

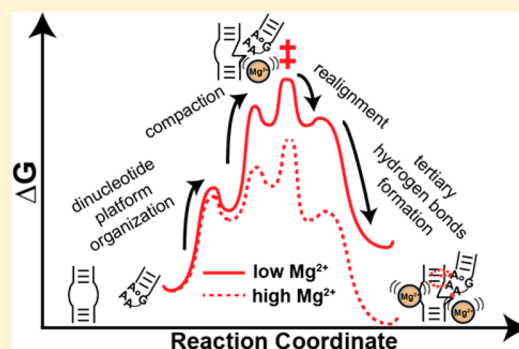
# Single-Molecule Fluorescence Reveals Commonalities and Distinctions among Natural and *in Vitro*-Selected RNA Tertiary Motifs in a Multistep Folding Pathway

Steve Bonilla,<sup>†,§</sup> Charles Limouse,<sup>‡,§</sup> Namita Bisaria,<sup>§,#</sup> Magdalena Gebala,<sup>§</sup> Hideo Mabuchi,<sup>‡</sup> and Daniel Herschlag<sup>\*,†,§,||,⊥</sup>

<sup>†</sup>Department of Chemical Engineering, <sup>‡</sup>Department of Applied Physics, <sup>§</sup>Department of Biochemistry, <sup>||</sup>Department of Chemistry, <sup>⊥</sup>Stanford ChEM-H, Stanford University, Stanford, California 94305, United States

## Supporting Information

**ABSTRACT:** Decades of study of the RNA folding problem have revealed that diverse and complex structured RNAs are built from a common set of recurring structural motifs, leading to the perspective that a generalizable model of RNA folding may be developed from understanding of the folding properties of individual structural motifs. We used single-molecule fluorescence to dissect the kinetic and thermodynamic properties of a set of variants of a common tertiary structural motif, the tetraloop/tetraloop-receptor (TL/TLR). Our results revealed a multistep TL/TLR folding pathway in which preorganization of the ubiquitous AA-platform submotif precedes the formation of the docking transition state and tertiary A-minor hydrogen bond interactions form after the docking transition state. Differences in ion dependences between TL/TLR variants indicated the occurrence of sequence-dependent conformational rearrangements prior to and after the formation of the docking transition state. Nevertheless, varying the junction connecting the TL/TLR produced a common kinetic and ionic effect for all variants, suggesting that the global conformational search and compaction electrostatics are energetically independent from the formation of the tertiary motif contacts. We also found that *in vitro*-selected variants, despite their similar stability at high  $Mg^{2+}$  concentrations, are considerably less stable than natural variants under near-physiological ionic conditions, and the occurrence of the TL/TLR sequence variants in Nature correlates with their thermodynamic stability in isolation. Overall, our findings are consistent with modular but complex energetic properties of RNA structural motifs and will aid in the eventual quantitative description of RNA folding from its secondary and tertiary structural elements.



## INTRODUCTION

Essential and ubiquitous biological functions, including pre-mRNA splicing and translation, require RNA molecules to fold into well-defined structures, bind specific proteins and ligands, and undergo precise conformational changes.<sup>1–4</sup> These properties of RNA are ultimately dictated by its sequence, and the importance of RNA's sequence–structure–function relationship to modern day biology and, likely, to early evolution has motivated the in-depth investigation of the RNA folding problem.<sup>5–9</sup>

Over the past decades, general features of RNA structure have emerged that simplify the dissection of RNA folding. First, the high stability of RNA secondary structure results in hierarchical folding, such that tertiary folding can be considered to arise from a conformational search among largely prefolded secondary structure elements.<sup>10–15</sup> Second, phylogenetic and structural studies have revealed recurring structural motifs in RNA.<sup>16–20</sup> These structural motifs include junction elements, such as four-way junctions and kink turns that steer emanating helices in preferred orientations, and tertiary motifs that

connect regions of RNA distant in primary and secondary structure.<sup>21–23</sup>

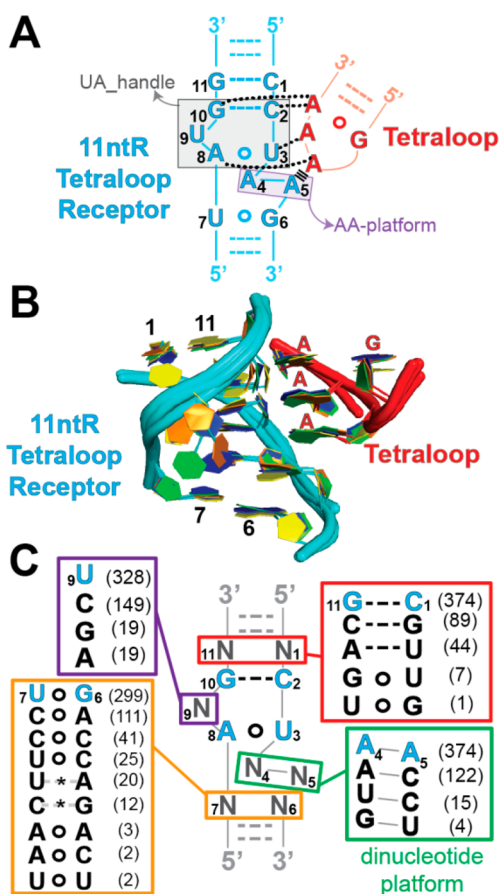
Observation of RNA motifs with superimposable structures embedded in different folded RNAs has led to the perspective that RNA is structurally modular and has suggested that understanding the folding properties of isolated RNA motifs might provide generalizable insights into RNA folding.<sup>12,24–32</sup> Thus, there has been considerable effort to dissect the properties and behaviors of certain common and tractable structural elements. For example, single-molecule and ensemble studies of isolated kink turns and four-way junctions have provided insights into their conformational preferences and dynamics and have revealed a high sensitivity of their folding properties to ionic conditions.<sup>33–36</sup> Particularly relevant to the current study are the pioneering single-molecule FRET (smFRET) studies of the canonical GAAA/11ntR tetraloop/tetraloop-receptor (TL/TLR) in isolation by Nesbitt, Pardi,

Received: August 19, 2017

Published: November 29, 2017

and colleagues.<sup>37–41</sup> Here, we build on these studies to describe the molecular events that underlie the formation of this tertiary structural motif and to provide insights into the kinetic and thermodynamic properties of different sequence variants and how these properties may influence their use in natural structured RNAs.

The GAAA/11ntR TL/TLR is a tertiary structural motif found in group I and II introns, RNase P, and cyclic-di-GMP riboswitches that consists of long-range hydrogen bond and base-stacking interactions between a GAAA tetraloop and an 11 nucleotide tetraloop-receptor (11ntR) (Figure 1A).<sup>16,24,42</sup>



**Figure 1.** Structural modularity and sequence diversity of the GAAA/11ntR tertiary motif. (A) Secondary structure of the canonical 11ntR tetraloop-receptor (cyan) and GAAA tetraloop (red). Watson-Crick (WC) and noncanonical base-pairs are marked as colored dashed lines and circles, respectively. Tertiary interactions are marked in black. UA\_handle<sup>29</sup> and AA-platform<sup>19</sup> submotifs are boxed. Residue numbering is used throughout. (B) Overlay of four crystal structures of the GAAA/11ntR motif. Backbone of tetraloop and tetraloop-receptor are colored as in (A). The structure of the TL/TLRs were extracted from crystal structures of the P4-P6 domain (PDB 1GID,<sup>47</sup> blue), RNase P (PDB 1NBS,<sup>48</sup> green), and *Azoarcus* group I intron (PDB 1ZZN,<sup>49</sup> yellow and orange). PDB 1NBS contains a tetraloop-receptor variant in which the canonical AA-platform is replaced by an AC-platform. (C) Sequence variation in 515 11ntR-like tetraloop-receptors identified in the sequence and secondary structure database of aligned group I introns.<sup>50</sup> Numbers in parentheses correspond to the total number of receptors containing the specified sequence variation. The complete sequences of the variants along with their observed frequencies are provided in Table S1. It is not known whether A-U and G-C at positions 6 and 7 (marked with \*) form canonical WC base-pairs.

Comparison of crystal structures of the GAAA/11ntR embedded within different RNAs show that the three-dimensional structure of the GAAA/11ntR motif is largely independent of structural context (Figure 1B), and solution NMR studies suggest that the crystallographically observed structure of the GAAA/11ntR is also maintained in solution.<sup>43</sup> Although most studies have focused on the canonical sequence of the GAAA/11ntR TL/TLR motif, phylogenetic comparisons of structured RNAs reveal a wealth of variants of the 11ntR tetraloop-receptor (Figure 1C),<sup>16,44</sup> and additional variants have been obtained through *in vitro* selection.<sup>44–46</sup>

We used smFRET to dissect the kinetic and thermodynamic properties of a set of sequence variants of the GAAA/11ntR in isolation, including variants found in natural structured RNAs and variants evolved previously by *in vitro* selection. Our results, in combination with prior studies of the canonical GAAA/11ntR, revealed multiple steps in the TL/TLR folding pathway. Differences in the ion sensitivity of the docking and undocking rate constants revealed sequence-dependent local conformational rearrangements prior to and after the formation of the docking transition state. These rearrangements and their accompanying electrostatic effects are separable from a common global conformational search and the electrostatics associated with global compaction. Finally, comparison of the thermodynamic stability of the TL/TLR variants at near-physiological ionic conditions provided insights into the relative occurrence of TL/TLR sequences in Nature.

## MATERIALS AND METHODS

**smFRET Instrumentation.** smFRET data were collected by two independent methods using (1) a conventional prism-based TIRF microscope equipped with an Andor Ultra iXon 897 EMCCD camera and (2) a confocal microscope equipped with avalanche photodiodes for single-photon counting (Supporting Methods). For the first setup, time-lapse images were recorded at an exposure time of 5 ms. Control experiments with varying acquisition frame rates suggested a slight underestimation (~15%) of the rate constants measured by the camera at conditions in which the dynamics were greater than ~70 s<sup>-1</sup>. We therefore used single-photon counting for conditions in which TL/TLR docking dynamics were expected to exceed 70 s<sup>-1</sup>. For these conditions, single-photon arrival times were recorded with ns accuracy and binned offline at 1 ms (Supporting Methods). Herein, data points obtained with the single-photon counting setup are marked with an internal “X” (see Figure 4 for example). Excellent agreement between data collected by camera (5 ms) and single-photon counting (1 ms) (Tables S4–S9) at overlapping conditions indicates that there were no significant time binning effects in the data analyzed (see Figure 4 for example), as did analysis of the overall relaxation rate constant ( $k_{\text{dock}} + k_{\text{undock}}$ ) obtained directly from photon counting (data not shown).

**Preparation of Constructs for smFRET Experiments.** Synthetic oligonucleotides (Table S2) with 5' amino-C6 modifications were purchased from Integrated DNA Technologies, and a standard protocol was followed to incorporate *N*-hydroxysuccinamide (NHS) functionalized Cy3B and Cy5 (GE Healthcare) fluorophores into the oligos.<sup>51</sup> After labeling, the oligos were purified by denaturing PAGE without the use of UV shadowing, which can damage RNA and result in artifactual covalent heterogeneity.<sup>52,53</sup> Prior to smFRET experiments, the corresponding oligos were mixed at 1 μM and annealed by heating to 70 °C for 3 min in 100 mM NaMOPS, pH 7.0, and 0.1 mM EDTA, followed by slow cooling to 4 °C at 0.1 °C/sec.

**smFRET Experiments.** Imaging buffers contained specified concentration of MgCl<sub>2</sub>, BaCl<sub>2</sub>, KCl, NaCl, and/or RbCl in addition to 50 mM MOPS, pH 8.0 (titrated with KOH, NaOH, or RbOH), 0.01 mM NaEDTA, 1 mM TROLOX, 2 mM protocatechuic acid, and 0.001 units/μL protocatechuate-3,4-dioxygenase. All smFRET experiments were performed at 23 ± 2 °C. Before taking smFRET

measurements at a specified salt condition, the flow channel was washed with 20 volumes of buffer containing the specified salt concentration followed by 2 volumes of imaging buffer with the specified salt concentration.

**Salt Activity Calculations.** The mean activity coefficients ( $\gamma_{\pm}$ ) of the monovalent salt solutions were calculated as described by Lepply et al.,<sup>54</sup> using the tables of empirical salt activity coefficients reported by Robinson and Stokes.<sup>55</sup> As there are no activity estimates for mixed chloride and MOPS salts, we calculated the mean activity using the total  $[M^+]$  as if it were present solely as the chloride salt; because the MOPS salts were present at low concentrations (40 mM cation), we expect this effect to be insignificant.

**Data Analysis.** Single-molecule traces were selected based on anticorrelation of donor and acceptor intensities, consistent total intensity throughout the trace, and single step photobleaching as previously described.<sup>25</sup> Consistent with prior smFRET studies of the canonical GAAA/11ntR, ~30% of the single molecules did not show FRET transitions within the observation time window.<sup>37,38</sup> This inactive population may be caused by synthesis errors and/or chemical damage, although a population of molecules trapped in an inactive conformation cannot be ruled out. Inactive molecules were removed from further analysis as done previously.<sup>38</sup>

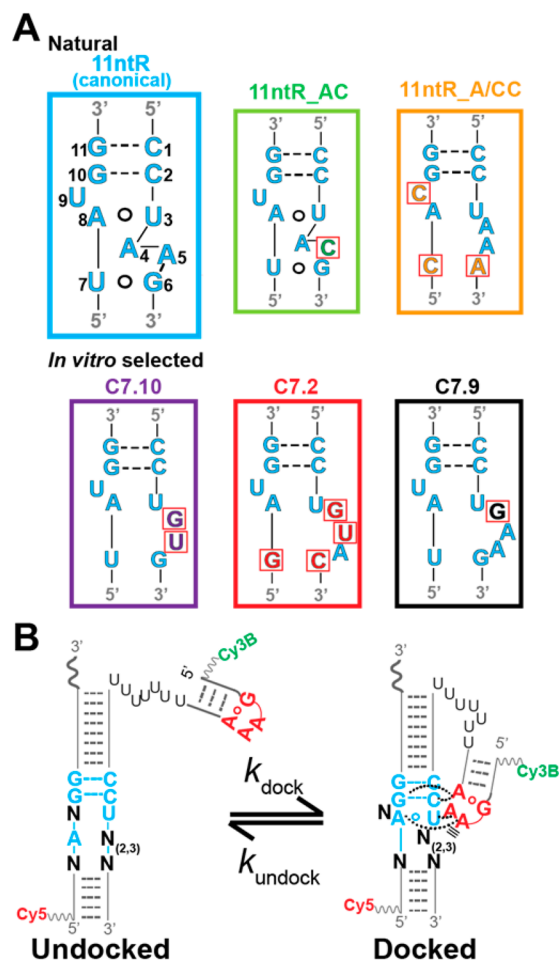
Bayesian analysis (Table S3) indicated that the smFRET traces are best described by two FRET states, consistent with previous observations of the canonical GAAA/11ntR smFRET minimal construct.<sup>37,40</sup> To extract kinetics, each single-molecule trace was fit to a two-state hidden Markov model (HMM) using the Single Molecule Analysis Research Tool (SMART), which fits the donor and acceptor intensities directly, as described previously.<sup>25,56–58</sup> Fitting the intensities directly allows for appropriately accounting for the noise in the donor and acceptor channels and reduces the amount of required preprocessing of the data. The rate and equilibrium constants and their errors for each construct and each solution condition are summarized in Tables S4–S9. smFRET data summaries for all TL/TLR variants and conditions are provided (Figures S9–S275). All raw data are available for download.<sup>59</sup>

The agreement between data obtained with the camera (5 ms binning) and single-photon counting (1 ms binning) setups at overlapping conditions demonstrated that there were no significant time binning effects (see for example Figure 4). Because of differences in background subtraction between the camera and the single-photon counting setups, the apparent average FRET states differed between the two experimental setups. There were sufficient overlapping data to indicate that the low FRET state value of  $0.28 \pm 0.01$  with the camera setup corresponded to an apparent FRET state value of  $0.44 \pm 0.01$  with the single-photon counting setup, and similarly for the high FRET state,  $0.57 \pm 0.01$  in camera setup corresponded to an apparent FRET state of  $0.68 \pm 0.02$  in the single-photon counting setup.

**Identification of Natural Variants of the GAAA/11ntR in Group I Introns.** The sequence alignments contained in the Group I Intron Sequence and Structure Database (GISSD)<sup>50</sup> were searched for 11ntR-like tetraloop-receptors at locations that are predicted to form TL/TLR interactions: LSb-P6, L9-P5, and L2-P8.<sup>24,60</sup> Based on previous structural characterization of the GAAA/11ntR and its component submotifs,<sup>29,44,45,61</sup> 11ntR-like tetraloop-receptors were defined as sequences containing (1) a putative UA\_handle submotif which consists of a noncanonical U-A base-pair that stacks over a Watson–Crick (WC) base-pair with an intervening bulging residue (Figure 1A),<sup>29</sup> (2) a conserved C<sub>2</sub>-G<sub>10</sub> base-pair, and (3) an internal asymmetrical loop capable of forming a dinucleotide platform submotif. For simplicity and to avoid ambiguity in the selection, sequences with insertions and deletions were not considered.

A list of all possible single to quintuple mutants of the canonical 11ntR were generated keeping C<sub>2</sub>-G<sub>10</sub> and U<sub>3</sub>-A<sub>8</sub> constant and with a WC or wobble base-pair at positions 1–11 (Figure 1C). These sequences were searched for in the P5, P6, and P8 helices of the 1789 group I introns contained in the database. To eliminate false positives, the criteria for 11ntR-like tetraloop-receptors described above were verified by visually inspecting the predicted secondary structures<sup>50</sup> of the introns containing the putative tetraloop-receptors. The database

was also searched for the *in vitro*-selected tetraloop-receptors C7.10, C7.2, and C7.9 described below (Figure 2A). These *in vitro*-selected

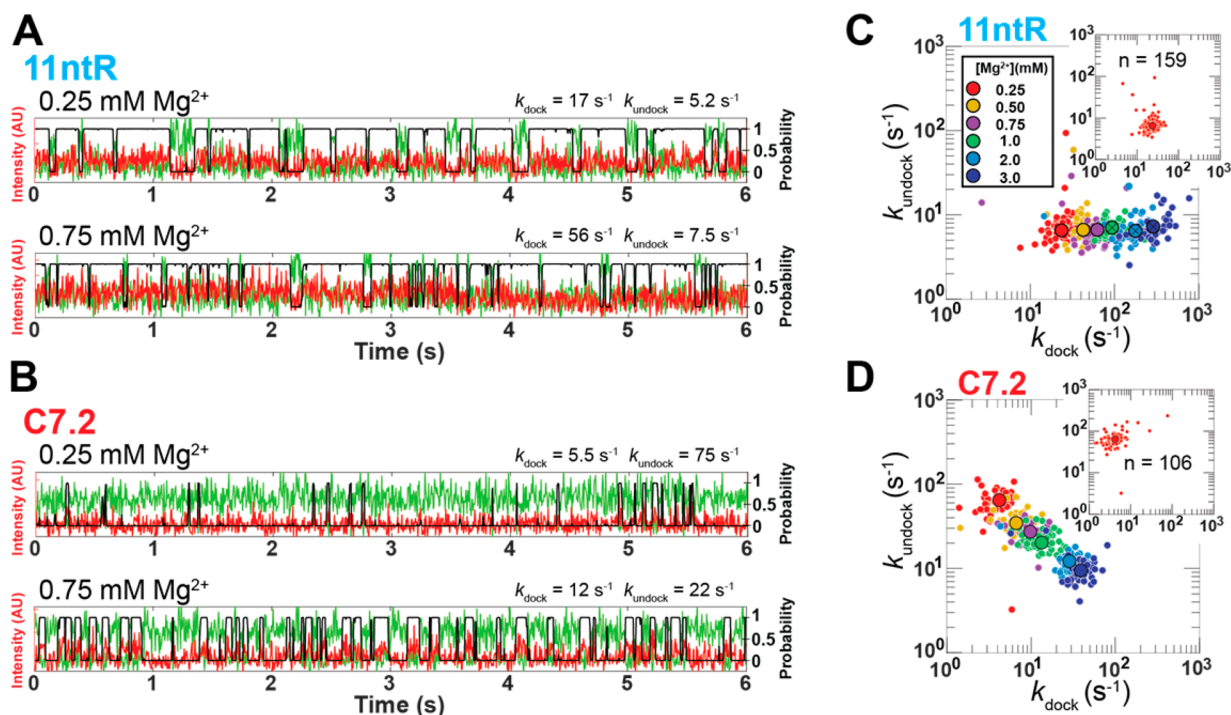


**Figure 2.** Natural and *in vitro*-selected TL/TLR variants were inserted into a minimal U<sub>7</sub>-tethered smFRET construct. (A) Sequence of the six tetraloop-receptor variants investigated. Residues that differ from the canonical 11ntR are boxed. Secondary structures of 11ntR and 11ntR\_AC in the docked state are known experimentally.<sup>47,48</sup> (B) Schematic of smFRET construct in which tetraloop-receptors in panel A were embedded. GAAA tetraloop (red) and residues in cyan are common to all of the TL/TLR variants investigated. This smFRET construct has been used in previous studies to characterize the canonical GAAA/11ntR.<sup>38,40</sup> The complete sequence of all constructs is provided in Table S2.

motifs were not found in the group I introns investigated, consistent with previous observations of the absence of these motifs in natural structured RNAs.<sup>24</sup>

## RESULTS

**Choice of TL/TLR Variants.** Comparison of the available sequences of group I and II introns, RNase P, and cyclic-di-GMP riboswitches reveals substantial sequence variation in the TL/TLR motif.<sup>44,50,62–65</sup> To examine this variation, we analyzed 1789 aligned group I intron sequences and found 515 identifiable 11ntR-like tetraloop-receptors. Figure 1C summarizes the most frequent base pairs, dinucleotide platforms, and residues that comprise the 11ntR-like tetraloop-receptors, and Table S1 lists the individual sequences. From these natural tetraloop-receptor variants, we chose to study the canonical 11ntR as well as tetraloop-receptors



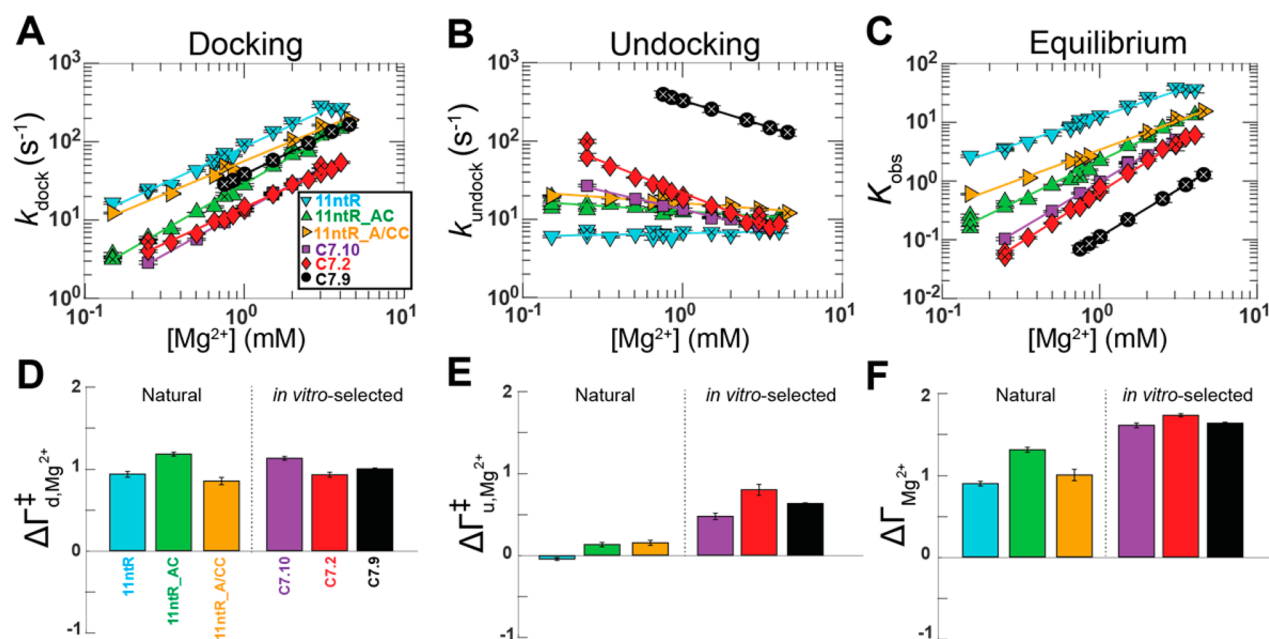
**Figure 3.** smFRET reveals distinct kinetic behaviors and  $\text{Mg}^{2+}$  dependences of TL/TLR in the  $U_7$ -tether minimal constructs (Figure 2). Sample smFRET traces of variants 11ntR (A) and C7.2 (B) at two  $\text{Mg}^{2+}$  concentrations. Green and red traces are donor and acceptor intensities respectively and black trace represents the probability of occupying the high FRET state as determined by a two-state hidden Markov model (HMM). The values of the fitted  $k_{\text{dock}}$  and  $k_{\text{undock}}$  for the sample traces are displayed above each trace. Traces were truncated at 6 s for ease of comparison. Sample traces for all variants and conditions are shown in Supporting Information (Figures S9–S275) and raw data are available to download.<sup>59</sup> (C–D) Median  $k_{\text{dock}}$  and  $k_{\text{undock}}$  values (large circles) for a population of individual molecules (small circles) at a range of  $\text{Mg}^{2+}$  concentrations (colors) for 11ntR (C) and C7.2 (D). For clarity, only a subset of  $\text{Mg}^{2+}$  concentrations and 50 randomly selected molecules at each  $[\text{Mg}^{2+}]$  are shown. Typically more than 100 molecules were measured at each  $\text{Mg}^{2+}$  concentration. The complete population of molecules at 0.25 mM  $\text{Mg}^{2+}$  is shown in the inset to better display the kinetic homogeneity. Data from measurements carried out by single-photon counting are marked with an internal “X”.

11ntR\_AC and 11ntR\_A/CC (Figure 2A), as they contain some of the most common mutations relative to the canonical 11ntR and have been shown to fold outside of their natural context.<sup>65</sup> Variant 11ntR\_AC differs from the canonical 11ntR only by the replacement of the canonical AA-platform with an AC-platform, and has a nearly identical three-dimensional conformation (Figure 1B and Figure S1). Variant 11ntR\_ACC differs from the canonical 11ntR by the replacement of the canonical  $G_6 \cdot U_7$  by  $A_6 \cdot C_7$  and replacement of the bulging  $U_9$  by  $C_9$ . Based on their position within the TL/TLR (Figure 1A) and their high occurrence in group I intron sequence variants (Figure 1C), these mutations are not expected to substantially alter the three-dimensional structure of the TL/TLR relative to the canonical GAAA/11ntR.

Several additional 11ntR-like tetraloop-receptor variants have been discovered through *in vitro* selection.<sup>44–46</sup> Of these, we chose C7.10, C7.2, and C7.9 for our studies (Figure 2A).<sup>44</sup> Previous studies showed that C7.10 and C7.2 bind to GAAA tetraloops with affinities comparable to that of the canonical 11ntR tetraloop-receptor within an engineered artificial dimer construct.<sup>45</sup> Nevertheless, they have not been found in natural RNAs.<sup>24,45,65</sup> Variant C7.10 differs from the canonical 11ntR by the replacement of the canonical AA-platform with a GU-platform. GAAA/11ntR and GAAA/C7.10 are expected to form similar tertiary interactions and conformations as they can both be inserted into the same structural scaffolds and have shown identical patterns of discrimination for different tetraloop sequences.<sup>45</sup> A computationally predicted structure supported by experimental chemical probing suggests that the

tertiary interactions and global conformations of GAAA/C7.2 and GAAA/11ntR are similar, but that the insertion of an additional residue into the platform of C7.2 alters the conformation of the residues around the platform of these two motifs (Figure S2).<sup>66</sup> We also studied the *in vitro*-selected C7.9 tetraloop-receptor that had not previously been characterized but differs in sequence relative to the canonical 11ntR only by the insertion of a G residue between positions 3 and 4 relative to the canonical 11ntR (Figure 2A). In total, we studied six tetraloop-receptor variants, three natural (11ntR, 11ntR\_AC, and 11ntR\_ACC) and three *in vitro*-selected (C7.10, C7.2, and C7.9) (Figure 2A).

**Measuring TL/TLR Docking Kinetics and Thermodynamics by smFRET.** The docking/undocking properties of the TL/TLR variants were investigated within the context of a minimal smFRET construct designed and used previously by Nesbitt, Pardi, and co-workers to characterize the folding properties of the canonical GAAA/11ntR.<sup>38,40</sup> In this construct (Figure 2B), the GAAA tetraloop is connected to its cognate tetraloop-receptor by a flexible  $U_7$  linker and TL/TLR docking/undocking is monitored by time-dependent fluctuations in FRET efficiency. Models with multiple FRET states were evaluated and Bayesian statistical analysis indicated that the FRET traces are best described by two FRET states, a low FRET state corresponding to the undocked TL/TLR, and a high FRET state corresponding to the docked TL/TLR (Table S3), consistent with previous observations for the canonical GAAA/11ntR.<sup>37,40</sup> Intermediates with too-short lifetimes or with FRET values only slightly different from those of the



**Figure 4.** Docking kinetics, thermodynamics and  $Mg^{2+}$  dependence of TL/TLR variants in  $U_7$ -tether minimal construct. (A–C) Median  $k_{dock}$ ,  $k_{undock}$  and  $K_{obs}$  values for each TL/TLR variant at a range of  $Mg^{2+}$  concentrations. All measurements were carried out in a background of 140 mM  $K^+$ . Data and errors are summarized in Table S4. Standard errors calculated from bootstrapping were smaller than the size of the symbols. Data points marked with an internal “X” were determined using single-photon counting. Linear fits were determined by least-squares.  $Mg^{2+}$  uptake prior to and after formation of the docking transition state (D and E, respectively) and net  $Mg^{2+}$  uptake (F) were obtained from the linear fits in panels A–C as given by eq 1 and eqs 2A and 2B.

docked and undocked states and with similar lifetimes would not be distinguishable as additional FRET states due to measurement limits and noise.

Average low and high FRET values were within error between the different TL/TLR variants and nearly constant across all salt types and concentrations investigated (Figure S3), suggesting similar global architectures of the TL/TLRs within the resolution of smFRET. Further, the absence of significant changes in the average FRET state values across ionic conditions is consistent with two-state folding behavior. Nevertheless, our results provide evidence for multiple intermediates along the folding pathway that do not accumulate, as described below.

Rate constants for TL/TLR docking and undocking,  $k_{dock}$  and  $k_{undock}$ , respectively, were extracted from the fluorescence traces of individual molecules using hidden Markov models (HMMs) as described previously.<sup>25,56,58,67</sup> Representative traces and HMM fits at two different ionic conditions for tetraloop-receptor variants 11ntR and C7.2 are shown in Figure 3A,B. For each construct and condition, sample traces and complete data summaries are presented in Supporting Information (Figures S9–S275). The docking equilibrium constant ( $K_{obs}$ ) for TL/TLR formation was determined from the ratio  $k_{dock}/k_{undock}$  and agreed with values obtained from integration of the areas under cumulative FRET distributions (Tables S4–S9). Our measurements of  $k_{dock}$  and  $k_{undock}$  for the canonical GAAA/11ntR agreed within 2-fold with those previously reported.<sup>38,40</sup>

**Individual TL/TLR Variants Displayed Homogeneous Molecular Behavior.** Kinetic heterogeneity in the folding and/or activity of structured RNAs has been reported previously and can arise from the coexistence of chemically identical RNA molecules with distinct conformations based on the interconversion of kinetic properties of individual RNA

molecules.<sup>68–70</sup> All of our TL/TLR variants displayed homogeneous kinetic behavior, with tightly clustered values of  $k_{dock}$  and  $k_{undock}$  under all conditions, consistent with previous observations for the canonical GAAA/11ntR.<sup>25</sup> The insets of Figure 3C,D show this behavior for tetraloop-receptor variants 11ntR and C7.2 at a single solution condition (0.25 mM  $Mg^{2+}$  and 140 mM  $K^+$ ), and analogous clusters for each TL/TLR variant under each solution condition tested are provided (Figures S9–S275). The observed homogeneous kinetic behavior provides no indication of conformational heterogeneity at the level of individual TL/TLR motifs and allowed us to precisely determine average  $k_{dock}$  and  $k_{undock}$  values and their dependences on salt concentration.

Figure 3C,D for tetraloop-receptors 11ntR and C7.2, respectively, illustrates these data and some of the differences in kinetic behavior that were observed. For example, at 0.25 mM  $Mg^{2+}$  (red points)  $k_{dock}$  was 6-fold larger for 11ntR than for C7.2 ( $23.4 \pm 0.6 \text{ s}^{-1}$  vs  $4.2 \pm 0.3 \text{ s}^{-1}$ ) and  $k_{undock}$  was 10-fold smaller ( $6.5 \pm 0.2$  vs  $65 \pm 7 \text{ s}^{-1}$ ). The  $Mg^{2+}$  dependence of the kinetic behavior also clearly differed between TL/TLR variants, with  $k_{undock}$  essentially constant across  $[Mg^{2+}]$  for tetraloop-receptor 11ntR but decreasing sharply for tetraloop-receptor C7.2. These differences in docking kinetics and  $Mg^{2+}$  dependences indicate differences in the folding pathway and electrostatic properties of the TL/TLR variants and motivated us to more closely examine and compare the kinetics, thermodynamics, and salt-dependence of the TL/TLR variants with the objective of developing physical models for the observed sequence effects on TL/TLR docking behavior and energetics.

**Comparison of Docking Kinetics and Thermodynamics of the TL/TLR Variants in  $Mg^{2+}$ .** The values of  $k_{dock}$ ,  $k_{undock}$ , and  $K_{obs}$  were determined for each of the six TL/TLR variants across a range of  $Mg^{2+}$  concentrations (Figure 4A–C).

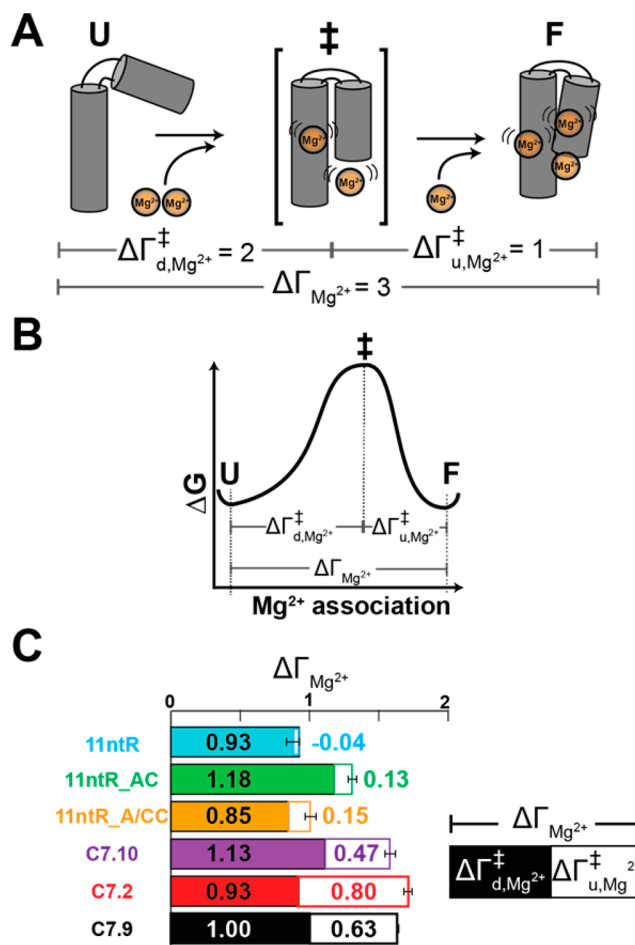
All  $k_{\text{dock}}$  values were slow (range of  $\sim 3\text{--}300\text{ s}^{-1}$ ) relative to an estimated rate of collision of  $\sim 10^5\text{--}10^6\text{ s}^{-1}$  between the tetraloop and the tetraloop-receptor based on measurements of end-to-end collisions in short single stranded oligonucleotides analogous to the flexible  $U_7$  tether joining the TL/TLR.<sup>71</sup> The relatively slow  $k_{\text{dock}}$  values suggest that all six tetraloop-receptor variants must undergo substantial unfavorable conformational reorganization prior to the formation of the docking transition state, as suggested previously for the 11ntR,<sup>41</sup> and differences in the magnitude of  $k_{\text{dock}}$  between variants indicate sequence-dependent differences in this reorganization and/or in the contacts that need to form prior to the docking transition state. Increasing  $\text{Mg}^{2+}$  substantially increased  $k_{\text{dock}}$  and did so similarly for all of the TL/TLR variants as evidenced by the near-parallel slopes in Figure 4A. In contrast, the  $\text{Mg}^{2+}$  dependence of  $k_{\text{undock}}$  was shallower than that for  $k_{\text{dock}}$  and exhibited larger variations in slope between TL/TLR variants (Figure 4B). Undocking of the *in vitro*-selected tetraloop-receptors (C7.10, C7.2, and C7.9; purple, red, and black symbols, respectively) exhibited stronger dependence on  $\text{Mg}^{2+}$  concentration than the naturally occurring variants.

Figure 4C shows the docking equilibrium,  $K_{\text{obs}}$ , for each variant across the range of  $\text{Mg}^{2+}$  concentrations. The variations in slope translate into different relative stabilities at different  $\text{Mg}^{2+}$  concentrations. Interestingly, the stabilities were more similar between TL/TLR variants at higher  $\text{Mg}^{2+}$ . Overall, the different  $\text{Mg}^{2+}$ -dependences of  $k_{\text{dock}}$ ,  $k_{\text{undock}}$  and  $K_{\text{obs}}$  reveal complexities in the folding behavior of the TL/TLR variants. Before exploring this further, we first lay out the classic thermodynamic framework for the interpretation of  $\text{Mg}^{2+}$  uptake in the following section.

**Interpreting  $\text{Mg}^{2+}$  Effects on  $K_{\text{obs}}$ ,  $k_{\text{dock}}$ , and  $k_{\text{undock}}$ .** RNA's high negative charge density leads to the formation of a dynamic sheath of ions around the RNA, referred to as the ion atmosphere, in which cations are accumulated and anions are depleted relative to bulk solution.<sup>72–76</sup> Conformational changes, such as a folding transition, result in ions being taken up or released from the ion atmosphere to compensate for changes in the charge density of RNA (Figure 5). According to a thermodynamic framework described by Record, Draper and co-workers,<sup>54,77</sup> the relationship between  $\text{Mg}^{2+}$  uptake and the thermodynamic equilibrium for an RNA conformational transition is given by the following:

$$\Delta\Gamma_{\text{Mg}^{2+}} = \frac{\partial \log K_{\text{obs}}}{\partial \log [\text{Mg}^{2+}]} \quad (1)$$

where  $\Delta\Gamma_{\text{Mg}^{2+}}$  is the average number of  $\text{Mg}^{2+}$  ions taken into the ion atmosphere as the RNA undergoes a conformational transition and  $K_{\text{obs}}$  is the apparent equilibrium constant for that conformational transition. This framework makes no assumptions about the nature of the RNA- $\text{Mg}^{2+}$  interactions and therefore cannot distinguish between nonspecific electrostatic interactions and specific binding.<sup>54,78</sup> The physical interpretation of eq 1 is valid when the  $\text{Mg}^{2+}$  titration is carried out in a high background concentration of monovalent salt such that the total concentration of anions and the activity coefficient of  $\text{Mg}^{2+}$  do not change significantly as the  $\text{Mg}^{2+}$  salt is added. Previous studies have shown that a ratio of 30:1  $\text{K}^+:\text{Mg}^{2+}$  is sufficient to meet these requirements,<sup>79</sup> and our measurements with  $\text{Mg}^{2+}$  concentrations ranging from 0.15 to 4.5 mM in a background of 140 mM  $\text{K}^+$  (100 mM KCl and 40 mM K-MOPS) meet these requirements. Thus, the slopes of the linear



**Figure 5.** Physical interpretation of the sensitivity of  $K_{\text{obs}}$ ,  $k_{\text{dock}}$ , and  $k_{\text{undock}}$  to  $\text{Mg}^{2+}$ . (A) Schematic of a model RNA undergoing a folding transition from an extended unfolded (U) state, through a compact transition state ( $\ddagger$ ), to a final folded (F) state.  $\text{Mg}^{2+}$  is taken up to compensate for changes in charge density as the RNA folds. To maintain charge neutrality,  $\text{Mg}^{2+}$  uptake must be accompanied by the uptake of anions and/or release of  $\text{K}^+$  into ion atmosphere (not shown). The net number of  $\text{Mg}^{2+}$  taken up ( $\Delta\Gamma_{\text{Mg}^{2+}}$ ) is the sum of  $\text{Mg}^{2+}$  taken up prior to ( $\Delta\Gamma_{\text{d,Mg}^{2+}}^{\ddagger}$ ) and after ( $\Delta\Gamma_{\text{u,Mg}^{2+}}^{\ddagger}$ ) formation of the transition state ( $\ddagger$ ;  $\Delta\Gamma_{\text{Mg}^{2+}} = \Delta\Gamma_{\text{d,Mg}^{2+}}^{\ddagger} + \Delta\Gamma_{\text{u,Mg}^{2+}}^{\ddagger}$ ). (B) Folding process in panel A is represented in a free-energy diagram. (C)  $\text{Mg}^{2+}$  uptake of TL/TLR variants in  $U_7$  smFRET construct. Overall bar gives  $\Delta\Gamma_{\text{Mg}^{2+}}$  and is divided into contributions from  $\Delta\Gamma_{\text{d,Mg}^{2+}}^{\ddagger}$  (filled) and  $\Delta\Gamma_{\text{u,Mg}^{2+}}^{\ddagger}$  (open). Error bars are standard errors of  $\Delta\Gamma_{\text{Mg}^{2+}}$ .

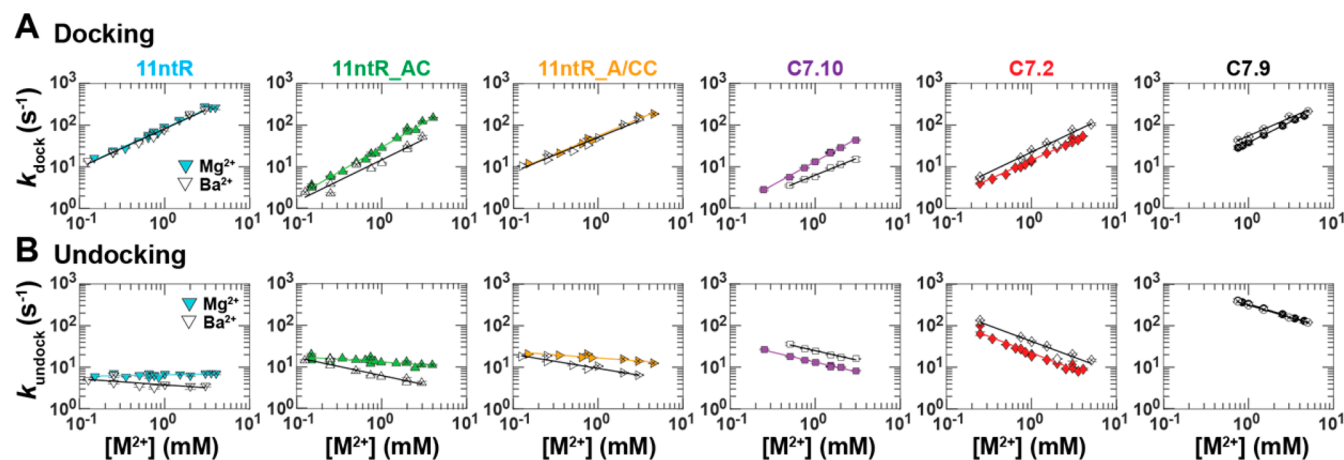
fits shown in Figure 4C correspond to the net number of  $\text{Mg}^{2+}$  ions taken up upon TL/TLR docking.

By measuring the  $\text{Mg}^{2+}$  dependence of the docking and undocking rate constants the net uptake of  $\text{Mg}^{2+}$  can be broken down into  $\text{Mg}^{2+}$  taken up prior to and after the formation of the docking transition state using the following relationships:

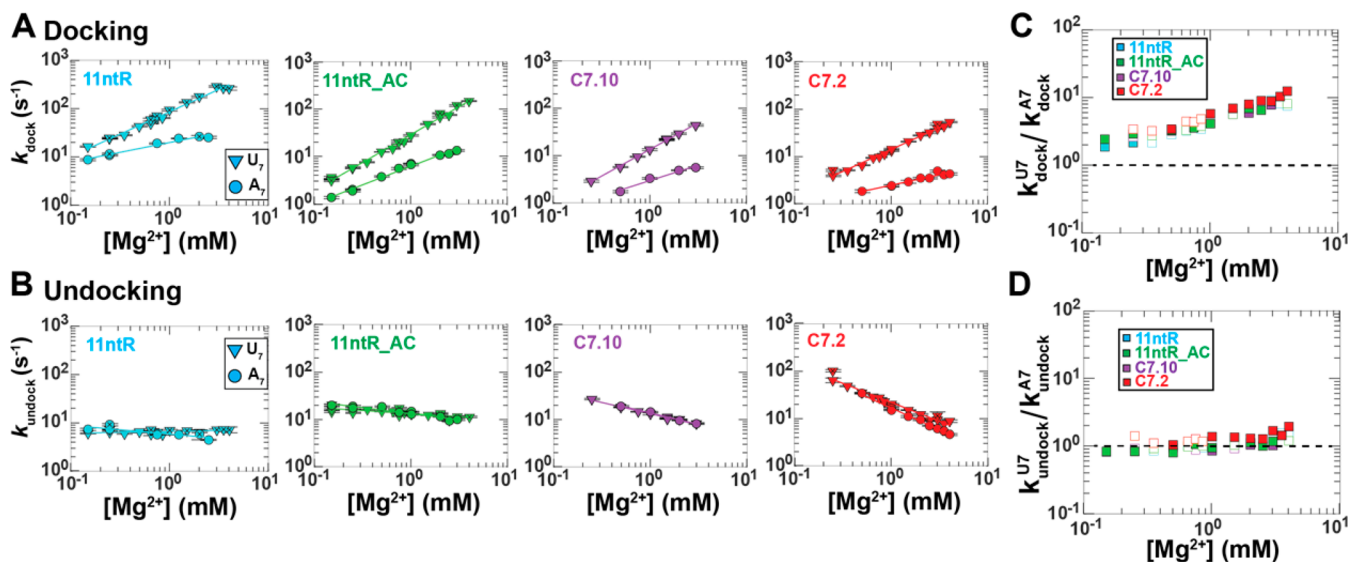
$$\Delta\Gamma_{\text{d,Mg}^{2+}}^{\ddagger} = \frac{\partial \log k_{\text{dock}}}{\partial \log [\text{Mg}^{2+}]} \quad (2A)$$

$$\Delta\Gamma_{\text{u,Mg}^{2+}}^{\ddagger} = - \frac{\partial \log k_{\text{undock}}}{\partial \log [\text{Mg}^{2+}]} \quad (2B)$$

where  $\Delta\Gamma_{\text{d,Mg}^{2+}}^{\ddagger}$  is the average number of  $\text{Mg}^{2+}$  ions taken up prior to the formation of the docking transition state and  $\Delta\Gamma_{\text{u,Mg}^{2+}}^{\ddagger}$  is the average number of  $\text{Mg}^{2+}$  ions taken up after the formation of the docking transition state as illustrated



**Figure 6.** Comparison of TL/TLR docking kinetics in Mg<sup>2+</sup> and Ba<sup>2+</sup> reveals weak but consistent sequence-dependent discrimination for different types of divalent cations. Median  $k_{\text{dock}}$  (A) and  $k_{\text{undock}}$  (B) values over a range of Mg<sup>2+</sup> (filled symbols) and Ba<sup>2+</sup> (open symbols) for each TL/TLR variant. Mg<sup>2+</sup> values are reproduced from Figure 4. Data and errors are summarized in Tables S4 and S5. Standard errors calculated from bootstrapping were smaller than the size of the symbols. Linear fits were determined by least-squares. All measurements were carried out in a background of 140 mM K<sup>+</sup>. Data collected with single-photon counting setup are marked with an internal “X”.



**Figure 7.** Conformational preferences and Mg<sup>2+</sup>-dependence of connecting tether produces a common kinetic and electrostatic effect on TL/TLR docking. (A,B) Median  $k_{\text{dock}}$  (A) and  $k_{\text{undock}}$  (B) values of four TL/TLR variants as a function of Mg<sup>2+</sup> in a background of 140 mM K<sup>+</sup> with an A<sub>7</sub> (circles) or U<sub>7</sub> (triangles) tether connecting the tertiary contact partners. Data and errors are summarized in Tables S4 and S6. Standard errors calculated from bootstrapping were smaller than the size of the symbols. Data collected with single-photon counting setup are marked with an internal “X”. Linear fits were determined by least-squares. (C–D) Ratio of the  $k_{\text{dock}}$  (C) and  $k_{\text{undock}}$  (D) values with the U<sub>7</sub> and A<sub>7</sub> tethers. Open symbols were obtained from interpolation using the linear fits in panels A and B. Dashed lines at a constant value of one are shown for reference.

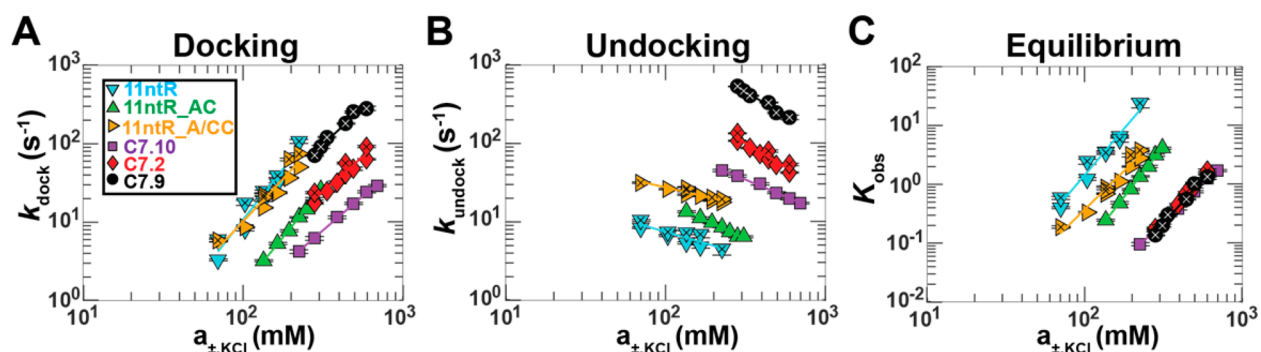
schematically in Figure 5A.<sup>80,81</sup> The values of  $\Delta\Gamma_{\text{d},\text{Mg}^{2+}}^{\ddagger}$  and  $\Delta\Gamma_{\text{u},\text{Mg}^{2+}}^{\ddagger}$  are given by the slopes of the linear fits shown in Figure 4A,B, respectively. For simplicity, RNA is represented as a static molecule and folding is depicted as a simple two-state process with an integer number Mg<sup>2+</sup> taken up in Figure 5A. In reality, the statistical nature of the ion atmosphere and RNA conformational ensembles typically leads to a noninteger Mg<sup>2+</sup> uptake.<sup>82</sup>

The linearity displayed in Figure 4A–C for all TL/TLR variants indicate that the number of Mg<sup>2+</sup> ions taken up is constant across the range of Mg<sup>2+</sup> concentrations investigated and suggests that for a given TL/TLR variant the docking process involves essentially the same undocked, transition, and docked states across Mg<sup>2+</sup> concentrations. Figure 4D–F

summarizes the values of  $\Delta\Gamma_{\text{Mg}^{2+}}$ ,  $\Delta\Gamma_{\text{d},\text{Mg}^{2+}}^{\ddagger}$ , and  $\Delta\Gamma_{\text{u},\text{Mg}^{2+}}^{\ddagger}$  obtained from the linear fits in Figure 4A–C.

Differences in Mg<sup>2+</sup> uptake between TL/TLR variants suggest differences in the electrostatic component of their folding energy landscape, as shown schematically in Figure 5B and discussed below.

**Using Ba<sup>2+</sup> to Test for Specific Ion Effects.** The nature of the interactions between RNA and cations is diverse, ranging from long-range nonspecific electrostatic interactions to specific interactions in which the cation is chelated by RNA ligands.<sup>76</sup> In some cases it is possible to infer ion-specific binding sites in RNA by measuring folding or catalytic activity in the presence of ions of different size or charge.<sup>74,83</sup> To learn more about the specificity of the Mg<sup>2+</sup> effects observed above we measured docking kinetics of the TL/TLR variants in the presence of



**Figure 8.** Docking kinetics and thermodynamics of TL/TLR variants in KCl. To account for nonideal electrolyte interactions, the data are plotted as a function of the mean activity ( $a_{\pm}$ ) of the salt solution. (A–C) Median  $k_{\text{dock}}$ ,  $k_{\text{undock}}$ , and  $K_{\text{obs}}$  values for each of the TL/TLR variants over a range of KCl concentrations. Coloring scheme is shown in panel A. Data and errors are summarized in Table S7. Standard errors determined from bootstrapping were smaller than the size of the symbols. Data marked with an internal “X” were determined by single-photon counting. Linear fits were determined by least-squares.

$\text{Ba}^{2+}$ . Due to the differences in size and coordination preferences of  $\text{Mg}^{2+}$  and  $\text{Ba}^{2+}$ ,  $\text{M}^{2+}$ -specific binding sites are expected to typically discriminate between these two divalent cations. For example, the metal core of the P4–P6 domain of the *Tetrahymena* group I intron discriminates between  $\text{Mg}^{2+}$  and  $\text{Ba}^{2+}$  by 2 orders of magnitude.<sup>84</sup>

Differences in  $k_{\text{dock}}$  and  $k_{\text{undock}}$  between  $\text{Mg}^{2+}$  and  $\text{Ba}^{2+}$  (Figure 6) were small for all variants (less than 3-fold), suggesting that the observed  $\text{M}^{2+}$  effects on  $k_{\text{dock}}$  and  $k_{\text{undock}}$  are largely due to nonspecific electrostatic interactions between the RNA and the divalent cations. The largely nonspecific nature of  $\text{M}^{2+}$  interactions with the TL/TLRs is consistent with the observation that the TL/TLR variants fold in the absence of divalent cations when enough  $\text{M}^+$  is present, as described below. However, small consistent differences between  $\text{Mg}^{2+}$  and  $\text{Ba}^{2+}$  that depended on the sequence of the TL/TLR variant were observed. For example, whereas  $\text{Ba}^{2+}$  decreased  $k_{\text{undock}}$  relative to  $\text{Mg}^{2+}$  in variants 11ntR, 11ntR\_AC, and 11ntR\_A/CC by up to 3-fold, it increased  $k_{\text{undock}}$  for variants C7.10 and C7.2 by 2-fold, and had no effect on  $k_{\text{undock}}$  for variant C7.9. These small idiosyncratic effects on the magnitude and  $\text{M}^{2+}$  concentration dependence of the docking kinetics suggest that differences in conformation and electrostatic potential between TL/TLR variants lead to small but measurable discrimination between divalent cations that differ in size and solvation properties.

**Conformational Preferences and Electrostatic Properties of Connecting Tether Produce a Common, Separable Effect on TL/TLR Variants.** We measured the docking kinetics for four of the TL/TLR variants with the tetraloop and tetraloop-receptor components connected by both an  $\text{A}_7$  and  $\text{U}_7$  tether (Figure 7). Consistent with prior results,<sup>37,38,40</sup> the elements connecting the tetraloop and the tetraloop-receptor did not significantly alter  $k_{\text{undock}}$  or its  $\text{Mg}^{2+}$  dependence (Figures 7B, Undocking). Thus, the ratio  $k_{\text{undock}}^{\text{U}_7}/k_{\text{undock}}^{\text{A}_7}$  is essentially unity across the range of  $\text{Mg}^{2+}$  concentrations investigated (Figure 7D). In contrast, there were differences in  $k_{\text{dock}}$  between the  $\text{U}_7$  and  $\text{A}_7$  tethers (Figure 7A, Docking). Nevertheless, the effect from the  $\text{U}_7$  versus  $\text{A}_7$  tether was identical for all of the variants at each  $\text{Mg}^{2+}$  concentration, as demonstrated by the superimposable  $k_{\text{undock}}^{\text{U}_7}/k_{\text{undock}}^{\text{A}_7}$  ratios for all four TL/TLRs at each  $\text{Mg}^{2+}$  concentration (Figure 7C). Thus, the differences in  $k_{\text{dock}}$  and its  $\text{Mg}^{2+}$  dependence between  $\text{U}_7$  and  $\text{A}_7$  constructs are intrinsic to the conformational preferences and  $\text{Mg}^{2+}$  dependence of the

tether itself and separable from the energetics and distinct  $\text{Mg}^{2+}$  dependences of the TL/TLR variants.

The slower docking with the  $\text{A}_7$  relative to the  $\text{U}_7$  tether (Figure 7A) is consistent with the less frequent end-to-end collisions and the higher persistence length observed for poly-A oligonucleotides relative to poly-U.<sup>71,85,86</sup> Presumably some of the base-stacking interactions between A bases in the  $\text{A}_7$  tether break prior to the docking transition state, or their presence alters the conformational landscape to make collisions less probable. In contrast, weaker base-stacking interactions between U bases within the  $\text{U}_7$  tether may not contribute significantly to the conformational preferences of the tether and, hence, to the activation energy for the formation of the compact docking transition state.

The common tether effects across TL/TLR sequence variants at each  $\text{Mg}^{2+}$  concentration (Figure 7C,D), contrasted with the idiosyncratic  $\text{Mg}^{2+}$ -dependence of the TL/TLR variants with a given tether (Figure 4A–C), indicate that the contribution to folding from the conformational search of the tether is independent from the intrinsic energetics of the TL/TLR. These results are consistent with the model that RNA motifs behave as separable modules. This observation and previous studies<sup>25,31</sup> underscore the possibility of developing a reconstitution model for RNA folding from the isolated properties of its junctions, helices, and tertiary motifs (see Discussion).

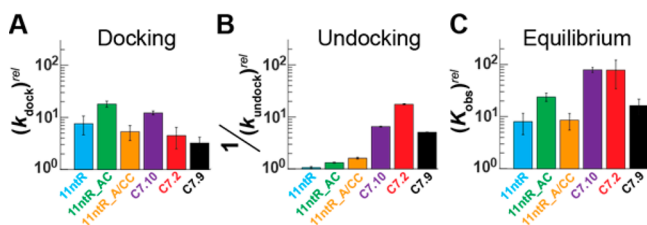
**Comparison of Docking Kinetics and Thermodynamics of the TL/TLR Variants in Monovalent Salts.** To further investigate the electrostatic energy landscape of the TL/TLR variants, we explored the docking kinetics and thermodynamics of the TL/TLR variants across a range of  $\text{K}^+$ ,  $\text{Na}^+$ , and  $\text{Rb}^+$  concentrations (Figure 8 and Figures S4 and S5). Previous studies provided evidence for a  $\text{K}^+$  binding site below the AA-platform of the canonical GAAA/11ntR tertiary motif that has a significant thermodynamic effect,<sup>87–89</sup> and other sequence-specific cation binding sites, whether inner or outer sphere, could form within the grooves and crevices of the TL/TLR variants and affect the kinetics and thermodynamics of TL/TLR formation.

The six TL/TLR variants exhibited qualitatively similar behavior in the presence of  $\text{K}^+$  as observed with  $\text{Mg}^{2+}$ , i.e., increasing  $k_{\text{dock}}$ , decreasing  $k_{\text{undock}}$ , and higher stability with increasing  $\text{K}^+$  concentration. As observed in  $\text{Mg}^{2+}$ , the 11ntR and 11ntR\_A/CC variants docked with greater  $k_{\text{dock}}$  values compared to other variants over the measured salt concen-



tration range, followed closely by variant C7.9 (Figure 8A). Interestingly, these three tetraloop-receptor variants have a pair of adjacent A residues in their asymmetric loops that can form AA-platforms, whereas the other variants have alternative sequences in those positions (Figure 2A); as discussed below, organization of the dinucleotide platform may influence the docking rate of TL/TLR motifs. The undocking behavior of the TL/TLR variants was also qualitatively similar in  $K^+$  and  $Mg^{2+}$ . For example, 11ntR displayed relatively low  $k_{\text{undock}}$  values ( $<10 \text{ s}^{-1}$ ) over the entire salt concentration range, whereas C7.9 displayed  $k_{\text{undock}}$  values that exceeded  $100 \text{ s}^{-1}$  over the entire concentration range (Figure 8B). Also as in  $Mg^{2+}$ , the natural TL/TLRs were more stable than the *in vitro*-selected variants in  $K^+$  (Figure 8C).

Despite these similarities, there are differences between TL/TLRs in  $K^+$  and  $Mg^{2+}$  that reveal a complex relationship between the sequence of the tetraloop-receptor and the relative effects of divalent and monovalent cations. For example, relative to  $K^+$  only, addition of 1 mM  $Mg^{2+}$  accelerates docking for all TL/TLR variants but it does so to slightly different extents for different TL/TLR variants (Figure 9A). Notably,



**Figure 9.** Relative effects of  $Mg^{2+}$  vs  $K^+$  on the docking kinetics and thermodynamics of the TL/TLR variants.  $k_{\text{dock}}$  (A),  $k_{\text{undock}}$  (B), and  $K_{\text{obs}}$  (C) in 140 mM  $K^+$  were compared with and without 1 mM  $Mg^{2+}$ .  $(k_{\text{dock}})^{\text{rel}}$  is defined as  $k_{\text{dock}}$  at 140 mM  $K^+$  and 1 mM  $Mg^{2+}$  divided by  $k_{\text{dock}}$  at 140 mM  $K^+$ ;  $(k_{\text{undock}})^{\text{rel}}$  and  $(K_{\text{obs}})^{\text{rel}}$  are similarly defined. Values for variants that were too weak to measure in 140 mM  $K^+$  ( $a_{\pm, KCl} = 103 \text{ mM}$ ) were extrapolated from the linear fits in Figure 8.

relative to  $K^+$ ,  $Mg^{2+}$  produces a stronger effect on the  $k_{\text{dock}}$  values of variants 11ntR\_AC (green) and C7.10 (purple) compared to other variants, and these two variants also had higher  $k_{\text{dock}}$  values in  $Mg^{2+}$  relative to  $Ba^{2+}$  (Figure 6). These observations are consistent with small but measurable specific interactions between  $Mg^{2+}$  and these two tetraloop-receptor variants prior to the formation of the docking transition state.

Similarly, there were sequence dependent differences in the relative effect of  $Mg^{2+}$  and  $K^+$  in  $k_{\text{undock}}$  (Figure 9B). For example, whereas  $k_{\text{undock}}$  was essentially the same in  $K^+$  and  $Mg^{2+}$  for variant 11ntR, *in vitro*-selected variants C7.10, C7.2, and C7.9 displayed much slower undocking in the presence of  $Mg^{2+}$ . Corresponding to these effects on  $k_{\text{dock}}$  and  $k_{\text{undock}}$ , there were  $\sim 10$ -fold differences in the relative effect of  $Mg^{2+}$  vs  $K^+$  on the stability of the TL/TLR variants (Figure 9C).

The differential effects of  $Mg^{2+}$  and  $K^+$  suggest differences in the electrostatic potential of the variants that may arise from local conformational differences between them. For example, conformations which place phosphate groups in close proximity may create high charge density “pockets” that favor  $Mg^{2+}$  uptake over  $K^+$  more than conformations without such electrostatic pockets.<sup>78,90</sup> The modest sequence-dependent differences in divalent (Figure 6) and monovalent (Figure S5) cation identity effects suggest that differences in the electrostatic potentials of the variants or in the positions of

functional groups can discriminate slightly for the size or the solvation properties of the cations.

## DISCUSSION

The growing database of RNA structures has revealed that complex RNAs are commonly built from a set of recurring structural motifs that largely preserve their three-dimensional structure irrespective of the context in which they are embedded.<sup>26–30,61,91</sup> From an energetic perspective, recent studies have shown that the canonical GAAA/11ntR tertiary motif undocks with the same energetic barrier, i.e., has the same rate constant, when embedded in RNAs of varying complexity.<sup>25</sup> These and additional observations suggest that kinetic and thermodynamic properties of isolated motifs may be applicable across structured RNAs containing those motifs, thereby providing a means to generalize and predict RNA folding behavior.<sup>24,25,31,92</sup> Below, we describe insights from our in-depth investigation of TL/TLR variants into the order of rearrangement in the folding pathway of TL/TLR docking, implications for the potential to quantitatively predict RNA folding kinetics and thermodynamics, and insights into the basis of Nature’s choice of TL/TLR variants.

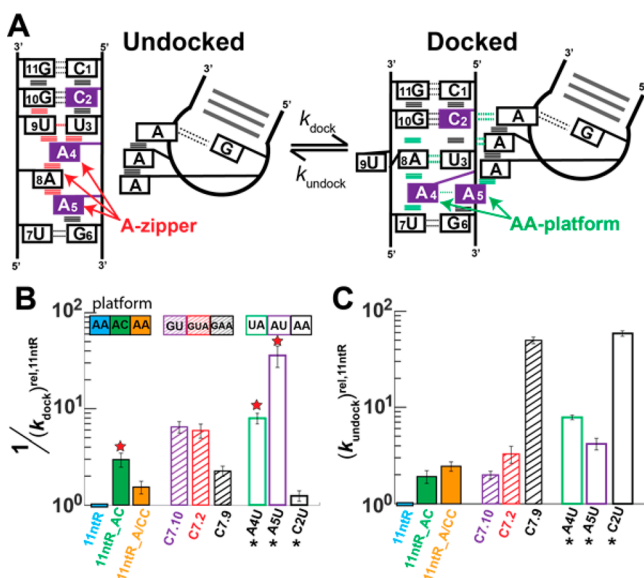
### Dissecting the Multistep Folding Pathway of TL/TLRs.

RNA folding studies to date have led to the suggestion that tertiary contact formation generally involves early transition states.<sup>93</sup> For example, ablation of individual tertiary interactions in the hairpin ribozyme increased  $k_{\text{undock}}$  without significant effects on  $k_{\text{dock}}$ , suggesting that these contacts form late in the folding process.<sup>80</sup> For GAAA/11ntR docking, a near-zero enthalpy change in going from the undocked state to the transition state provided support for the late formation of tertiary contacts in the overall folding of this RNA tertiary motif.<sup>40</sup> Tertiary folding of P4–P6 RNA, an independently folding component from the *Tetrahymena* group I intron, was also suggested to have an early transition state, although the reaction step or steps that were reported on by the attached fluorescent dye could not be ascertained.<sup>94</sup>

Notwithstanding the evidence for the late formation of tertiary contacts, RNA tertiary folding events are much slower than diffusional collision, strongly suggesting a need for substantial rearrangements preceding the formation of the transition state that would likely involve the breaking and formation of interactions.<sup>95</sup>

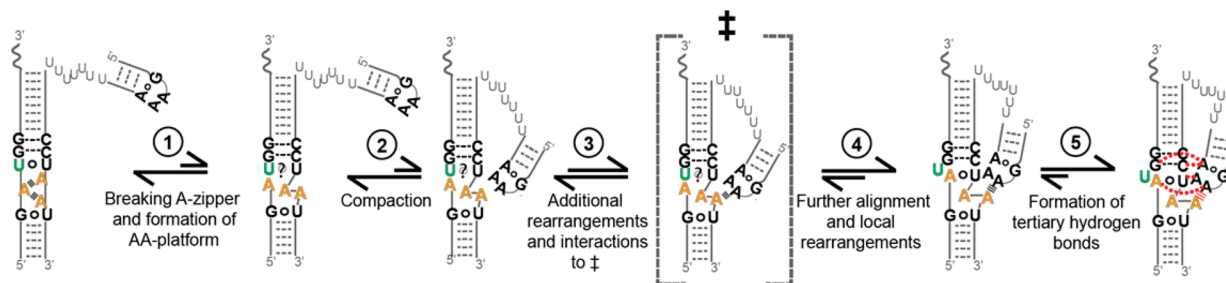
For GAAA/11ntR docking, the bimolecular association rate constant ( $k_{\text{assoc}}$ ) is 5 orders of magnitude slower than diffusional collision.<sup>95</sup> Indeed, NMR and X-ray crystal structures reveal that the undocked and docked conformations of the 11ntR tetraloop-receptor differ substantially, whereas the GAAA tetraloop appears to be unchanged (Figure 10A).<sup>43,96</sup> Thus, a reasonable model invokes contributions to the observed slow docking from some of these conformational rearrangements prior to the formation of the docking transition state. Other tetraloop-receptor variants docked with similar or lower  $k_{\text{dock}}$  values compared to 11ntR (Figures 4 and 8), suggesting similar unfavorable conformational rearrangements prior to the attainment of the docking transition state. To test this model and to ascertain which rearrangements might occur prior to the docking transition state, we compared the docking kinetics of these tetraloop-receptor variants along with three 11ntR single-point mutants studied previously (Figure 10B).<sup>31</sup>

If residues rearrange or make interactions prior to attainment of the docking transition state, then mutations of those residues



**Figure 10.** Conformational rearrangements of canonical 11ntR tetraloop-receptor and kinetic effect of sequence variations relative to 11ntR. (A) Schematic of the experimentally determined conformations of the undocked (left) and docked (right) canonical 11ntR tetraloop-receptor.<sup>24,47,96</sup> The GAAA tetraloop does not undergo significant conformational changes upon docking.<sup>24,30</sup> For the 11ntR tetraloop-receptor, in the undocked state A<sub>4</sub>, A<sub>8</sub>, and A<sub>5</sub> form a series of stacking interactions (A-zipper) that break upon formation of the AA-platform present in the docked state.<sup>96</sup> Base-stacking and hydrogen-bond interactions in the undocked state that need to break prior to attaining the docked state are colored in red, and interactions that form in the docked state are colored in green. Residues that were mutated to U in a previous study<sup>31</sup> are colored in purple.  $k_{\text{dock}}$  (B) and  $k_{\text{undock}}$  (C) of variants relative to the canonical 11ntR at 1 mM Mg<sup>2+</sup> and 140 mM K<sup>+</sup> (except for single point mutants marked with ‘\*’, which were measured previously<sup>31</sup> in 1 mM Ba<sup>2+</sup> and 140 mM K<sup>+</sup>; measurements across a range of [M<sup>2+</sup>] suggest that TL/TLRs display only minor kinetic differences in Mg<sup>2+</sup> vs Ba<sup>2+</sup> (Figure 6)). The sequences of the residues that comprise the putative platforms are shown in small boxes on the top of left panel. Variants that are identical to the 11ntR except for a single mutation to the AA-platform are marked with a red star.

are predicted to have an effect on  $k_{\text{dock}}$ . Conversely, mutations to residues that participate in rearrangements or make interactions only after the docking transition state are predicted to affect  $k_{\text{undock}}$  and to have no significant effect on  $k_{\text{dock}}$ .



**Figure 11.** Schematic model for TL/TLR folding for the GAAA/11ntR. The canonical 11ntR is shown starting with a schematic of the experimentally determined unbound form of the tetraloop-receptor and ending with the known structure of the docked GAAA/11ntR.<sup>47,96</sup> The features of this pathway are supported by mutational and ionic effects described in the text. Nevertheless, this model simplifies the many rearrangements that must occur in folding, and the experimental data distinguish only between effects that occur prior and subsequent to the rate-limiting transition state (‡). The U-U base pair in the unbound state is expected to be weak<sup>43</sup> and in the simplest model it breaks to allow rearrangement of the A residues to form the dinucleotide platform.

Despite the limitations in the interpretation of this type of mutational effects (i.e.,  $\Phi$ -value analysis as developed for protein folding) that have been discussed previously,<sup>97,98</sup> the results herein for TL/TLR docking suggest a single dominant pathway and provide evidence for events that occur early and late in this pathway. For example, had multiple disparate pathways been present, any single mutation would be expected to affect only one or a few of these pathways and give a small rate effect. Instead, our kinetic comparisons between tetraloop-receptor variants reveal substantial effects and provide evidence for substantial rearrangements of the residues that constitute the dinucleotide platform prior to attaining the docking transition state and for formation of tertiary hydrogen bonds after the formation of the docking transition state. A model summarizing these results is presented in Figure 11 and described below and in the next section.

Relative to the canonical 11ntR, single point mutations to residues that make up the AA-platform substantially reduce  $k_{\text{dock}}$  (Figure 10B, red stars). Tetraloop-receptor C7.10 and C7.2 which have GU and GUA platforms respectively, also docked more slowly than 11ntR (Figure 10B, hashed purple and red). In contrast, variants 11ntR\_A/CC, C7.9, and C2U that each have two adjacent A residues capable of forming AA-platforms exhibited  $k_{\text{dock}}$  values within 2–3-fold of the canonical 11ntR value (Figure 10B, filled orange, hashed black, and open black). These results suggest that rearrangements and interactions that involve the dinucleotide platform occur, at least in part, prior to the docking transition state. These rearrangements presumably involve breaking base-stacking and hydrogen bond interactions between residues of the tetraloop-receptor to expose the residues that form the dinucleotide platform and allow docking to occur (e.g., A-zipper in the undocked conformation of 11ntR, Figure 10A). The previous observation of near-zero net enthalpy change in the transition from the undocked to the transition state of the GAAA/11ntR<sup>40</sup> most simply suggests that the magnitude of the enthalpy changes of broken and newly formed interactions prior to the formation of the docking transition state are comparable. The proposed rearrangements of the dinucleotide platform of the 11ntR (i.e., AA-platform) prior to the docking transition state are illustrated in Step 1 of the model in Figure 11.

Mutations to the dinucleotide platform relative to the canonical 11ntR also increased  $k_{\text{undock}}$ , albeit modestly (2–8-fold), and these effects could arise from further reorganization

of or additional interactions with the dinucleotide platform after the attainment of the docking transition state (Figure 11, Step 4). Nevertheless, effects from mutations at other positions had larger effects on  $k_{\text{undock}}$ .

We consider these larger effects on  $k_{\text{undock}}$  in terms of the last step in the folding model of Figure 11 (Step 5). Mutation C2U is predicted to affect the network of hydrogen bonds (i.e., A-minor motif<sup>61</sup>) between the GAAA tetraloop and the C<sub>2</sub>-G<sub>10</sub> base pair in the docked state of the GAAA/11ntR (Figure 10A). Relative to the canonical 11ntR, this mutation had a large effect on  $k_{\text{undock}}$  but a negligible effect on  $k_{\text{dock}}$  suggesting that the hydrogen bond network to the wild-type C<sub>2</sub>-G<sub>10</sub> base pair forms after the transition state in the folding pathway (Figure 11, Step 5). Late formation of the tertiary hydrogen bonds between the tetraloop and the tetraloop-receptor is also consistent with the previous finding of a large favorable enthalpy change after the docking transition state.<sup>40</sup>

Additional information can be obtained from the insertion of a G residue between residues 3 and 4 relative to 11ntR, in variant C7.9, which is expected to alter the relative position of the U<sub>3</sub>·A<sub>8</sub> noncanonical base pair and the AA-platform (Figure 10A). As the U<sub>3</sub>·A<sub>8</sub> and the AA-platform both make tertiary interactions with the GAAA tetraloop, the ~50-fold increase in  $k_{\text{undock}}$  with only ~2.5-fold effect on  $k_{\text{dock}}$  from the insertion of G between U<sub>3</sub>·A<sub>8</sub> and the AA-platform suggest that the tertiary interactions at one or both of these sites form or rearrange subsequent to the docking transition state, consistent with the order of events in the folding model of Figure 11.

In summary, our kinetic analysis at near-physiological ionic conditions in combination with previous studies of the canonical GAAA/11ntR suggests a dominant pathway for the formation of the GAAA/11ntR in which reorganization of the dinucleotide platform occurs early in the folding pathway, before the formation of the docking transition state, followed by additional rearrangements and formation of tertiary hydrogen bonds after the docking transition state. For simplicity, we depict formation of the tertiary hydrogen bond network involving the conserved C<sub>2</sub>-G<sub>10</sub> as the final step in docking (Figure 11, Step 5). As described in the next section, differential ionic effects provide evidence for additional rearrangements prior to and subsequent to the formation of the docking transition state.

**Common and Specific Electrostatic Effects in the TL/TLR Folding Pathway.** In a seminal study, the rate constant for docking of the hairpin ribozyme was shown to depend strongly on the concentration of Mg<sup>2+</sup>, whereas the rate constant for undocking was largely independent of [Mg<sup>2+</sup>].<sup>80</sup> These observations, along with electrostatic modeling, led to a simple model in which Mg<sup>2+</sup> is taken up as the helices containing the tertiary contact partners come together to form a compact transition state, with the charge density of the RNA not changing significantly in subsequent folding steps. Previous studies of the isolated GAAA/11ntR showed similar Mg<sup>2+</sup> effects, i.e., a strong dependence of docking on Mg<sup>2+</sup> and little or no dependence of undocking on Mg<sup>2+</sup>, suggesting that this simple model might be broadly applicable in RNA folding.<sup>37,40</sup>

Our systematic comparative analysis of multiple TL/TLR variants under a range of ionic conditions has allowed us to test and extend this model. As described below, our results indicate a more complex electrostatic landscape in which Mg<sup>2+</sup> uptake occurs at multiple steps along the folding pathway; nevertheless, contributions from electrostatics and ion interactions can be assigned to separate energetic terms within a framework

that ultimately may allow quantitative prediction of RNA folding kinetics and thermodynamics.<sup>25,31</sup>

Figure 5C summarizes the average number of Mg<sup>2+</sup> ions taken up prior to ( $\Delta\Gamma_{\text{d},\text{Mg}^{2+}}^{\ddagger}$ ) and subsequent to ( $\Delta\Gamma_{\text{u},\text{Mg}^{2+}}^{\ddagger}$ ) the docking transition state, and the overall Mg<sup>2+</sup> uptake ( $\Delta\Gamma_{\text{Mg}^{2+}}^{\ddagger}$ ) for each of the TL/TLR variants studied herein.

For all variants, a large fraction of the total Mg<sup>2+</sup> uptake occurs prior to the formation of the docking transition state, with similar values of  $\Delta\Gamma_{\text{d},\text{Mg}^{2+}}^{\ddagger}$  ( $1.02 \pm 0.16$ ) between TL/TLR variants (Figure 5C, filled). This observation suggests a common large change in charge density as the tertiary contact partners come together to form a compact transition state (Figure 11, Step 2) and is consistent with the previous studies of the hairpin ribozyme and the canonical GAAA/11ntR mentioned above.<sup>37,40,80</sup> Nevertheless, we observe small but significant differences in  $\Delta\Gamma_{\text{d},\text{Mg}^{2+}}^{\ddagger}$  values between variants (Figure 5C, filled) and small sequence-dependent differences in the relative effects of distinct divalent (Figure 6) and monovalent (Figure S6) cations on  $k_{\text{dock}}$ . These observed sequence-dependent effects suggest that, in addition to compaction, local conformational rearrangements prior to the transition state alter the charge density of the TL/TLR and result in cation uptake (Figure 11, Steps 1 and 3).

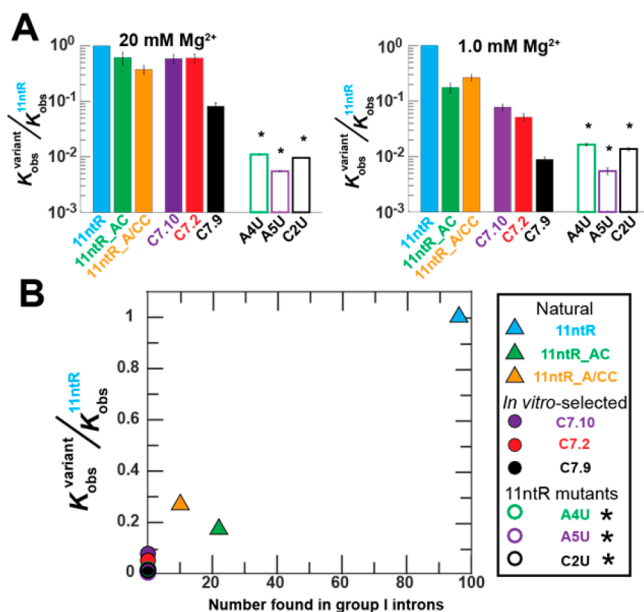
Idiosyncratic effects are even more evident in Mg<sup>2+</sup> uptake subsequent to the formation of the docking transition state ( $0.40 \pm 0.40$ , Figure 5C, open bars). In contrast to early results that suggested minimal charge density changes after the formation of the docking transition state, our analysis revealed the presence of such effects for most of the sequence variants. These effects are caused by the combined effect of local conformational rearrangements, alignment of the tertiary contact partners and ion binding that occurs after the formation of the docking transition state (Figure 11, Step 4).

Despite sequence dependent differences in ionic effects before and after the formation of the docking transition state, substituting the U<sub>7</sub> tether with A<sub>7</sub> gave a uniform effect on  $k_{\text{dock}}$  and no effect on  $k_{\text{undock}}$  (Figure 7), suggesting that it will be possible to consider the electrostatics of junctions and compaction separately from the intrinsic electrostatics and ion-binding of the tertiary contacts. In essence, this separability represents the core of the Reconstitution Hypothesis for RNA folding, which posits that the folding of a complex RNA can be understood from the isolated behavior of its constituent helices, junctions, and tertiary structural motifs.<sup>31</sup> The findings herein provide additional evidence in support of the Reconstitution Hypothesis, and indicate that additional sequence-dependent electrostatic effects will need to be incorporated in the Reconstitution model.

**Correspondence between Thermodynamic Stability and Biological Occurrence of Tetraloop-Receptors.** The *in vitro*-selected variants studied herein are not observed in natural structured RNAs, even though the selection experiments that yielded these variants also produced the naturally occurring 11ntR variants and these variants were shown to have similar thermodynamic stability.<sup>24,44,65</sup> As the *in vitro* selection and prior thermodynamic measurements were carried out at high Mg<sup>2+</sup> (15–100 mM) and at low K<sup>+</sup> (0–50 mM) relative to physiological salt conditions,<sup>44,45</sup> we compared the thermodynamic stability of the TL/TLR variants at salt conditions closer to physiological.

Our measurements with the minimal smFRET TL/TLR construct support the prior observations of similar thermodynamic stability of natural and *in vitro*-selected variants at high

Mg<sup>2+</sup> concentrations (Figure 12A, left); however, under salt concentrations that approximate physiological conditions (1



**Figure 12.** Thermodynamic stability of TL/TLR variants in isolated smFRET construct at near physiological ionic conditions correlates with observed frequency in group I introns. (A) Stability of TL/TLR variants in smFRET construct relative to the canonical 11ntR at high and near-physiological Mg<sup>2+</sup> concentrations (values with “\*” were measured previously<sup>31</sup> in Ba<sup>2+</sup>). Values at 20 mM Mg<sup>2+</sup> (left) for TL/TLR variants studied here were extrapolated from linear fits in Figure 4C. All measurements were carried out in a background of 140 mM K<sup>+</sup>. (B) Relative stability of TL/TLR variants in 1 mM Mg<sup>2+</sup> and 140 mM K<sup>+</sup> (except for variants with “\*” in figure legend which were measured at 1 mM Ba<sup>2+</sup> and 140 mM K<sup>+</sup> previously<sup>31</sup>) plotted against their observed frequency in the Group I Intron Sequence and Structure Database (GISSD).<sup>50</sup>

mM Mg<sup>2+</sup>; 140 mM K<sup>+</sup>), the naturally occurring canonical 11ntR was >10-fold more stable than the *in vitro*-selected variants (Figure 12A, right). Under these conditions, two natural variants that are less common in natural RNAs than the canonical 11ntR (11ntR\_AC and 11ntR\_A/CC) were also less stable than the 11ntR, though still more stable than the tetraloop-receptors not observed in Nature. Similarly, the three previously studied 11ntR single-point mutants that have not been observed in Nature are also highly destabilizing (Figure 12A, open bars). Overall, there appears to be a relationship between the thermodynamic stability of a tetraloop-receptor variant and its usage in natural structured RNAs, as shown in our analysis of group I introns (Figure 12B).

Though thermodynamic stability can account for the relative usage of the 11ntR-like TL/TLRs (Figure 12B), other types of TL/TLR motifs that are commonly found in natural RNA sequences, such as GNRA tetraloops docking into tandem Watson–Crick base pairs, are thermodynamically much weaker than the naturally occurring TL/TLRs described herein.<sup>24</sup> These distinct TL/TLRs may be used by Nature to favor different tertiary orientations, to facilitate conformational rearrangements, and in cases where there is sufficient stability from other structural elements.

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b08870.

Supporting Methods, Tables S1–S9, Figures S1–S275 (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*herschla@stanford.edu

### ORCID

Steve Bonilla: 0000-0002-6526-7158

### Present Address

#Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, United States

### Funding

S.B., N.B. and M.G. were supported by the NIH grant P01GM066275 to D.H.; S.B. was supported by the NSF GRFP under grant DGE-114747; C.L. was supported by NSF CMMI-0856205 to H.M.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Dawn Kellogg Merriman, Brant Gracia, Isabel Strohkendl, Sabine van Schie, Dmitri Pavlichin, and all members of the Herschlag lab for their comments on the paper. We thank Rhiju Das for sharing a PDB file of the computationally predicted structure of GAAA/C7.2. S.B. thanks Xuesong Shi and Ben Allred for helpful discussions and advice.

## ■ REFERENCES

- (1) Wan, R.; Yan, C.; Bai, R.; Huang, G.; Shi, Y. *Science* **2016**, 353 (6302), 895–904.
- (2) Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. *Science* **2000**, 289 (5481), 905–920.
- (3) Steitz, T. A. *Nat. Rev. Mol. Cell Biol.* **2008**, 9 (3), 242–253.
- (4) Stagno, J. R.; Liu, Y.; Bhandari, Y. R.; Conrad, C. E.; Panja, S.; Swain, M.; Fan, L.; Nelson, G.; Li, C.; Wendel, D. R.; White, T. A.; Coe, J. D.; Wiedorn, M. O.; Knoska, J.; Oberthuer, D.; Tuckey, R. A.; Yu, P.; Dyba, M.; Tarasov, S. G.; Weierstall, U.; Grant, T. D.; Schwieters, C. D.; Zhang, J.; Ferré-D’Amaré, A. R.; Fromme, P.; Draper, D. E.; Liang, M.; Hunter, M. S.; Boutet, S.; Tan, K.; Zuo, X.; Ji, X.; Barty, A.; Zatsepin, N. A.; Chapman, H. N.; Spence, J. C. H.; Woodson, S. A.; Wang, Y.-X. *Nature* **2016**, 541 (7636), 242–246.
- (5) Herschlag, D. *J. Biol. Chem.* **1995**, 270 (36), 20871–20874.
- (6) Li, R.; Zhu, H.; Luo, Y. *Int. J. Mol. Sci.* **2016**, 17 (5).
- (7) Joyce, G. F.; Orgel, L. E. *RNA world* **1999**, 49–77.
- (8) Breaker, R. R. *Cold Spring Harb. Perspect. Biol.* **2012**, 4 (2).
- (9) Wachowius, F.; Attwater, J.; Holliger, P. Q. *Rev. Biophys.* **2017**, 50, e4.
- (10) Gracia, B.; Xue, Y.; Bisaria, N.; Herschlag, D.; Al-Hashimi, H. M.; Russell, R. *J. Mol. Biol.* **2016**, 428 (20), 3972–3985.
- (11) Bailor, M. H.; Sun, X.; Al-Hashimi, H. M. *Science* **2010**, 327 (5962), 202–206.
- (12) Tinoco, I.; Bustamante, C. *J. Mol. Biol.* **1999**, 293 (2), 271–281.
- (13) Brion, P.; Westhof, E. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, 26, 113–137.
- (14) Schroeder, R.; Barta, A.; Semrad, K. *Nat. Rev. Mol. Cell Biol.* **2004**, 5 (11), 908–919.
- (15) Woodson, S. A. *Annu. Rev. Biophys.* **2010**, 39 (1), 61–77.
- (16) Costa, M.; Michel, F. *EMBO J.* **1995**, 14 (6), 1276–1285.
- (17) Massire, C.; Jaeger, L.; Westhof, E. *RNA* **1997**, 3 (6), 553–556.

- (18) Jaeger, L.; Verzemnieks, E. J.; Geary, C. *Nucleic Acids Res.* **2009**, *37* (1), 215–230.
- (19) Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Szwedczak, A.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273* (5282), 1696–1699.
- (20) Grabow, W. W.; Zhuang, Z.; Shea, J. E.; Jaeger, L. *Wiley Interdiscip. Rev. RNA* **2013**, *4* (2), 181–203.
- (21) Leontis, N. B.; Lescoute, A.; Westhof, E. *Curr. Opin. Struct. Biol.* **2006**, *16* (3), 279–287.
- (22) Hendrix, D. K.; Brenner, S. E.; Holbrook, S. R. Q. *Rev. Biophys.* **2005**, *38* (3), 221–243.
- (23) Butcher, S. E.; Pyle, A. M. *Acc. Chem. Res.* **2011**, *44* (12), 1302–1311.
- (24) Fiore, J. L.; Nesbitt, D. J. Q. *Rev. Biophys.* **2013**, *46* (3), 223–264.
- (25) Bisaria, N.; Greenfeld, M.; Limouse, C.; Pavlichin, D. S.; Mabuchi, H.; Herschlag, D. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (34), E4956–E4965.
- (26) Wu, L.; Chai, D.; Fraser, M. E.; Zimmerly, S. *PLoS One* **2012**, *7* (11), e49225.
- (27) Klein, D. J.; Schmeing, T. M.; Moore, P. B.; Steitz, T. A. *EMBO J.* **2001**, *20* (15), 4214–4221.
- (28) Nasalean, L.; Stombaugh, J.; Zirbel, C. L.; Leontis, N. B. *Non-Protein Coding RNAs* **2009**, *13*, 1–26.
- (29) Jaeger, L.; Verzemnieks, E. J.; Geary, C. *Nucleic Acids Res.* **2009**, *37* (1), 215–230.
- (30) Heus, H. A.; Pardi, A. *Science* **1991**, *253* (5016), 191–194.
- (31) Bisaria, N.; Greenfeld, M.; Limouse, C.; Mabuchi, H.; Herschlag, D. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (37), E7688–E7696.
- (32) Grabow, W.; Jaeger, L. *F1000Prime Rep.* **2013**, *5*, 46.
- (33) Goody, T. A.; Melcher, S. E.; Norman, D. G.; Lilley, D. M. J. *RNA* **2004**, *10* (2), 254–264.
- (34) Shi, X.; Huang, L.; Lilley, D. M. J.; Harbury, P. B.; Herschlag, D. *Nat. Chem. Biol.* **2016**, *12* (3), 146–152.
- (35) Hohng, S.; Wilson, T. J.; Tan, E.; Clegg, R. M.; Lilley, D. M. J.; Ha, T. J. *Mol. Biol.* **2004**, *336* (1), 69–79.
- (36) Melcher, S. E.; Wilson, T. J.; Lilley, D. M. J. *RNA* **2003**, *9* (7), 809–820.
- (37) Hodak, J. H.; Downey, C. D.; Fiore, J. L.; Pardi, A.; Nesbitt, D. J. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (30), 10505–10510.
- (38) Downey, C. D.; Fiore, J. L.; Stoddard, C. D.; Hodak, J. H.; Nesbitt, D. J.; Pardi, A. *Biochemistry* **2006**, *45* (11), 3664–3673.
- (39) Fiore, J. L.; Holmstrom, E. D.; Fiegand, L. R.; Hodak, J. H.; Nesbitt, D. J. *J. Mol. Biol.* **2012**, *423* (2), 198–216.
- (40) Fiore, J. L.; Holmstrom, E. D.; Nesbitt, D. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (8), 2902–2907.
- (41) Holmstrom, E. D.; Fiore, J. L.; Nesbitt, D. J. *Biochemistry* **2012**, *51* (18), 3732–3743.
- (42) Qin, P. Z.; Butcher, S. E.; Feigon, J.; Hubbell, W. L. *Biochemistry* **2001**, *40* (23), 6929–6936.
- (43) Davis, J. H.; Tonelli, M.; Scott, L. G.; Jaeger, L.; Williamson, J. R.; Butcher, S. E. *J. Mol. Biol.* **2005**, *351* (2), 371–382.
- (44) Costa, M.; Michel, F. *EMBO J.* **1997**, *16* (11), 3289–3302.
- (45) Geary, C.; Baudrey, S.; Jaeger, L. *Nucleic Acids Res.* **2008**, *36* (4), 1138–1152.
- (46) Ohuchi, S. P.; Ikawa, Y.; Nakamura, Y. *Nucleic Acids Res.* **2008**, *36* (11), 3600–3607.
- (47) Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273* (5282), 1678–1685.
- (48) Krasilnikov, A. S.; Yang, X.; Pan, T.; Mondragón, A. *Nature* **2003**, *421* (6924), 760–764.
- (49) Stahley, M. R.; Strobel, S. A. *Science* **2005**, *309* (5740), 1587–1590.
- (50) Zhou, Y.; Lu, C.; Wu, Q. J.; Wang, Y.; Sun, Z. T.; Deng, J. C.; Zhang, Y. *Nucleic Acids Res.* **2008**, *36* (SUPPL.1), D31–D37.
- (51) Greenfeld, M.; Herschlag, D. *Methods Enzymol.* **2013**, *530*, 281–297.
- (52) Greenfeld, M.; Solomatin, S. V.; Herschlag, D. *J. Biol. Chem.* **2011**, *286* (22), 19872–19879.
- (53) Kladwang, W.; Hum, J.; Das, R. *Sci. Rep.* **2012**, *2*, 517.
- (54) Leipply, D.; Lambert, D.; Draper, D. E. *Methods Enzymol.* **2009**, *469*, 433–463.
- (55) Robinson, R. A.; Stokes, A. H. *Electrolyte Solutions*, Second Rev.; Dover, 2002.
- (56) Greenfeld, M.; Pavlichin, D. S.; Mabuchi, H.; Herschlag, D. *PLoS One* **2012**, *7* (2), e30024.
- (57) Gebala, M.; Bonilla, S.; Bisaria, N.; Herschlag, D. *J. Am. Chem. Soc.* **2016**, *138* (34), 10925–10934.
- (58) Shi, X.; Bisaria, N.; Benz-Moy, T. L.; Bonilla, S.; Pavlichin, D. S.; Herschlag, D. *J. Am. Chem. Soc.* **2014**, *136* (18), 6643–6648.
- (59) Raw data for every TL/TLR variant and solution condition were deposited at <https://purl.stanford.edu/hj41Sjd0699>.
- (60) Ishikawa, J.; Fujita, Y.; Maeda, Y.; Furuta, H.; Ikawa, Y. *Methods* **2011**, *54* (2), 226–238.
- (61) Nissen, P.; Ippolito, J. A.; Ban, N.; Moore, P. B.; Steitz, T. A. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (9), 4899–4903.
- (62) Dai, L.; Toor, N.; Olson, R.; Keeping, A.; Zimmerly, S. *Nucleic Acids Res.* **2003**, *31* (1), 424–426.
- (63) Brown, J. W.; Haas, E. S.; Gilbert, D. G.; Pace, N. R. *Nucleic Acids Res.* **1994**, *22* (17), 3660–3662.
- (64) Sudarsan, N.; Lee, E. R.; Weinberg, Z.; Moy, R. H.; Kim, J. N.; Link, K. H.; Breaker, R. R. *Science* **2008**, *321* (5887), 411–413.
- (65) Afonin, K. A.; Lin, Y. P.; Calkins, E. R.; Jaeger, L. *Nucleic Acids Res.* **2012**, *40* (5), 2168–2180.
- (66) Sripakdeevong, P.; Kladwang, W.; Das, R. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (51), 20573–20578.
- (67) Blanco, M.; Walter, N. G. *Methods Enzymol.* **2010**, *472* (10), 153–178.
- (68) Solomatin, S. V.; Greenfeld, M.; Chu, S.; Herschlag, D. *Nature* **2010**, *463* (7281), 681–684.
- (69) Smith, K. D.; Lipchock, S. V.; Ames, T. D.; Wang, J.; Breaker, R. R.; Strobel, S. A. *Nat. Struct. Mol. Biol.* **2009**, *16* (12), 1218–1223.
- (70) Solomatin, S. V.; Greenfeld, M.; Herschlag, D. *Nat. Struct. Mol. Biol.* **2011**, *18* (6), 732–734.
- (71) Wang, X.; Nau, W. M. *J. Am. Chem. Soc.* **2004**, *126* (3), 808–813.
- (72) Bai, Y.; Greenfeld, M.; Travers, K. J.; Chu, V. B.; Lipfert, J.; Doniach, S.; Herschlag, D. *J. Am. Chem. Soc.* **2007**, *129* (48), 14981–14988.
- (73) Gebala, M.; Giambaşu, G. M.; Lipfert, J.; Bisaria, N.; Bonilla, S.; Li, G.; York, D. M.; Herschlag, D. *J. Am. Chem. Soc.* **2015**, *137* (46), 14705–14715.
- (74) Lipfert, J.; Doniach, S.; Das, R.; Herschlag, D. *Annu. Rev. Biochem.* **2014**, *83* (1), 813–841.
- (75) Chu, V. B.; Bai, Y.; Lipfert, J.; Herschlag, D.; Doniach, S. *Curr. Opin. Chem. Biol.* **2008**, *12* (6), 619–625.
- (76) Draper, D. E. *RNA* **2004**, *10* (3), 335–343.
- (77) Record, M. T.; Zhang, W.; Anderson, C. F. *Adv. Protein Chem.* **1998**, *51*, 281–353.
- (78) Misra, V. K.; Draper, D. E. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (22), 12456–12461.
- (79) Leipply, D.; Draper, D. E. *Biochemistry* **2010**, *49* (9), 1843–1853.
- (80) Bokinsky, G.; Rueda, D.; Misra, V. K.; Rhodes, M. M.; Gordus, A.; Babcock, H. P.; Walter, N. G.; Zhuang, X. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (16), 9302–9307.
- (81) Vander Meulen, K. A.; Butcher, S. E. *Nucleic Acids Res.* **2012**, *40* (5), 2140–2151.
- (82) Gluick, T. C.; Gerstner, R. B.; Draper, D. E. *J. Mol. Biol.* **1997**, *270* (3), 451–463.
- (83) Frederiksen, J. K.; Li, N.-S.; Das, R.; Herschlag, D.; Piccirilli, J. A. *RNA* **2012**, *18* (6), 1123–1141.
- (84) Travers, K. J.; Boyd, N.; Herschlag, D. *RNA* **2007**, *13* (8), 1205–1213.
- (85) Sim, A. Y. L.; Lipfert, J.; Herschlag, D.; Doniach, S. *Phys. Rev. E* **2012**, *86* (2), 21901.

- (86) McIntosh, D. B.; Duggan, G.; Gouil, Q.; Saleh, O. A. *Biophys. J.* **2014**, *106* (3), 659–666.
- (87) Basu, S.; Rambo, R. P.; Strauss-Soukup, J.; Cate, J. H.; Ferre-D'Amare, A. R.; Strobel, S. A.; Doudna, J. A. *Nat. Struct. Biol.* **1998**, *5* (11), 986–992.
- (88) Lambert, D.; Leipply, D.; Shiman, R.; Draper, D. E. *J. Mol. Biol.* **2009**, *390* (4), 791–804.
- (89) Bisaria, N.; Herschlag, D. *Biochem. Soc. Trans.* **2015**, *43* (2), 172–178.
- (90) Chin, K.; Sharp, K. A.; Honig, B.; Pyle, A. M. *Nat. Struct. Biol.* **1999**, *6* (11), 1055–1061.
- (91) Petrov, A. I.; Zirbel, C. L.; Leontis, N. B. *RNA* **2013**, *19* (10), 1327–1340.
- (92) Qin, H.; Sosnick, T. R.; Pan, T. *Biochemistry* **2001**, *40* (37), 11202–11210.
- (93) Mitra, S.; Brenowitz, M. *Nucleic Acid–Metal Ion Interact.* **2009**, 221–259; DOI: 10.1039/9781847558763-00221.
- (94) Silverman, S. K.; Cech, T. R. *RNA* **2001**, *7* (2), 161–166.
- (95) Herschlag, D.; Allred, B. E.; Gowrishankar, S. *Curr. Opin. Struct. Biol.* **2015**, *30*, 125–133.
- (96) Butcher, S. E.; Dieckmann, T.; Feigon, J. *EMBO J.* **1997**, *16* (24), 7490–7499.
- (97) Sánchez, I. E.; Kiefhaber, T. *J. Mol. Biol.* **2003**, *334* (5), 1077–1085.
- (98) Bartley, L. E.; Zhuang, X.; Das, R.; Chu, S.; Herschlag, D. *J. Mol. Biol.* **2003**, *328* (5), 1011–1026.