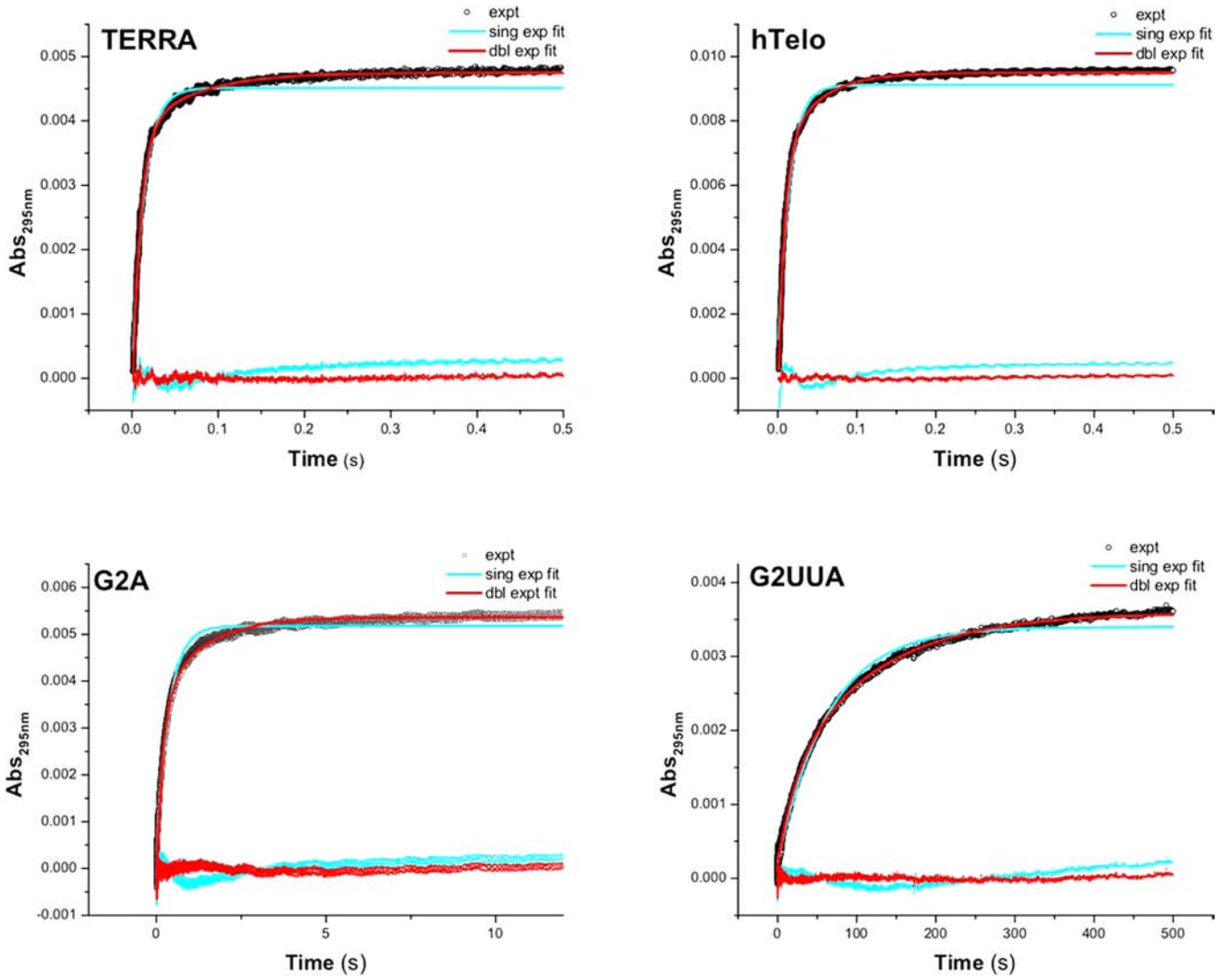


Figure 3. Representative kinetics traces for folding reactions between the quadruplex-forming oligonucleotides and K^+ fit to single or double exponential processes. Residual plots [blue dots (single), red dots (double)] show deviations between experimental and fitted curves.

sensitive to 260 nm, for example, the stacking of Watson−Crick paired bases.

While the ^{1}H NMR spectrum of G2UUA in 100 mM KCl shows that the equilibrium mixture of structures is likely to include a G-quadruplex structure, the circular dichroism (CD), electrophoretic, and UV melting data suggest that the concomitant existence of a competing structure is probable (Figure S2). CD signals (minimum at 240 nm, maximum at 260 nm) that are consistent with the formation of known parallel G-quadruplex structures[39] were detectable in 100 mM KCl, but a dominant minimum near 270 nm becomes more apparent as $[K^+]$ is increased. The non-denaturing gel electrophoresis of the folded G2UUA oligonucleotide showed two distinct species and provided further evidence for at least two different classes of folded structure, while multimerization was ruled out due to the absence of a slow migrating band in the gel. The UV thermal melting and annealing curves for G2UUA in 100 mM KCl were also non-superimposable, which could support the presence of multiple species. Collectively, these results suggest that the G2UUA oligonucleotide possesses the ability to fold into more than one structure.

The trinucleotide repeat $(UUA)_{17}$ forms semistable hairpins existing as three conformers.[33] It is therefore conceivable that the additional structure formed from the G2UUA oligonucleotide may be due to a hairpin with non-canonical G:U or G:G base pairing. Interestingly, it was observed that loop variants G2UAU and G2AUU also showed multiple species in non-denaturing gels, while G2AAU, G2AUA, and G2UAA migrated as single bands (Figure S2). This suggests that having two Us in each loop brings about the structural polymorphism within the affected sequences, and it appears that the specific sequence of the loops may be responsible for the properties of the G2UUA sequence.

Collectively, the data for G2UUA may hint at a departure from the pathway to folded structures followed by the other quadruplex-forming sequences, suggesting it would be of particular interest to see if such differences manifest in the stopped-flow kinetics conditions.

Folding Kinetics. *TERRA and hTelo Demonstrate Similar Folding Time Scales, and 3-Quartet G-quadruplexes Fold Faster than 2-Quartet G-quadruplexes.* Folding reactions were performed by rapid mixing of oligonucleotide in buffer with potassium chloride in buffer at 25 °C in a stopped-flow

dx.doi.org/10.1021/ja309851t | J. Am. Chem. Soc. 2012, 134, 19297−19308

Table 2. Kinetic Parameters[a] for K^+-Initiated Folding of G-Quadruplex Forming Sequences

oligo	k_1 (s^{-1})	k_2 (s^{-1})	$\tau_1{}^c$ (ms)	$\tau_2{}^c$ (ms)	$k_{1,max}$ (s^{-1})	K_{eq} (mM)
TERRA	124.2 ± 9.6	18.0 ± 1.8	5	55	187	39
hTelo	148.1 ± 13.3	25.1 ± 3.2	7	40	425	127
G2A	7.6 ± 0.7	1.4 ± 0.03	131	714	13	82
G2UUA[b]	0.05 ± 0.0.1	0.009 ± 0.0004	20×10^3	111×10^3	–	–

[a]Determined by stopped-flow rapid mixing with folding reaction concentrations of 5 μM oligonucleotide and 90 mM KCl in 10 mM Tris (pH 7) at 25 °C. [b]Conditions as above except for 500 mM KCl concentration. [c]Where τ is the relaxation time constant.

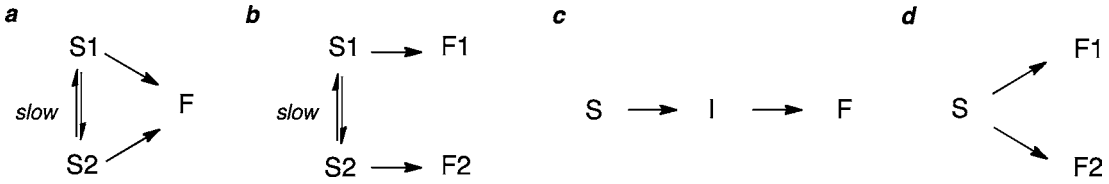


Figure 4. Possible mechanistic interpretations for the observed biphasic kinetics in the folding reactions of TERRA, hTelo, and G2A G-quadruplexes. S, I, and F represent starting, intermediate, and final folding states, respectively.

device to give final reaction concentrations of 5 μM oligonucleotide and 90 mM KCl (buffer = 10 mM Tris, pH 7.0). The progress of each folding reaction was followed over time by recording the UV absorbance signal at 295 nm, which is a measure of G-quartet stacking and therefore G-quadruplex structure formation.[34] The quadruplex-forming oligonucleotides were allowed to fold until the absorbance signal approached a constant value, signifying that the G-quadruplex folding reaction had reached equilibrium.

Kinetics traces showing changes at 295 nm as a function of time for hTelo, TERRA, G2A, and G2UUA are presented in Figure 3. The folding reactions for hTelo and TERRA were complete within 400 ms after mixing, whereas the folding of G2A required an appreciably longer time frame of ~5 s. In the case of G2UUA, the mixing reactions were conducted with a higher concentration of 500 mM KCl to ensure that the experiments allowed for measurement of G-quadruplex folding. Under these conditions the absorbance changes required a much longer time of 800 s to reach completion.

Folding Involves Two Kinetically Resolvable Steps. The pseudo-first order conditions (oligonucleotide ≪ $[K^+]$) allowed us to fit the kinetic traces to eq 2, which describes a sum of exponentials with observable rate constants. For all the quadruplex-forming oligonucleotides, curve fitting of the kinetic traces indicated that two exponential processes optimally described the folding reactions (adjusted $R^2 > 0.995$). Fitting the kinetics data to a single exponential equation consistently led to noticeable deviations of the fitted curve from the experimental data. Residual plots for both single and double exponential fits are included in Figure 3 for comparison.

For folding in the presence of 90 mM KCl, the observed rate constants, denoted k_1 and k_2 (Table 2), were similar for TERRA (124.2 ± 9.6, 18.0 ± 1.8 s^{-1}) and hTelo (148.1 ± 13.3, 25.1 ± 3.2 s^{-1}), while G2A was >15-fold slower (7.6 ± 0.7, 1.4 ± 0.03 s^{-1}). Despite the higher $[K^+]$ of 500 mM, k_1 and k_2 for G2UUA (0.05 ± 0.0.1, 0.009 ± 0.0004 s^{-1}) were 3 orders of magnitude slower compared to TERRA.

The biphasic character of the folding reactions may have several origins (Figure 4). If the starting state (S) ensemble consists of two kinetically distinct populations separated by significant energy barriers (i.e., their interchange is slow with respect to the time resolution of the analytical method, such that both populations are detectable), then parallel pathways may account for the two rate constants observed in this study. This scenario could have states S1 and S2 converting to the same final folded state (F) (Figure 4a) or converting to two distinct folded states (F1 and F2) (Figure 4b). On the other hand, the biphasic character may point to a folding intermediate (I) that is kinetically distinct from the starting and folded states. In this scenario, the intermediate is obligatory along a sequential folding pathway and is likely to represent an ensemble of metastable, partially folded structures (Figure 4c). A kinetically homogeneous starting state could also branch off into two different folded states (F1 and F2) (Figure 4d).

Under the initial conditions of our study (10 mM Tris, pH 7.0), the starting ensemble is expected to possess secondary structure since the high ionic strength of the buffer will induce the rapid collapse transitions from the fully denatured state that characterize the early phase of nucleic acid folding.[40] Given that multiple conformations of folded G-quadruplex structures can coexist at equilibrium, at least in the case of structures formed from DNA, there is potential for this stabilized starting ensemble to also consist of the multiple populations. If interconversion between such populations is slow relative to folding, then the kinetically distinct populations in the starting state ensemble required by Figure 4a and 4b can arise. However, it has already been demonstrated that multiple G-quadruplex conformations in the equilibrium folded state may not necessarily lead to multiphasic kinetics.[18] Both d[TTGGG-(TTAGGG)$_3$A] and d[AGGG(TTAGGG)$_3$] have been shown to fold with monophasic kinetics in the presence of K^+ (UV detection), even though the former forms a hybrid topology,[41] and the latter forms a mixture of G-quadruplex conformations.[38] Figure 4d can be ruled out by taking into account that TERRA, G2A, and hTelo all demonstrated biphasic kinetics in folding. If conformational heterogeneity in the folded hTelo G-quadruplex (multiple hybrid topologies)[38] was to dictate a branched folding pathway, then less complex folding kinetics for the all parallel TERRA[17] and G2A[42] G-quadruplexes might have been expected. This suggests that a mixture of completely folded G-quadruplex topologies can appear as one kinetic species and that the energetic barriers between conformers are not sufficient to make them kinetically distinguishable.

Therefore for the TERRA, hTelo, and G2A oligonucleotides, the analysis suggests that a sequential folding mechanism is the most plausible, although the design of our experiments does

not allow non-sequential pathways to be absolutely ruled out. Our conclusion of a sequential folding mechanism is consistent with observations made by others.[19,43] The biphasic kinetics observed for G2UUA, on the other hand, is more likely to be attributed to a nonsequential pathway since the evidence supports an equilibrium mixture of a G-quadruplex with nonquadruplex structures.

G-quadruplex Folding is Dominated by Entropy at Physiological Temperature. A common property of biomolecular folding pathways is that a higher temperature does not necessarily lead to a faster reaction, resulting in a deviation from the Arrhenius[44]/Eyring[31] type kinetics usually associated with simple reactions.[45] To investigate whether G-quadruplex folding is consistent with the folding of complex biomolecules, folding reactions for hTelo and TERRA were performed between 5 and 45 °C to characterize the temperature dependence of the rate constants. The trend for both TERRA and hTelo G-quadruplex folding was nonlinear, where a maximum in k_1 and k_2 was reached near 25 °C before decreasing ~3-fold between 25 and 45 °C (Figure S3). Curvatures in Arrhenius and Eyring plots are a hallmark of protein[45] and peptide[46] folding as well as DNA[47] and RNA[48] hairpin loop closures and indicate that folding is dominated by enthalpy at low temperatures and entropy at high temperatures.[49] Anomalous Eyring plots have also been observed for DNA G-quadruplex folding.[18]

Here, linear regression fits of the low (5−25 °C, $1/T$ > 0.00335) and high (25−45 °C, $1/T$ < 0.00335) temperature regions of the Eyring plot (Figure 5) for k_1 was performed to estimate $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ (Table S1), which collectively describe the thermodynamic quasi-equilibrium between the ground and transition states.[31] The fitting highlighted the existence of negative activation enthalpies of folding ($\Delta H^{\ddagger}$) for temperatures above 25 °C ($\Delta H^{\ddagger}_{(k_1,\mathrm{TERRA})}$ = −33.0 kJ mol^{-1}; $\Delta H^{\ddagger}_{(k_1,\mathrm{hTelo})}$ = −38.9 kJ mol^{-1}). An observation of $\Delta H^{\ddagger}$ < 0 may have implications for the folding mechanism of G-quadruplexes, especially if G-quadruplexes are to be considered as an assembly of hairpins which in themselves are known to fold with non-Arrhenius kinetics.[47,48]

K^+ Binding is Prerequisite for Folding and K^+ Stabilizes Folding Steps. In this study, K^+ is expected to both promote the folding of the G-quadruplex and to bind to the folded structure, so the role it plays in the kinetic mechanism should be considered. Folding reactions for TERRA, hTelo, and G2A were performed with [K^+] from 5 to 500 mM to examine the rate dependence of G-quadruplex folding on potassium. Both k_1 and k_2 increased with greater [K^+] and reached saturating rates at high [K^+] (~200−500 mM, depending on oligonucleotide) (Figure 6). Cations (Na^+ and K^+) bound by G-quartets in the $[d(G_3T_4G_3)]_2$ dimeric G-quadruplex have been found to exchange with ions in the bulk solution at a rate greater than 10^3 s^{-1}.[50] Since this is at least an order of magnitude faster than the fastest rate constant observed here (i.e., k_1), we would not expect the dynamics of cation exchange to influence the observed rates unless the exchange process significantly differs for intramolecular G-quadruplexes. The observation of a rate dependence on [K^+] does, however, imply that K^+ affects the free energy of the transition states involved in the two kinetic steps associated with k_1 and k_2.[51] In protein folding studies, saturation kinetics can be indicative of conformational selection mechanisms whereby the ligand 'selects' a conformation close to that of the bound form, from the ensemble of conformations

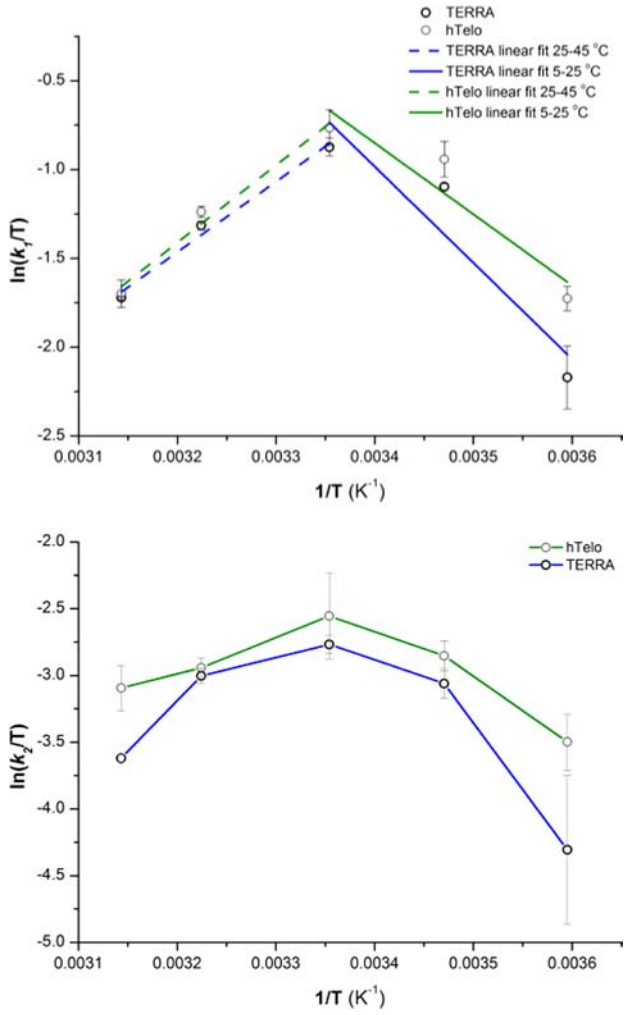


Figure 5. Eyring plots for k_1 and k_2 for folding of TERRA and hTelo. For the k_1 data set, separate linear regressions were fit to the low- and high-temperature regions. For the k_2 data set, the lines connect the data points only. Error bars are given as the standard deviation of two measurements.

populated.[52,53] Our data could therefore be consistent with G-quadruplex folding whereby early on K^+ binds to a fraction of molecules within the starting ensemble that are suitably preorganized for ligand binding.

We found that the saturation behavior between k_1 and K^+ was best described by a hyperbolic relationship where n in eq 4 was fixed at 1. Such a hyperbolic dependence has previously been mechanistically interpreted as a rapid, cation-dependent pre-equilibration step where a complex between the oligonucleotide and a single K^+ is formed, before folding begins.[18] Here, it can be considered as the formation of the first potassium-bound species, S_K−K^+, which is the starting point for G-quadruplex folding under the experimental conditions:

$$S + K^+ \overset{K_{eq}}{\rightleftharpoons} S_K - K^+ \quad (5)$$

The parameters describing the pre-equilibrium are shown in Table 2. TERRA shows the greatest affinity for K^+ (K_{eq} = 39 mM), followed by G2A (K_{eq} = 82 mM), and then hTelo (K_{eq} = 127 mM). It is interesting to note while a higher [K^+] was required to fold G2A compared to hTelo during the equilibrium titration (i.e., $K_{0.5,\mathrm{G2A}}$[1.8 mM] > $K_{0.5,\mathrm{hTelo}}$[0.12

dx.doi.org/10.1021/ja309851t | J. Am. Chem. Soc. 2012, 134, 19297−19308

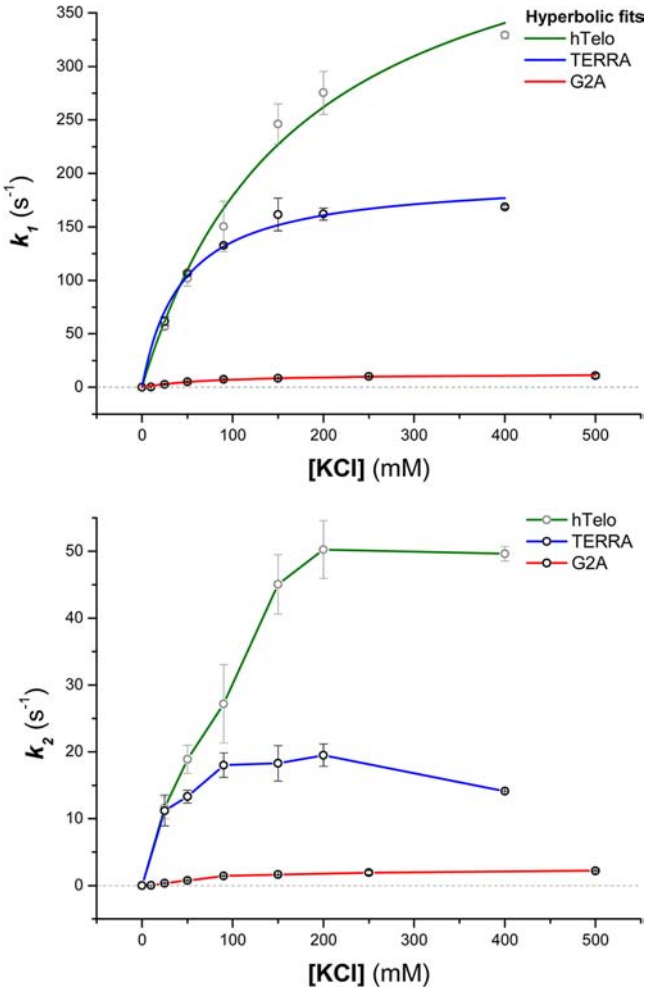


Figure 6. Observed rate constants k_1 and k_2 as a function of $[K^+]$ concentration for folding of TERRA, hTelo, and G2A. Using eq 4, k_1 data points were fit to a hyperbolic function. For the k_2 plot, the lines connect the data points only. Error bars are given as the standard deviation of two measurements.

mM]), the reverse was true for establishing the pre-equilibrium in the kinetics experiments ($K_{eq,G2A}$[82 mM] < $K_{eq,hTelo}$[127 mM]). On the other hand, TERRA demonstrated the lowest K_{eq}[39 mM] and $K_{0.5}$[0.03 mM] among the oligonucleotides. The trends indicate that: (a) the binding affinity of the first K^+ to partially folded G-quadruplexes to establish the pre-equilibirum is higher for RNA than for DNA, irrespective of how many G-quartets the RNA G-quadruplex will ultimately adopt; and (b) this information is obscured in measurements made at thermodynamic equilibrium, since overall folding and ligation is achieved at a lower $[K^+]$ for 3-quartet G-quadruplexes than for 2-quartet G-quadruplexes, irrespective of whether they are formed from RNA or DNA. Furthermore, the highest value for $k_{1,max}$ was achieved by hTelo, demonstrating that binding affinity of K^+ to the starting ensemble is not correlated to the kinetics of the first folding step.

Minimal Kinetic Mechanism. The analysis so far suggests that the kinetic folding mechanism for the G-quadruplexes formed from the TERRA, hTelo, and G2A oligonucleotides follows a multistep sequential pathway involving at least one folding intermediate and two kinetic steps, which is preceded by a pre-equilibration step where K^+ binding takes place. Taking these components into account, a minimal kinetic mechanism that is consistent with the analysis is presented in eq 6:

$$S + K^+ \overset{K_{eq}}{\rightleftharpoons} S_K - K^+ \xrightarrow{k_{1,max}} I - K^+ \xrightarrow{k_2} F - nK^+ \quad (6)$$

where the observed rate constant k_1 comprises the first kinetic step (the formation of I) described by $k_{1,max}$ and K_{eq}. The second kinetic step is described by the observed rate constant, k_2. It should be noted that the absolute assignment of the faster observed rate constant to the first step and the slower observed rate constant to the second step would require the concentration of the I species to be monitored over time. An accumulation of the I species would support the fast first step/slow second step case, while we would expect a very low steady-state concentration of I to be established in the slow first step/fast second step case.[54] In the limiting case of $k_1 \ll k_2$, the relaxation process would manifest as a single exponential, since I would transition to F as soon as I was formed, which was not observed here. Although our experiments do not provide concentration profiles of each species over time, our assignment of a fast first step and slow second step in the kinetic mechanism is consistent with our observation of a transient maximum in the refolding probability timecourse for an intermediate species in single-molecule studies of TERRA folding (Shi et al., to be submitted for publication). Furthermore, in thermal unfolding studies of the telomeric DNA G-quadruplex, intermediates were sufficiently populated for optical detection by Gray et al.[43] Finally, n represents the number of specifically bound K^+ (assumed $n = 1$ (G2A) and $n = 2$ (TERRA, hTelo)).

Our kinetic mechanism implies folding intermediates. Evidence for stable folding intermediates from telomeric DNA sequences has been provided under thermodynamic conditions by calorimetric measurements[55] and singular value decomposition analyses of optical spectra vs temperature matrices.[43] Using a molecular dynamics approach, Mashimo et al. demonstrated that triple-stranded helical structures are energetically feasible intermediates along the telomeric DNA G-quadruplex folding pathway.[56] Triplexes formed from three telomeric DNA repeats have subsequently been experimentally observed as a mechanically stable species in single-molecule optical tweezer experiments, and the species also coexists with the fully folded G-quadruplex.[57]

Proposed Folding Pathway. We now propose a folding pathway that is consistent with the kinetic mechanism and considers possible structures within each kinetic state (Figure 7). The first event could involve the binding of K^+ to collapsed structures within the starting ensemble, perhaps selecting for suitably preorganized conformations, such as hairpins. As implicated by the hyperbolic dependence of the folding rate on $[K^+]$, such a binding event would result in oligonucleotides complexed with a single K^+. This is significant because the triplex intermediates proposed by Mashimo et al.[56] were initially derived from GG base paired hairpin structures that are also stabilized by one K^+. The structures in the S_K–K^+ kinetic state could therefore potentially take the form of cation-stabilized hairpins, which could then rapidly fold into more stable triplex intermediates to give I–K^+. The rate-limiting second kinetic step could then represent the formation of G-quartets and slow conformational rearrangements to give the final folded G-quadruplex. For the RNA G-quadruplexes TERRA and G2A, the rearrangements could, for example,

dx.doi.org/10.1021/ja309851t | J. Am. Chem. Soc. 2012, 134, 19297–19308

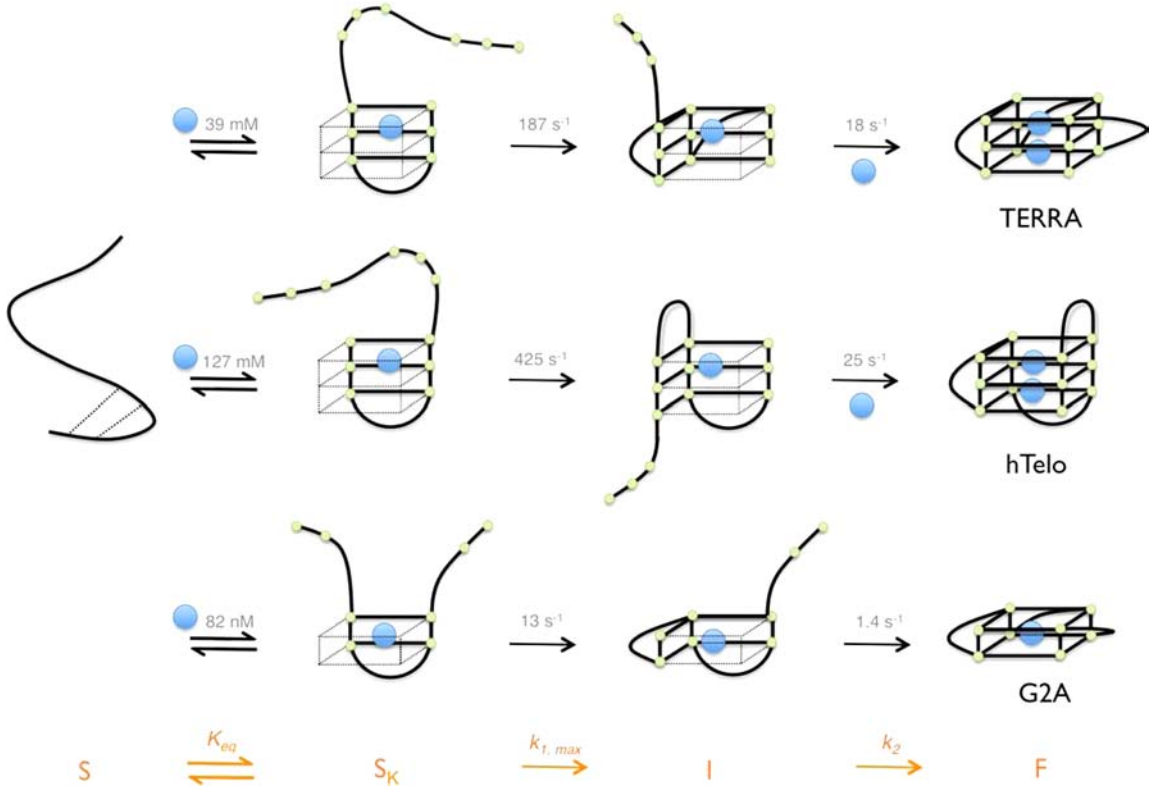


Figure 7. Proposed pathway for K^+-initiated G-quadruplex folding featuring hairpin and triplex intermediates in 90 mM KCl. Each kinetic species (S, S_K, I, F) is an ensemble of structures. The actual structures presented here are possible structures within the ensemble. The dashed lines in the S state signify that the collapsed oligonucleotides possess secondary structure. The dashed lines in the S_K and I states outline the G-quartet framework and do not indicate interactions between bases. The large blue spheres represent K^+, while the small green spheres indicate guanines involved in G-quartet formation.

involve the formation of double chain reversal loops. For hTelo, the final step could account for the final mixture of hybrid topologies (types 1−3) expected in K^+ solution.[38] The observation of accelerated folding rates with increasing $[K^+]$ suggests that K^+-stabilized structural changes feature in each folding step.

Although triplex folding intermediates were originally proposed for the folding of telomeric DNA sequences, the ab initio calculations of the stabilization energies from which the triplexes were hypothesized were performed on guanines only.[56] Therefore the finding that the stabilization energy of a G-triplet (−50.6 kcal mol^{-1}) is higher than a GG-pair (−18.4 kcal mol^{-1}) and comparable to a G-quartet (−67.8 kcal mol^{-1}) may also support the formation of triplex intermediates during RNA G-quadruplex folding. A difference, however, is that RNA G-quadruplexes are well-known to adopt parallel topologies with double chain reversal loops. This means that all of the glycosylic conformations are of the *anti* form, in contrast to the mixture of *anti* and *syn* glycosylic conformations that feature in the hairpins and triplexes proposed by Mashimo et al.[56] The *anti* glycosylic conformation is generally preferred in structured nucleic acids due to the lower steric clash between the base and the sugar, and this effect is enhanced in RNA because of the additional presence of a 2′-OH on the sugar. However, it is possible for the base to be in the less preferred *syn* conformation if constrained by the surrounding structure. The mixture of glycosylic conformations in the hairpin and triplex structures proposed by Mashimo et al. therefore arises from the nature of the loops that connect the G-tracts in these structures. These loops ultimately map directly to the two lateral loops that comprise the final hybrid type-1 and -2 G-quadruplex topologies for which the folding pathway was originally proposed.

If the folding of RNA G-quadruplexes were to also pass through a triple stranded intermediate, then it is conceivable that the energetically preferred *anti* glycosylic conformations may be present from early on in the folding process, in which case the hairpins and/or triplexes will possess folded back loop regions that are consistent with the double chain reversal loops found in the final parallel G-quadruplex structures. Alternatively, the *anti* conformations may potentially be attained via a rearrangement of the triplex intermediates of the type hypothesized for the telomeric DNA sequences.

Comparisons Between Different G-quadruplexes. The design of the oligonucleotides allows a direct comparison between: (a) a RNA and DNA G-quadruplex of the same sequence (TERRA vs hTelo); (b) a 2-quartet RNA G-quadruplex and a 3-quartet RNA G-quadruplex (TERRA vs G2A); and (c) 2-quartet RNA G-quadruplexes with different loop lengths (G2A vs G2UUA).

TERRA vs hTelo. Since the rate-limiting step characterized by k_2 is similar for both TERRA and hTelo, the largest difference between the folding of these two G-quadruplexes is in the nature of the initial K^+ binding event and subsequent isomerization to the I intermediate. The >3-fold lower K_{eq} for TERRA over hTelo suggests that the affinity of K^+ for RNA G-quadruplexes is greater than for DNA G-quadruplexes of the same sequence. The consequence is that $k_{1,max}$ is attainable at a lower $[K^+]$ for TERRA than hTelo. However, this does not necessarily lead to a faster rate of formation of I since $k_{1,max}$ is

dx.doi.org/10.1021/ja309851t | J. Am. Chem. Soc. 2012, 134, 19297−19308

noticeably higher for hTelo (425 s^{-1}) than for TERRA (187 s^{-1}). This suggests that K^+ contributes more to the cation-stablilized structural changes that take place between the S_K (e.g., hairpin) and the I (e.g., triplex) kinetic species during the folding of the DNA G-quadruplex compared to the RNA G-quadruplex. The observation that $k_{1,max}$ is reached for TERRA (and G2A) at a lower $[K^+]$ than hTelo is consistent with previous findings that the 2′-OH in RNA G-quadruplexes can partake in stabilizing intramolecular interactions.[58] This is not possible in hTelo so the ability of K^+ to facilitate folding of this sequence is more pronounced (i.e., higher observed rate constant as $[K^+]$ increases).

The kinetic distinction could additionally arise from differences in properties of the nucleic acid (sugar moiety, T vs U) and/or the conformational properties of each kinetic species (*anti* vs *syn* glycosylic conformations and resulting loop orientations). Taking into account that the property measured in the folding experiments is the rate of G-quartet formation, conformations that encourage the most stable base stacking interactions in the transition state linking S_K and I are likely to promote the fastest folding reaction. Although the *anti* glycosylic conformation is preferred between a sugar and its base, the energetics of base stacking favors an *anti*:*syn* arrangement between adjacent nucleotides over *anti*:*anti*.[56] This suggests that the first kinetic step along the hTelo folding pathway involves the creation of a greater proportion of *anti*:*syn* stacking interactions in the triplex (I state) intermediate compared to TERRA, resulting in a higher value for $k_{1,max}$. The triplex intermediates proposed by Mashimo et al. for telomeric DNA sequences maximize *anti*:*syn* stacking between the 5 *anti* and 4 *syn* glycoylic conformations.[56] On the other hand, triplex intermediates formed from the TERRA oligonucleotide might contain more (or only) *anti*:*anti* stacking interactions. This would be consistent with a lower $k_{1,max}$ value and lends support to the hypothesis that the double chain reversal loops found in the final parallel TERRA G-quadruplex are already in the triplex.

To achieve the final folded G-quadruplex structure, folding back of the final G-tract and binding of a second K^+ needs to take place. The comparable k_2 values observed for TERRA and hTelo suggest that a common folding event occurs during this kinetic step for both G-quadruplexes. Such a folding event might be expected to involve the formation of the final loop in a double chain reversal orientation to account for the parallel topology of the TERRA G-quadruplex and the hybrid (3 + 1) topologies of the hTelo G-quadruplex. For hTelo, it has been reported that the addition of a double chain reversal loop to a triplex intermediate is preferred over a lateral loop since it minimizes steric hindrance with the central lateral loop.[56,59]

TERRA vs G2A. The observation of a >15-fold difference in folding rates between TERRA and G2A at 90 mM KCl suggests that the number of G-quartets involved in K^+-initiated quadruplex formation can dictate the rate of folding. Specifically, the result suggests that events which encourage the stacking of G-quartets (or partially formed G-quartets) during folding have the ability to lower activation energies of folding, such that a higher number of G-quartets involved in G-quadruplex formation will favor a faster folding reaction. This is not surprising since the stacking of two G-quartets has been calculated to contribute an average of −2.2 kcal mol^{-1} toward the free energy of (DNA) G-quadruplex formation, compared to only −0.2 kcal mol^{-1} from loops.[60] This may explain the slower, and more complex, folding of telomeric DNA quadruplexes previously observed by Gray and Chaires with Na^+ compared to K^+,[18] since K^+ can facilitate the association of neighboring G-quartets due to its ability to bind interstitially. Such assistance from the cation is not available with Na^+ since it coordinates to the center of a G-quartet owing to its smaller size.

It is also notable that the "two-state" folding character that was apparent in the equilibrium titration of G2A did not translate into a single exponential kinetic folding process. Indeed, it has previously been observed that the number of phases in a folding reaction need not correlate between equilibrium and kinetic folding conditions. Gray and Chaires[18] observed monophasic kinetics for the sequence d[AGGG-$(GGGTTA)_3$)], despite equilibrium data supporting the presence of stable populated intermediate states characteristic of multiphasic folding.

G2A vs G2UUA. Despite the higher $[K^+]$ of 500 mM, folding for G2UUA was 2 orders of magnitude slower than G2A (and 3 orders of magnitude slower compared to TERRA). While the longer folding time is consistent with the observation that 2-quartet G-quadruplexes fold slower than 3-quartet G-quadruplexes, the equilibrium observation of an uncharacteristic spectral response to increasing $[K^+]$ and multiple classes of structure points to more complicated folding dynamics that may additionally contribute to the long folding time. Whereas the high stability of 2-quartet RNA G-quadruplexes with shorter loop lengths, such as in G2A, favors the adoption of a G-quadruplex fold, the combination of only two G-quartets and longer loop length in G2UUA appears to render it less able to compete effectively against alternatively folded structures. Consistent with this observation, a Van't Hoff treatment of the 295 nm annealing curve in 100 mM $[K^+]$ gave a free energy of formation of only −0.8 kcal mol^{-1} at 310 K.

Biological Implications. The kinetic constants for the rate-limiting steps of G-quadruplex folding obtained in in 90 mM $[K^+]$ convert to relaxation times of 55, 60, and 700 ms (hTelo, TERRA, and G2A, respectively). Since these values are within the region of rates of eukaryotic DNA replication (∼20 ms/nucleotide) and transcription (∼200 ms/nucleotide),[61] there exists the possibility that partially synthesized G-quadruplex sequences may participate in folding events concurrent with DNA and RNA synthesis. Programs, such as Kinefold,[62] attempt to address this by predicting cotranscriptional folding pathways of RNA molecules, however the algorithms do not yet consider G-quadruplex structure. Recently, we showed that the telomeric repeat binding factor 2[63] (TRF2) protein binds telomeric RNA only when it adopts the G-quadruplex structure and not as the single stranded nucleic acid.[64] TRF2 is an essential component of the shelterin complex that protects telomere ends.[65] The potential shown here for cotranscriptional folding of a telomeric RNA G-quadruplex suggests that the G-quadruplex-dependent binding mechanism of TRF2 has a temporal as well as physical basis.

Given the similar time scales of folding for TERRA and hTelo in the presence of near physiological $[K^+]$, an interesting consideration is the formation of hybrid G-quadruplexes with contributions from both DNA and RNA telomeric repeats. It is known that telomeric RNA colocalizes at the telomere,[14,66] and it has been proposed that TERRA regulates telomere function by sequestering telomeric DNA (the substrate for telomerase).[67] One hypothesis for telomere regulation is through the formation of intermolecular DNA−RNA hybrid G-quadruplexes. These have been observed *in vitro*,[68] and model hybrid

systems have been shown to provide a protective effect for telomere ends in cells.[69] Such hybrids have also been shown to bind to the TRF2 protein with nanomolar affinity.[64] The similar limiting rate measurements made here for the folding of the TERRA and hTelo G-quadruplexes arguably support the association of newly synthesized telomeric RNA repeats with locally unwound single stranded telomeric DNA repeats behind the RNA polymerase and support a potential regulatory mechanism of telomeric function that implicates the formation of hybrid G-quadruplexes. The formation of hybrid G-quadruplexes from nascent RNA and nontemplate DNA has also been implicated in regulating the initiation of DNA replication in human mitochondria.[70]

The capacity of DNA and RNA G-quadruplexes to rapidly adopt structure is likely to be central to their ability to regulate biological function. This may be especially true for RNA G-quadruplexes since regulatory mRNAs and regulatory non-coding RNA tend to have much shorter half-lives (<4 h, some as short as 1 min) than those encoding housekeeping genes (typically >4 h).[71] While 3-quartet RNA G-quadruplexes are generally considered as inhibitory structures due to their high stability, the diverse folding times observed in this study among 2-quartet RNA G-quadruplexes coupled with their generally lower stability[25,72] raise the potential for alternative roles.

There are now an increasing number of examples of 2-quartet RNA G-quadruplexes being reported with putative function,[25,28,72] with the emerging theme being that these G-quadruplexes can serve as switchable and tunable motifs. For example, the activity of gene reporter constructs with inserted 2-quartet RNA G-quadruplexes has been shown to change in response to temperature in a fashion that correlates with G-quadruplex stability, leading to their description as artificial thermoswitches.[25] In contrast, equivalent constructs with 3-quartet G-quadruplexes maintained the same activity irrespective of temperature.

The G-rich segment GGAGGAGGGGGAGGAGGA from the internal ribosome entry site (IRES) of vascular endothelial growth factor (VEGF) mRNA[28] provides a naturally occurring example. This sequence can potentially form >20 different 2-quartet RNA quadruplexes of varying loop lengths and was found to be essential for initiation of cap-independent translation. It was proposed that the sequence redundancy enables the IRES to fine-tune its activity via utilization of different subsets of G-quadruplexes that support translation initiation to varying degrees. Which subset of G-quadruplexes is used will depend on the demands for the VEGF protein at the time. It is anticipated that a criterion for such a regulation mechanism will be the ability of the 2-quartet RNA G-quadruplexes to form within a biologically meaningful time yet not be of such high stability that G-quadruplex structure persists beyond the time it is required. Sequences possessing complex folding behaviors similar to G2UUA may therefore still be of biological relevance, where the combination of moderately fast folding kinetics, conformational diversity, and mild thermodynamic stability can be exploited as advantageous design elements in systems that require complex regulation.

CONCLUSIONS

The equilibrium and kinetics data for K^+-dependent G-quadruplex folding suggest complex folding mechanisms involving partially liganded (K^+) intermediate states. The time scales of folding for TERRA and hTelo are comparable to rates associated with DNA replication and transcription, raising the possibility for coreplicative and cotranscriptional folding. By way of example, this specifically lends support for a quadruplex-dependent hypothesis for telomere regulation. The sharp contrast in folding behavior between G2A and G2UUA suggests that any biological roles for 2-quartet RNA G-quadruplexes will be very sequence dependent.

ASSOCIATED CONTENT

Supporting Information

Annotated UV difference spectra for G2A, biophysical characterization of G2UUA, CD spectra, plots of temperature dependence of rate constants. This material is available free of charge via the Internet at http://pubs.acs.org

AUTHOR INFORMATION

Corresponding Author

sb10031@ cam.ac.uk

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

A.Y.Q.Z. was supported by a studentship from Trinity College. The Balasubramanian laboratory is core-funded by Cancer Research U.K. We thank Dr. Pierre Murat for reading the manuscript and for conducting the NMR experiments.

REFERENCES

(1) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2005**, *33*, 2908.
(2) Todd, A. K.; Johnston, M.; Neidle, S. *Nucleic Acids Res.* **2005**, *33*, 2901.
(3) Kostadinov, R.; Malhotra, N.; Viotti, M.; Shine, R.; D'Antonio, L.; Bagga, P. *Nucleic Acids Res.* **2006**, *34*, D119.
(4) Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. *Nucleic Acids Res.* **2008**, *36*, 6260.
(5) Bugaut, A.; Balasubramanian, S. *Nucleic Acids Res.* **2012**, *40*, 4727.
(6) Shen, W.; Gao, L.; Balakrishnan, M.; Bambara, R. A. *J. Biol. Chem.* **2009**, *284*, 33883.
(7) Hai, Y.; Cao, W.; Liu, G.; Hong, S.-P.; Elela, S. A.; Klinck, R.; Chu, J.; Xie, J. *Nucleic Acids Res.* **2008**, *36*, 3320.
(8) Subramanian, M.; Rage, F.; Tabet, R.; Flatter, E.; Mandel, J.-L.; Moine, H. *EMBO Rep.* **2011**, *12*, 697.
(9) Wanrooij, P. H.; Uhler, J. P.; Simonsson, T.; Falkenberg, M.; Gustafsson, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 16072.
(10) Di Antonio, M.; Rodriguez, R.; Balasubramanian, S. *Methods* **2012**, *57*, 84.
(11) Collie, G. W.; Parkinson, G. N. *Chem. Soc. Rev.* **2011**, *40*, 5867.
(12) Blackburn, E. H. *Cell* **2001**, *106*, 661.
(13) Mergny, J.-L.; Riou, J.-F.; Maillieet, P.; Teulade-Richou, M.-P.; Gilson, E. *Nucleic Acids Res.* **2002**, *30*, 839.
(14) Azzalin, C. M.; Reichenbach, P.; Khoriauli, L.; Giulotto, E.; Lingner, J. *Science* **2007**, *318*, 798.
(15) Schoeftner, S.; Blasco, M. A. *Nat. Cell Biol.* **2008**, *10*, 228.
(16) Xu, Y.; Kaminaga, K.; Komiyama, M. *J. Am. Chem. Soc.* **2008**, *130*, 11179.
(17) Martadinata, H.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 2570.
(18) Gray, R. D.; Chaires, J. B. *Nucleic Acids Res.* **2008**, *36*, 4191.
(19) Gray, R. D.; Petraccone, L.; Trent, J. O.; Chaires, J. B. *Biochemistry* **2010**, *49*, 179.
(20) Liu, W.; Fu, Y.; Zheng, B.; Cheng, S.; Li, W.; Lau, T. C.; Liang, H. *J. Phys. Chem. B* **2011**, *115*, 13051.
(21) Liu, W.; Zhu, H.; Zheng, B.; Cheng, S.; Fu, Y.; Li, W.; Lau, T. C.; Liang, H. *Nucleic Acids Res.* **2012**, *40*, 4229.
(22) Zhang, A. Y. Q.; Bugaut, A.; Balasubramanian, S. *Biochemistry* **2011**, *50*, 7251.

(23) Joachimi, A.; Benz, A.; Hartig, J. S. *Bioorg. Med. Chem.* **2009**, *17*, 6811.
(24) Zhang, D.-H.; Fujimoto, T.; Saxena, S.; Yu, H.-Q.; Miyoshi, D.; Sugimoto, N. *Biochemistry* **2010**, *49*, 4554.
(25) Wieland, M.; Hartig, J. S. *Chem. Biol.* **2007**, *14*, 757.
(26) Mashima, T.; Matsugami, A.; Nishikawa, F.; Nishikawa, S.; Katahira, M. *Nucleic Acids Res.* **2009**, *37*, 6249.
(27) Mullen, M. A.; Assmann, S. M.; Bevilacqua, P. C. *J. Am. Chem. Soc.* **2012**, *134*, 812.
(28) Morris, M. J.; Negishi, Y.; Pazsint, C.; Schonhoft, J. D.; Basu, S. *J. Am. Chem. Soc.* **2010**, *132*, 17831.
(29) Phan, A. T. *FEBS J.* **2010**, *277*, 1107.
(30) Mergny, J.-L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515.
(31) Eyring, H. *Chem. Rev.* **1935**, *17*, 65.
(32) Nagata, T.; Sakurai, Y.; Hara, Y.; Mashima, T.; Kodaki, T.; Katahira, M. *FEBS J.* **2012**, *279*, 1456.
(33) Sobczak, K.; Michlewski, G.; de Mezer, M.; Kierzek, E.; Krol, J.; Olejniczak, M.; Kierzek, R.; Krzyzosiak, W. J. *J. Biol. Chem.* **2012**, *285*, 12755.
(34) Mergny, J.-L.; Phan, A.-T.; Lacroix, L. *FEBS Lett.* **1998**, *435*, 74.
(35) Misra, V. K.; Draper, D. E. *J. Mol. Biol.* **2002**, *317*, 507.
(36) Weiss, J. N. *FASEB J.* **1997**, *11*, 835.
(37) Solomatin, S. V.; Greenfeld, M.; Herschlag, D. *Nat. Struct. Mol. Biol.* **2011**, *18*, 732.
(38) Lim, K. W.; Amrane, S.; Bouaziz, S.; Xu, W.; Mu, Y.; Patel, D. J.; Luu, K. N.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 4301.
(39) Vorlickova, M.; Kejnovska, I.; Sagi, J.; Renciuk, D.; Bednarova, K.; Motlova, J.; Kypr, J. *Methods* **2012**, *57*, 64.
(40) Woodson, S. A. *Ann. Rev. Biophys.* **2010**, *39*, 61.
(41) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 9963.
(42) Matsugami, A.; Ouhashi, K.; Kanagawa, M.; Liu, H.; Kanagawa, S.; Uesugi, S.; Katahira, M. *J. Mol. Biol.* **2001**, *313*, 255.
(43) Gray, R. D.; Buscaglia, R.; Chaires, J. B. *J. Am. Chem. Soc.* **2012**, *134*, 16834.
(44) Arrhenius, S. *Z. Phys. Chem.* **1889**, *4*, 226.
(45) Oliveberg, M.; Tan, Y. J.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 8926.
(46) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196.
(47) Wallace, M. I.; Ying, L. M.; Balasubramanian, S.; Klenerman, D. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5584.
(48) Zhang, W. B.; Chen, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1931.
(49) Dobson, C. M.; Sali, A.; Karplus, M. *Angew. Chem., Int. Ed.* **1998**, *37*, 868.
(50) Hud, N. V.; Smith, F. W.; Anet, F. A. L.; Feigon, J. *Biochemistry* **1996**, *35*, 15383.
(51) Bodenreider, C.; Kiefhaber, T. *J. Mol. Biol.* **2005**, *351*, 393.
(52) Weikl, T. R.; von Deuster, C. *Proteins* **2009**, *75*, 104.
(53) Kiefhaber, T.; Bachmann, A.; Jensen, K. S. *Curr. Opin. Struct. Biol.* **2012**, *22*, 21.
(54) Fersht, A. R. *Structure and Mechanism in Protein Science*; W. H. Freeman and Company: New York, 1998.
(55) Boncina, M.; Lah, J.; Prislan, I.; Vesnaver, G. *J. Am. Chem. Soc.* **2012**, *134*, 9657.
(56) Mashimo, T.; Yagi, H.; Sannohe, Y.; Rajendran, A.; Sugiyama, H. *J. Am. Chem. Soc.* **2010**, *132*, 14910.
(57) Koirala, D.; Mashimo, T.; Sannohe, Y.; Yu, Z.; Mao, H.; Sugiyama, H. *Chem. Commun.* **2012**, *48*, 2006.
(58) Collie, G. W.; Haider, S. M.; Neidle, S.; Parkinson, G. N. *Nucleic Acids Res.* **2010**, *38*, 5569.
(59) Okamoto, K.; Sannohe, Y.; Mashimo, T.; Sugiyama, H.; Terazima, M. *Bioorg. Med. Chem.* **2008**, *16*, 6873.
(60) Olsen, C. M.; Gmeiner, W. H.; Marky, L. A. *J. Phys. Chem. B* **2006**, *110*, 6962.
(61) Isambert, H. *Methods* **2009**, *49*, 189.
(62) Xayaphoummine, A.; Bucher, T.; Isambert, H. *Nucleic Acids Res.* **2005**, *2005*, W605.
(63) Broccoli, D.; Smogorzewska, A.; Chong, L.; de Lange, T. *Nat. Genet.* **1997**, *17*, 231.
(64) Biffi, G.; Tannahill, D.; Balasubramanian, S. *J. Am. Chem. Soc.* **2012**, *134*, 11974.
(65) Palm, W.; de Lange, T. *Annu. Rev. Genet.* **2008**, *42*, 301.
(66) Xu, Y.; Suzuki, Y.; Ito, K.; Komiyama, M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 14579.
(67) Redon, S.; Reichenbach, P.; Lingner, J. *Nucleic Acids Res.* **2010**, *38*, 5797.
(68) Xu, Y.; Suzuki, Y.; Kaminaga, K.; Komiyama, M. *Nucleic. Acids. Symp. Ser.* **2009**, *53*, 63.
(69) Xu, Y.; Yang, J.; Ito, K.; Ishizuka, T.; Katada, H.; Komiyama, M.; Hayashi, T. *J. Biol. Chem.* [Online early access]. DOI: 10.1074/jbc.M112.34203; http://www.jbc.org/content/early/2012/09/25/jbc.M112.342030 (accessed November 8, 2012).
(70) Wanrooij, P. H.; Uhler, J. P.; Shi, Y.; Westerlund, F.; Falkenberg, M.; Gustafsson, C. M. *Nucleic Acids Res.* **2012**, *40*, 10334.
(71) Tani, H.; Mizutani, R.; Salam, K. A.; Tano, K.; Ijiri, K.; Wakamatsu, A.; Isogai, T.; Suzuki, Y.; Akimitsu, N. *Genome Res.* **2012**, *37*, 6716.
(72) Mullen, M. A.; Olson, K. J.; Dallaire, P.; Major, F.; Assmann, S. M.; Bevilacqua, P. C. *Nucleic Acids Res.* **2010**, *38*, 8149.