**ABSTRACT:** Bacterial riboswitches couple small-molecule ligand binding to RNA conformational changes that widely regulate gene expression, rendering them potential targets for antibiotic intervention. Despite structural insights, the ligand-mediated folding mechanisms of riboswitches are still poorly understood. Using single-molecule fluorescence resonance energy transfer (smFRET), we have investigated the folding mechanism of an H-type pseudoknotted preQ1 riboswitch in dependence of Mg²⁺ and three ligands of distinct affinities. We show that, in the absence of Mg²⁺, both weakly and strongly bound ligands promote pseudoknot docking through an induced-fit mechanism. By contrast, addition of as low as 10 μM Mg²⁺ generally shifts docking toward conformational selection by stabilizing a folded-like conformation prior to ligand binding. Supporting evidence from transition-state analysis further highlights the particular importance of stacking interactions during induced-fit and of specific hydrogen bonds during conformational selection. Our mechanistic dissection provides unprecedented insights into the intricate synergy between ligand- and Mg²⁺-mediated RNA folding.

1. INTRODUCTION

RNA is remarkable in its ability to recognize a diverse array of small molecules, as exemplified by the recently discovered class of noncoding RNAs known as riboswitches. Riboswitches are highly structured motifs commonly found in the 5′-untranslated regions of bacterial messenger RNAs (mRNAs), where they regulate gene expression mainly through mechanisms such as transcriptional termination and repression of translational initiation. Riboswitches are structurally modular in nature consisting of a ligand binding aptamer domain followed by an expression platform, the conformation of which dictates the level of gene expression. The aptamer domains of riboswitches adopt a variety of architectures and vary greatly in size and complexity. Ligand binding by the aptamer domain stabilizes alternative RNA conformations in the downstream expression platform to turn gene expression either off or on. Gene regulation by riboswitches therefore depends on the efficient folding of the aptamer as mediated by ligand binding. Despite the availability of a diversity of high-resolution structures, our understanding of the molecular mechanisms of the ligand-dependent dynamic folding of riboswitches is far from complete.

Molecular recognition processes in general are classified into two distinct mechanisms—induced-fit (IF) and conformational selection (CS)—which are commonly referred to as “binding first” and “folding first” processes, respectively. In IF, ligand binding to an unfolded conformation promotes folding, whereas in CS, the ligand selects high-affinity, folded-like structures from an ensemble and thus shifts the conformational equilibrium toward them. Distinguishing these mechanisms with ensemble-averaging techniques has been challenging so that single-molecule probing is rapidly gaining popularity for deciphering coupled binding and folding mechanisms. In particular, single-molecule fluorescence resonance energy transfer (smFRET) has been recently used to measure the impact of increasing ligand concentration on the closing and opening rate constants of maltose-binding and ABC transporter proteins as a way to distinguish IF from CS.

For riboswitches as a dynamic class of molecules, smFRET microscopy can provide valuable kinetic information about the folding and unfolding structural transitions that are crucial for ligand-dependent modulation of gene expression. Consequently, smFRET has been used to investigate the conformational states adopted by various riboswitches and their global dynamics. In most smFRET studies, the effect of the cognate ligand on the structure and dynamics of the riboswitch is investigated in either the presence or absence of Mg²⁺, which serves as the physiologically most relevant cofactor for tertiary structure formation of polyanionic RNA molecules. A common theme emerging from these studies is that Mg²⁺ alone can promote folded-like conformations that are further stabilized significantly upon ligand binding. However, in the absence of a

**Received:** April 10, 2015
**Revised:** September 16, 2015
**Published:** October 15, 2015
rigorous kinetic analysis of the conformational dynamics at varying ligand concentrations, the folding mechanisms of riboswitches have remained obscure and prone to interpretation based on assumptions.1 Recently, smFRET investigation of the effector-dependent conformational dynamics in proteins and DNAs showed that folding can be mechanistically dissected using kinetic analysis.9,10,12

Here, we have used smFRET kinetic analysis to investigate the coupled ligand binding and folding mechanism of the class-I preQ₁ riboswitch aptamer from B. subtilis (Bsu) (Figure 1a–e).13,14 The Bsu preQ₁-ribo switch is present in the 5′-untranslated region of genes coding for enzymes involved in the biosynthesis of a hypermodified nucleoside, queuosine.13 Queuosine occupies the anticodon wobble position of certain tRNAs, where it is critical for preventing errors during translation, and is also implicated in bacterial virulence and carcinogenesis in humans.15,16 The Bsu preQ₁ riboswitch senses the intracellular concentrations of the free queuosine intermediates preQ₀ (7-aminomethyl-7-deazaguanine) and (possibly) the closely related preQ₂ (7-cyano-7-deazaguanine) to terminate transcription of the genes via ligand-mediated stabilization of the aptamer domain.17 It features the smallest 34 nt length and adopts an H-sense the intracellular concentrations of the free queuosine structure make this aptamer a preferred model system to understand the molecular underpinnings of riboswitch folding.17 It features the smallest stabilization of the aptamer domain.17 It features the smallest structure even in the absence of ligand. On the basis of computational folding predictions suggesting that preQ₁ preferentially binds to the docked state, we proposed that this ligand under near-physiological Mg²⁺ conditions primarily binds through the CS (folding first) mechanism.

In the available crystal and NMR structures, the aptamer binds preQ₁ within its core where the ligand stacks between G5 and G11, forms a Watson–Crick base pair with C17, and interacts with U6 of L1 and A30 in L3, using almost all of its available hydrogen-bonding capacities (Figure 1d).17,18 In addition to its high-affinity (K_D ≈ 20 nM) cognate ligand preQ₁, the Bsu aptamer binds the related ligands preQ₀ and guanine less tightly with K_D values of ~100 and ~500 nM, respectively.13 The three ligands have distinct hydrogen-bonding capacities around ring position 7, where preQ₁ displays a methylamine group, while preQ₀ and guanine exhibit a nitrile group and only a ring nitrogen, respectively (Figure 1e). Reasoning that these distinctions may impact the balance between stacking and hydrogen-bonding interactions with the RNA, we compared by smFRET the docking (k_dock) and undocking (k_undock) rate constants to probe the coupling of aptamer folding with binding of these three ligands. Most strikingly, we found that ligand-induced folding shifts from the IF to the CS mechanism at Mg²⁺ concentrations as low as 10 μM.

2. RESULTS

2.1. Bsu Aptamer Folds through an Induced-Fit Mechanism in the Absence of Mg²⁺. The aptamer used for smFRET contained a 5′-biotin, 3′-Dy547, and an internal Cy5 on U19 (Figure 1a,e,c). This labeling scheme leads to a high-FRET (~0.9) state in the pseudoknotted structure and a mid-FRET (~0.65) state in the ligand-free prefolded conformation.21 Transitions between these two FRET states were observed as before,21 reporting the rate constants k_dock and k_undock under a broad range of buffer conditions (Tables S1 and S2). The solution structure of the Bsu preQ₁-aptamer was solved in the absence of Mg²⁺, showing that Mg²⁺ is not required for ligand binding by the aptamer.17 Therefore, we first studied the dynamics of the aptamer in the absence of Mg²⁺ (in a buffer containing 50 mM Tris-HCl, pH 7.5, 100 mM KCl) to probe the binding mechanism and decouple the divalent metal ion’s stabilizing effect on the folded conformations from the impact of ligand alone. Individual smFRET traces showed multiple transitions between the two FRET states at varying concentrations of each of the three ligands (Figure 2 and Figure S1). From these smFRET traces, dwell times in each state were extracted and fit well with single-exponential functions to extract k_dock and k_undock (see Materials and Methods and Figure S3). For preQ₁, the k_dock value increased from 0.47 s⁻¹ at 50 nM to reach a maximum of ~3.1 s⁻¹ at 10 μM with a hyperbolic dependence on the ligand concentration. A similar trend was observed for the weaker ligands preQ₀ and guanine; for preQ₀, the k_dock value increased from 0.43 s⁻¹ at 250 nM to 2.6 s⁻¹ at 10 μM, whereas for guanine, k_dock increased from 0.36 s⁻¹ at 250 nM to 1.41 s⁻¹ at 10 μM. For all three ligands, k_undock remained almost constant with only a marginal decrease at the highest concentrations for preQ₁; the average k_undock for preQ₁, preQ₀, and guanine was measured at 0.5, 1.1, and 2.4 s⁻¹, respectively. An accelerating k_dock and invariant k_undock with increasing ligand concentration suggest that, in the absence of Mg²⁺, the Bsu aptamer binds all three ligands via the IF mechanism.5,6,11

2.2. Mg²⁺ Promoted Aptamer Folding through a Conformational Selection Mechanism. Next, we probed the impact of a near-physiological (1 mM) concentration of
Mg$^{2+}$, an essential cofactor of tertiary contact formation during RNA folding, on the ligand-dependent folding of the preQ1 aptamer. Mg$^{2+}$ promotes compact, folded-like aptamer conformations in the absence of ligand (Figure S5). In the presence of 1 mM Mg$^{2+}$, the cognate ligand preQ1 increased $k_{\text{dock}}$ from ~1.1 s$^{-1}$ at 10 nM to 3.1 s$^{-1}$ at 1 μM (Figure 3 and Figure S2). Similarly, $k_{\text{dock}}$ increased from 0.60 s$^{-1}$ at 25 nM to 2.9 s$^{-1}$ at 10 μM preQ0 and from 0.47 s$^{-1}$ at 50 nM to 1.6 s$^{-1}$ at 10 μM guanine. Remarkably, in stark contrast to the no-Mg$^{2+}$ data, the $k_{\text{undock}}$ values decreased with increasing concentration for each of the three ligands. Specifically, for the cognate ligand preQ1, $k_{\text{undock}}$ decreased from ~1.2 s$^{-1}$ at 10 nM to 0.41 s$^{-1}$ at saturating 1 μM preQ1. For preQ0 and guanine, the $k_{\text{undock}}$ values decreased even more significantly from ~1.6 to 0.42 s$^{-1}$ and from ~2.0 to 0.47 s$^{-1}$, respectively (Figure 3). (We note that heterogenic kinetics were observed both for $k_{\text{dock}}$ and $k_{\text{undock}}$ in the presence of Mg$^{2+}$, and to some extent for $k_{\text{dock}}$ in the absence of Mg$^{2+}$, where data were slightly better fit with double than single exponentials (Figures S3 and S4). Importantly, the average rate constants $k_{\text{dock}}$ and $k_{\text{undock}}$ from double-exponential fits, while slightly smaller, showed the same differences in folding mechanism as the values obtained from single-exponential fits, which were used throughout for consistency among data sets). This kinetic signature of a ligand-dependent decrease in $k_{\text{undock}}$ is a characteristic of $C_{G}$ and shows that Mg$^{2+}$ shifts the aptamer-folding mechanism for all three ligands toward CS. However, the increase in $k_{\text{dock}}$ suggests that IF may also be taking place at higher ligand concentrations, as shown recently, providing initial evidence that, in the presence of Mg$^{2+}$, both the CS and IF folding mechanisms are operating among the population of molecules.

2.3. Single-Molecule Transition State Analysis of the Ligand-Mediated Riboswitch Folding. To gain deeper insights into the differences in the aptamer-folding pathways for the three ligands, we performed single-molecule transition-state analysis (smTSA) of the aptamer folding. TSA (or Φ-value analysis) is a powerful method that has been extensively used to study the nature of protein-folding transition states. By measuring the effect of conservative point mutations on the kinetics and thermodynamics of folding, structural information about the transition state can be probed. For a folding (docking, in this case) reaction, a Φ-value refers to the ratio of the change in the transition-state free energy ($ΔG_{\text{ts}}^{*}$) and the change in the equilibrium free energy ($ΔG_{\text{eq}}^{*}$) of folding between the wild-type and the mutant ($Φ = ΔG_{\text{ts}}^{*}/ΔG_{\text{eq}}^{*}$). From smFRET experiments, the activation and the equilibrium free energies can be obtained from the individual docking and undocking rate constants as $ΔG_{\text{dock}}^{*} = RT \ln \left( k_{\text{dock}}/k_{\text{mutant}} \right)$ and $ΔG_{\text{undock}}^{*} = RT \ln \left( k_{\text{undock}}/k_{\text{mutant}} \right)$, where $k_{\text{dock}}$ and $k_{\text{mutant}}$ are the equilibrium constants for docking in the wild-type and mutant, respectively. A mutation that only accelerates the undocking (unfolding) kinetics without affecting the docking kinetics results in $Φ = 0$, implying that the contacts made by a residue are not yet formed in the transition state. In contrast, a mutation that decreases docking without affecting undocking leads to $Φ > 1$, signifying that the tertiary interactions of the residue are fully established.
in the transition state. Therefore, the \( \Phi \) value provides a measure of the relative formation of the native tertiary interactions in the transition state at the site of mutation. By making multiple mutations at different sites in a molecule, TSA can also be used to model the structure of the folding transition state.\(^{20}\) For RNA folding, TSA has also been applied, although sparsely, to probe the effect of mutations, metal ions, and denaturants, using both ensemble and single-molecule FRET on a few noncoding RNA structures, but has not yet been applied to riboswitch folding.\(^{6,20}\)

Although TSA has been extensively used to study the folding of single domain proteins, it can also be applied to investigate biomolecular reactions where folding is coupled to ligand binding.\(^{30}\) For example, TSA recently has been applied to probe the coupled binding and folding mechanisms of intrinsically disordered proteins (IDPs), protein oligomerization, and aggregation.\(^{51-53}\) In these studies, mutations were introduced into one of the binding partners and their effects on the nature of the folding transition state were probed. This approach gives information on the structure of the initial “encounter complex”, thereby providing insights into the mechanism of folding.\(^{30}\) We here use a similar approach, albeit the external ligand in our case is a small molecule to which we introduce subtle conservative “mutations”.

In our smTSA, we calculated \( \Delta \Delta G_{\text{dock}} \) values that characterize the thermodynamic difference of docking associated with binding the alternate, weaker ligands relative to the cognate (“wild-type”) preQ\(_2\), as well as \( \Delta \Delta G_{\text{dock}} \) values that reflect the transition-state destabilization by the ligand mutations, all at their respective saturating concentrations. The \( \Phi \) values then quantify the fraction of ligand–RNA interactions lost through mutation that were already present in the transition state of preQ\(_2\)-mediated docking. By measuring the free energies at saturating ligand concentration, this analysis effectively treats the ligand as the integral part of the docked RNA structure that it is. In addition, the use of rate constants at saturating concentrations of all three ligands compensates for the differences in their binding affinities to the aptamer.\(^{13}\) Because the ligand-binding pocket is fully formed only upon docking of the 3’-tail to the P1–L1 stem-loop to adopt the compact pseudoknot, this analysis effectively quantifies the contacts made by the ligand in the transition state of aptamer folding.

In the absence of Mg\(^{2+}\), the values of \( k_{\text{dock}}^\infty \) and \( k_{\text{undock}}^\infty \) which are the rate constants for docking and undocking for the aptamer in the presence of saturating concentrations of preQ\(_2\), are 3.46 and 0.50 s\(^{-1}\), respectively. This gives an equilibrium constant for docking, \( K_{\text{dock}} = k_{\text{dock}}^\infty / k_{\text{undock}}^\infty \), of \( \sim 6.92 \) and a free energy of docking, \( \Delta G_{\text{dock}} = -RT \ln (K_{\text{dock}}) \), of approximately \( \sim 1.14 \) kcal/mol (Figure 4a and Table 1). On “mutating” the ligand to preQ\(_1\), \( k_{\text{dock}}^\infty \) slightly decreases to \( \sim 3.28 \) s\(^{-1}\) while the \( k_{\text{undock}}^\infty \) value increases almost 2-fold to \( \sim 1.06 \) s\(^{-1}\),, resulting in a lower \( K_{\text{dock}} \) value of \( \sim 3.09 \). The observation of a similar \( k_{\text{dock}}^\infty \) value for preQ\(_1\) as compared to preQ\(_2\) and a faster \( k_{\text{undock}}^\infty \) then leads to \( \Phi \approx 0.07 \pm 0.02 \) (Figure 4a; for a discussion of the accuracy of the smTSA analysis presented here, please see Supporting Information), suggesting that the docking transition states associated with binding the two ligands are nearly identical. This observation implies that the specific ligand–RNA contacts made by the exocyclic 7-methylamine group of preQ\(_1\) are not yet formed in the transition state. From the crystal structure, these contacts can be proposed as the one hydrogen bond that 7-methylamine forms with GS and the pro-Rp oxygen of G11 in P2 (Figure 1d). This finding also supports the notion that the aptamer folds via IF only after either of the two ligands has already bound to the undocked conformation (primarily via stacking interactions). Similarly, upon binding of guanine, the transition state is only modestly destabilized with a 2-fold smaller \( k_{\text{dock}}^\infty \) of \( \sim 1.72 \) s\(^{-1}\) as compared to the significant destabilization of the docked state, with a 5-fold faster \( k_{\text{undock}}^\infty \) value of \( \sim 2.38 \) s\(^{-1}\), resulting in \( \Phi \approx 0.31 \pm 0.15 \) (Figure 4a and Table 1). The fact that this value is nonzero may be due to the guanine-mediated transition state lacking a distinct stabilizing feature of both preQ\(_1\) and preQ\(_2\) such as the van der Waals (stacking) interaction provided by a substituent on ring position 7; alternatively, binding of guanine in the absence of Mg\(^{2+}\) may start shifting the folding mechanism toward CS.

Next, we utilized smTSA to probe the folding transition state in the presence of 1 mM Mg\(^{2+}\). Mg\(^{2+}\) significantly increases the equilibrium constant for docking, \( K_{\text{dock}} \), for preQ\(_1\) and guanine by more than 2-fold and 6-fold, respectively, while it only modestly increases preQ\(_2\)-dependent docking (Table 1). For preQ\(_2\) and guanine, this increase in \( K_{\text{dock}} \) is brought about largely by a significant decrease in the undocking rate constant, \( k_{\text{undock}}^\infty \). smTSA then indicates that the transition-state energies for all ligands shift similarly to those of their respective docked states (Figure 4b, leading to \( \Phi \) values of \( \sim 0.89 \pm 0.25 \) and \( \sim 0.78 \pm 0.17 \) for preQ\(_1\) and guanine, respectively. That is, a majority of ligand–RNA interactions are already established in the transition state, rendering it folded-like (undocked) and thus primed to bind ligand via CS. Notably, the fractional (nonunity) \( \Phi \) values provide further evidence consistent with a partitioning of the conformational ensemble into parallel folding pathways, as expected for a mixed CS/IF mechanism in the presence of Mg\(^{2+}\) in which a majority, but not all, RNA molecules have largely formed their liganding contacts in the transition state.\(^{6,27}\)

![Figure 4](https://example.com/figure4.png)

**Figure 4.** Single-molecule transition-state analysis of ligand-dependent riboswitch folding. Free energy diagrams for the ligand-dependent folding of the Bau aptamer in (a) the absence of Mg\(^{2+}\) and (b) the presence of Mg\(^{2+}\), drawn to scale. \( \Delta \Delta G_{\text{dock}} \) and \( \Delta \Delta G_{\text{undock}} \) for guanine-mediated docking as compared to preQ\(_1\) are shown.

2.4. Strong Influence of Mg\(^{2+}\) on the preQ\(_1\)-Mediated Folding Mechanism. Our kinetic analysis of ligand-depend-
ent conformational dynamics showed that, under a near-
physiological Mg$^{2+}$ concentration of 1 mM, the folding
mechanism shows features of CS as evident from the decreasing
$k_{\text{undock}}$ at higher ligand concentrations (Figure 3c), in contrast to
the IF mechanism at zero Mg$^{2+}$. To find the Mg$^{2+}$
concentration where the shift from IF to CS occurs, we
determined the rate constants of docking and undocking for the
cognate ligand preQ$_1$ at the three intermediate Mg$^{2+}$
concentrations of 1, 10, and 100 μM (Figure 5). Under all

columns, the IF mechanism at zero Mg$^{2+}$. To
k
CS pathway across the riboswitch population at Mg$^{2+}$
(left of the axis break, colored in red) is showing the zero [Mg$^{2+}$]
and a cooperativity constant $k_{1/2}$ for [Mg$^{2+}$] of 9 ± 1 μM Mg$^{2+}$
and yielded a $K_{1/2}$ of 9 ± 1 μM Mg$^{2+}$, with a cooperativity constant of $n = 1.47$. These findings suggest that at least one high-affinity Mg$^{2+}$ ion is required for 3′-tail docking to form the folded state ensemble that the ligand selectively binds during the CS mechanism.

3. DISCUSSION

Folding of riboswitches is intricately linked to ligand bind-
ing.$^{5,14}$ Such coupled binding and folding mechanisms are
crucial for effective gene regulation, especially for riboswitches
functioning at the level of transcription termination that entails
kinetic competition between alternate folding and tran-
scription.$^{34}$ Although the folding of proteins coupled to ligand
binding is actively investigated, similar studies on RNA are
relatively few. Over the past decade, our understanding of the
molecular mechanisms employed by RNA to recognize a
chemically diverse array of ligands was advanced by a wealth of
high-resolution crystal structures of various riboswitch classes
bound to their ligands.$^{1}$ In contrast, only a few structures of
ligand-free riboswitches were successfully solved, offering
limited insights into the plausible ligand-binding mechanisms
of riboswitches.$^{3,35}$ These structures likely represent only a
subset of compact, crystallizable conformations of a large
ligand-free ensemble sampled by the RNA, as suggested by
lower-resolution solution measurements (such as small-angle X-
ray scattering (SAXS)).$^{35}$ In addition, static crystal structures
cannot explain the mode of transition from a ligand-free (apo)
conformation to the ligand-bound (holo) conformation that is
crucial to understanding riboswitch mechanism and function.$^{4}$

NMR spectroscopy has been used to probe the conformations
sampled by riboswitches and to study ligand-binding mechanisms.$^{5,36,37}$ Recently, smFRET microscopy has emerged as a powerful tool to characterize the structure,
dynamics, and ligand-induced folding of riboswitches.$^{5,11,18}$ In
the present study, we have performed a smFRET-based kinetic
analysis of the preQ$_1$-1 aptamer from B. subtilis to probe the
influence of Mg$^{2+}$ on the folding mechanism in the presence of
ligands of varying affinities.

Our results demonstrate that physiological Mg$^{2+}$ concentra-
tions govern the folding pathway of the Bsu aptamer toward
its native ligand-bound state. On the basis of our smFRET
kinetic and transition-state analyses, we propose a model in
which ligand and Mg$^{2+}$ synergistically fold the Bsu aptamer
(Figure 6). In the absence of Mg$^{2+}$ and ligand, the aptamer
largely exists in an unfolded hairpin conformation, as previously
suggested from smFRET analysis of the Bsu aptamer and

Table 1. Rate Constants and Free Energies for smTSA

<table>
<thead>
<tr>
<th>ligand</th>
<th>$k_{\text{dock}}$ ($s^{-1}$)</th>
<th>$k_{\text{undock}}$ ($s^{-1}$)</th>
<th>$k_{\text{dock}}$</th>
<th>$\Delta G_{\text{dock}}$ (kcal/mol)</th>
<th>$\Delta G_{\text{undock}}$ (kcal/mol)</th>
<th>$\Delta G_{\text{undock}}$ (kcal/mol)</th>
<th>$\phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>preQ$_1$</td>
<td>3.46</td>
<td>0.50</td>
<td>6.92</td>
<td>0.57</td>
<td>0.47</td>
<td>0.20</td>
<td>0.07 ± 0.02</td>
</tr>
<tr>
<td>guanine</td>
<td>1.72</td>
<td>2.38</td>
<td>0.72</td>
<td>0.19</td>
<td>1.53</td>
<td>0.41</td>
<td>0.15</td>
</tr>
<tr>
<td>preQ$_0$</td>
<td>3.44</td>
<td>0.41</td>
<td>8.39</td>
<td>1.25</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>guanine</td>
<td>2.83</td>
<td>0.42</td>
<td>6.73</td>
<td>1.12</td>
<td>0.13</td>
<td>0.12</td>
<td>0.25</td>
</tr>
<tr>
<td>preQ$_1$</td>
<td>2.10</td>
<td>0.47</td>
<td>4.47</td>
<td>0.88</td>
<td>0.37</td>
<td>0.29</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Figure 5. Effect of Mg$^{2+}$ on (a) $k_{\text{dock}}$ and (b) $k_{\text{undock}}$ as a function of
preQ$_1$ concentration. The curves represent noncooperative ($n = 1$)
Hill-equation fits to the data at varying Mg$^{2+}$ concentrations as
indicated. (c) Semilogarithmic plot of the ratio $k_{\text{dock}}/k_{\text{undock}}$ against
the Mg$^{2+}$ concentration, fit with a Hill equation with a
fixed y-intercept of 1, yielding a saturation ratio of $\sim 2.9$, a $K_{1/2}$ for [Mg$^{2+}$]
of $9 \pm 1$ μM, and a cooperativity coefficient of $n = 1.47$. The first point in the plot
(left of the axis break, colored in red) is showing the zero [Mg$^{2+}$] value.

for the half-titration point $K_{1/2}$, we plotted the ratio $k_{\text{undock}}/k_{\text{dock}}$ at the different Mg$^{2+}$ concentrations as shown in
Figure 5c. The data, when fit with a Hill equation with $y$-intercept of 1
(since $k_{\text{undock}}$ is independent of the preQ$_1$ concentration at zero
Mg$^{2+}$), showed saturating behavior even at 100 μM Mg$^{2+}$ and
yielded a $K_{1/2}$ of $9 \pm 1$ μM Mg$^{2+}$, with a cooperativity constant of $n = 1.47$. These findings suggest that at least one high-affinity Mg$^{2+}$ ion is required for 3′-tail docking to form the folded state ensemble that the ligand selectively binds during the CS mechanism.

3. DISCUSSION

Folding of riboswitches is intricately linked to ligand bind-
ing.$^{5,14}$ Such coupled binding and folding mechanisms are
crucial for effective gene regulation, especially for riboswitches
functioning at the level of transcription termination that entails
kinetic competition between alternate folding and tran-
scription.$^{34}$ Although the folding of proteins coupled to ligand
binding is actively investigated, similar studies on RNA are
relatively few. Over the past decade, our understanding of the
molecular mechanisms employed by RNA to recognize a
chemically diverse array of ligands was advanced by a wealth of
high-resolution crystal structures of various riboswitch classes
bound to their ligands.$^{1}$ In contrast, only a few structures of
ligand-free riboswitches were successfully solved, offering
limited insights into the plausible ligand-binding mechanisms
of riboswitches.$^{3,35}$ These structures likely represent only a
subset of compact, crystallizable conformations of a large
ligand-free ensemble sampled by the RNA, as suggested by
lower-resolution solution measurements (such as small-angle X-
ray scattering (SAXS)).$^{35}$ In addition, static crystal structures
cannot explain the mode of transition from a ligand-free (apo)
conformation to the ligand-bound (holo) conformation that is
crucial to understanding riboswitch mechanism and function.$^{4}$

NMR spectroscopy has been used to probe the conformations
sampled by riboswitches and to study ligand-binding mechanisms.$^{5,36,37}$ Recently, smFRET microscopy has emerged as a powerful tool to characterize the structure,
dynamics, and ligand-induced folding of riboswitches.$^{5,11,18}$ In
the present study, we have performed a smFRET-based kinetic
analysis of the preQ$_1$-1 aptamer from B. subtilis to probe the
influence of Mg$^{2+}$ on the folding mechanism in the presence of
ligands of varying affinities.

Our results demonstrate that physiological Mg$^{2+}$ concentra-
tions govern the folding pathway of the Bsu aptamer toward
its native ligand-bound state. On the basis of our smFRET
kinetic and transition-state analyses, we propose a model in
which ligand and Mg$^{2+}$ synergistically fold the Bsu aptamer
(Figure 6). In the absence of Mg$^{2+}$ and ligand, the aptamer
largely exists in an unfolded hairpin conformation, as previously
suggested from smFRET analysis of the Bsu aptamer and
shown by NMR for a similar preQ1-I aptamer from F. nucleatum (Fnu). Under these conditions, the aptamer forms a transient conformational ensemble (the gray lines and the double-headed arrows are drawn to convey the dynamic nature of the ensemble). Ligand (yellow) binding to this unfolded conformation promotes folding via the IF pathway. In the presence of Mg2+, the aptamer samples compact, loosely pseudoknotted, folded-like conformations that are captured via CS and further stabilized by the ligand. High ligand concentrations favor IF, whereas CS dominates at high Mg2+ concentrations due to the increased population of compact conformations.

Mg2+ in general is known to promote fluctuations in riboswitches in the absence of ligand into compact native-like conformations that it then stabilizes; once sufficiently stabilized, these conformations can form the basis for CS by the ligand. Indirect evidence exists that even small differences in aptamer sequence and Mg2+ concentration may have a significant effect on the nature of the ligand-mediated folding mechanism. For example, while kinetic data suggest an IF mechanism for the pbuE adenine riboswitch in 0.5 mM Mg2+, the structurally similar spt guanine riboswitch was proposed to follow a CS model in 1 mM Mg2+. Similarly, we suggested that the structurally highly similar Bsu and Tfe preQ1-I riboswitches to fold through distinct CS and IF mechanisms in 1 mM Mg2+, respectively. Our finding of an exquisite sensitivity of the folding mechanism of the Bsu riboswitch to the Mg2+ concentration, with a transition from IF to CS folding around just 9 μM (Figure 5c), is arguably a particularly poignant example of such fine-tuning.

The observation of a Mg2+-promoted CS mechanism for the folding of the Bsu aptamer also raises the question of how a ligand reaches its binding pocket for which crystal and NMR structures reveal a largely buried solvent-accessible surface. Studies of enzymes containing a "lid" over their substrate binding sites have shown that the presence of even a small population of partially closed, ligand-free conformations are sufficient for the folding to proceed through either CS or IF mechanisms. In analogy, NMR studies of the Bsu aptamer have indicated that loop L2 is highly dynamic in the pseudoknot conformation, which is also supported by its missing density in the crystal structure, and may thus act as such a lid on the binding pocket in a loose pseudoknot with an almost completely formed P2.

From the limited mechanistic smFRET studies of riboswitches, no consensus has yet emerged between the mode of gene regulation and the mechanism of ligand binding. Transcriptionally acting riboswitches are thought to be kinetically controlled, where the ligand-mediated conformational change has to occur cotranscriptionally within a narrow time window before RNA polymerase passes the intrinsic terminator hairpin. For such a riboswitch under kinetic control, an IF mechanism may be advantageous since the rate of folding upon ligand binding may be significantly faster than the rate of transient fluctuations of ligand-free conformations during CS. In addition, for transcriptionally acting riboswitches the ability during IF to weakly bind the ligand even before the entire aptamer domain has been completely transcribed may provide extra time to commit to a gene regulatory decision. In the case of the transcriptional Bsu preQ1-I riboswitch studied here, the distance between the 3′-ends of the aptamer and the terminator hairpin is ~25 nt, leaving a time window of only ~1 s within which the ligand-bound pseudoknot has to stably fold to terminate transcription. Our observation of similarly fast pseudoknot docking/undocking kinetics under physiological Mg2+ concentrations (Figure 3), which may be further accelerated under the crowded molecular conditions found in vivo, suggests that the IF mechanism can in fact be suitable for such fast transcriptional regulation.

While the cognate ligand preQ1 with the highest affinity showed a relatively small decrease in kundock with increasing ligand concentration, the weaker near-cognate ligands preQ2 and in particular guanine showed more significant decreases (Figure 3). This is consistent with the notion that, under conditions where both open and closed conformations of the
RNA coexist, the weaker the ligand, the more it binds via CS, where compact, partially closed, high-affinity conformations are selected from the ensemble over low-affinity open conformations from which the weak ligand presumably dissociates too rapidly for IF to occur.41

Fractional Φ-values as obtained in our smTSA (Figure 4) are classically interpreted to correlate with the degree of native contacts formed in a transition state.27 More precisely, such a fractional Φ-value represents an average value for the transition-state ensemble (TSE) derived from parallel folding pathways and gives a measure of the closeness of the TSE to the folded state. Evidence for multiple folding pathways also can be found in the mildly heterogeneous kinetics observed in our data, which are more pronounced in the presence of Mg2+ (Figures S3 and S4). In fact, tight Mg2+ binding sites are thought to be a major cause of heterogeneous folding of nucleic acid structures containing helical junctions and/or loops.42–44

Our Bsu aptamer contains three loops of varying sizes, with an extended loop L2 that is typicall of H-type pseudoknots.18 Differential occupancy of Mg2+ around these loops and P2 could thus cause the observed heterogeneous kinetics.42–44 In contrast, the small heterogeneity mostly in kundock in the absence of Mg2+ may stem from a ligand-free open conformation that exists as a broad ensemble exhibiting different degrees of P1–L1 interaction with the 3’ A-rich tail, as suggested by our previous smFRET studies.51 Notably, our use of average rate constants from single-exponential fits then yields insights into the nature of the overall folding mechanism that can be related with those obtained from classical, ensemble-averaging kinetic measurements.30

Previous smFRET studies on riboswitches only probed the influence of the single cognate ligand on the folding pathway.13,14 However, cells contain many closely related metabolites at concentrations similar to or higher than the cognate ligand. In certain cases, it was suggested that noncognate ligands with lower affinity can in principle bind to and affect the gene regulation of riboswitches.13,14 Our work directly addresses this notion for the preQ1-I riboswitch. Interestingly, our kinetic data show that in the absence of Mg2+ the equilibrium constant for docking, kundock, follows the expected trend of binding affinities, with preQ1 > preQ0 > guanine (Figure 4a).13 While kundock in the presence of near-physiological Mg2+ follows a similar trend, the difference between the ligands decreases significantly, with the nearest noncognate ligand preQ1 adopting an equilibrium constant similar to that of preQ0 (Figure 4b). This suggests that, under in vivo conditions, preQ0 may stabilize the native conformation of the riboswitch to act as a “proxy” for preQ1. Given that preQ0 is an immediate upstream intermediate of preQ1 in the biosynthetic pathway controlled by the riboswitch in its entirety, such redundancy may serve the role of more robustly regulating gene expression.

Distinguishing the IF and CS folding mechanisms traditionally has been difficult using ensemble methods so that many studies have incorrectly posited that folding occurred via a CS mechanism based solely on the observation of folded-like conformations in the absence of ligand.6,45 Importantly, it has been noted that studying the effects of ligand on the relaxation kinetics of (un)folding (kundock) using ensemble experiments cannot unambiguously distinguish between the two mechanisms.8 By contrast, single-molecule techniques, including force spectroscopy using optical tweezers and smFRET as demonstrated here, are ideal tools for dissecting the individual rate constants kundock and kdock and thus for directly unraveling the ligand binding mechanism(s). We have introduced smTSA to further strengthen our conclusions while highlighting the importance of stacking interactions during IF when the ligand binds to the open (undocked) conformation, as well as of specific hydrogen bonds during CS when the ligand binds to the folded-like (Docked) conformation of the preQ1 aptamer. Future three-color smFRET studies employing a ligand labeled with a third fluorophore or a dark quencher may be applied to visualize the arrival time of the ligand relative to pseudoknot docking.42 However, labeling may be nontrivial to accomplish because most ligands are tightly shielded in their RNA binding pockets so that fluorophore appendages may easily interfere with binding. For certain riboswitches, such as the hydroxocobalamin riboswitch, the ligand itself acts as a quencher and, therefore, can report on the nature of the ligand-binding conformation at the single-molecule level.30 Alternatively, ensemble flux calculations using rigorous analysis of kinetic measurements could be carried out but have only been applied recently to study ligand-binding coupled protein folding.51,61

In summary, we have shown that combined kinetic and transition-state analyses using smFRET are a powerful toolset to dissect the exquisite interdependence of ligand- and Mg2+-mediated folding mechanisms of a riboswitch. We anticipate that this work will pave the way for deciphering the coupled ligand binding and folding pathways of many riboswitch RNAs, which will render them attractive antibiotic drug targets.

4. MATERIALS AND METHODS

4.1. Labeling and Purification of RNA. The Bsu preQ1 riboswitch aptamer51,52 with the sequence 5′-UGCGGGAGAGG-UUUCUAGC(5-N-U)ACACCCUCUAUAAAAAACUAA-3′ was chemically synthesized by Dharmacon, Inc. (Fayette, Colorado) with a 5′-biotin, 3′-Dy547, and an internal 5-aminoallyl uridine (5-N-U) at U12. The sequence shown in bold is the minimal aptamer construct that was crystallized, and the numbering used is consistent with this construct.18 The RNA was deprotected following the manufacturer’s protocol and labeled with Cy5-NHS ester (GE Healthcare) as previously described.14 Excess dye was removed using a NAP-5 gel filtration column (GE Healthcare) followed by ethanol precipitation. The doubly labeled RNA was dissolved in autoclaved, deionized water and used for performing smFRET experiments.

4.2. Single-Molecule FRET. Single-molecule FRET experiments were performed using prism-based total internal reflection fluorescence (TIRF) microscopy.11,14 The doubly labeled RNA with a 5′-biotin was immobilized onto the surface of clean quartz slides containing a microfluidic channel using the biotin–streptavidin interaction. Prior to immobilization, the aptamer was folded by heating a low concentration (20–50 pM) of RNA at 90 °C for 1 min in 1× smFRET buffer (50 mM Tris-HCl, pH 7.5, 100 mM KCl) without Mg2+ and allowed to slowly cool down to room temperature over 15 min. The heat-annaled RNA was used for immobilization onto the slide surface, and the unbound molecules were washed away using the 1× smFRET buffer (50 mM KCl). Ligand titrations were performed on the same slide in the 1× smFRET buffer with or without 1 mM Mg2+. An enzymatic oxygen scavenging system consisting of 5 mM protocatechuc acid (PCA), protocatechuic-3,4-dioxygenase (PCD), and 2 mM Trolax (6-hydroxy-2,5,7,8-tetramethyl-chroman-2-carboxylic acid) was used to extend the life of fluorophores and to reduce photobleaching.14,64 Molecules were imaged at ~16 Hz time resolution using an intensified charge-coupled device camera (ICCD, I-Pentamax, Princeton Instruments). A 532 nm green laser was used to excite Dy547, and fluorescence from both Dy547 and Cy5 were recorded from which FRET efficiency was calculated as I55/ (I55 + I54), where I55 and I54 stand for the background-corrected intensities of Dy547 (donor) and Cy5 (acceptor), respectively. Raw movies were processed using IDL (Research Systems, Inc.) to generate smFRET time traces that were further analyzed using custom written software.
MATLAB (The Math Works) scripts. smFRET traces displaying single-step photobleaching, a signal-to-noise ratio of >4, and a total fluorescence intensity \((I_0 + I_f)\) of >300 (arbitrary units) were manually selected for kinetic analysis. The traces were idealized with a two-state model using hidden-Markov modeling (HMM) with a segmental k-means algorithm as implemented in the program QbB.48,49 Dwell times in the undocked and docked states were extracted from all the idealized traces, and the cumulative dwell time distributions were fit with a single-exponential function to obtain the rate constants \(k_{\text{dock}}\) and \(k_{\text{undock}}\) respectively. A minimum of 70 smFRET traces showing multiple transitions were included for kinetic analysis for every experiment so that more than 250 dynamic traces were used to obtain the mean and standard deviation (SD) of all the rate constants. Heterogeneity in the kinetic behavior was observed in some conditions, mainly in the docking rates in the presence of Mg\(^{2+}\). In such cases, fitting with a single-exponential function reports an average rate constant value. Ligand-titration experiments in the presence or absence of Mg\(^{2+}\) were performed for each ligand in triplexes, and the mean ± SD for the rate constants were plotted as Figures 2 and 3. Cross-correlation analysis of the smFRET traces in the absence of ligand was performed as described previously.\(^{50}\)

## 4.3. Single-Molecule Transition-State Analysis (smTSA)

Single-molecule transition-state (or Φ-value) analysis was carried out as previously described.\(^{26}\) The rate constants of docking and undocking at saturating ligand concentrations \(k_{\text{dock}}\) and \(k_{\text{undock}}\) were used to calculate changes in free energy barrier for docking, \(\Delta \Delta G_{\text{dock}}\), and the overall equilibrium stability, \(\Delta \Delta G_{\text{undock}}\) of preQ- and guanine-bound complexes as compared to the preQ-bound RNA at 295 K. The \(k_{\text{dock}}\) versus \([L]\) data in the absence or presence of Mg\(^{2+}\) as shown in Figures 2 and 3 were fit with the following single-site (noncooperative) binding Hill equation to yield \(k_{\text{dock}}\):

\[
k_{\text{dock}} = \frac{[L] \times (k_{\text{dock}}^\infty - k_{\text{dock}}^0)}{[L] + K_{1/2}} + k_{\text{dock}}^0
\]

(1)

In the absence of Mg\(^{2+}\), the ligand-free preQ-1 aptamer exists mainly in a hairpin conformation without any tertiary interactions as shown by NMR and supported by our previous smFRET analysis, indicating that the molecules exist mainly in a static form.\(^{52}\) The rate constants of docking and undocking \(k_{\text{dock}}\) and \(k_{\text{undock}}\) were used to calculate the changes in free energy barrier for docking, \(\Delta \Delta G_{\text{dock}}\), and the overall equilibrium stability, \(\Delta \Delta G_{\text{undock}}\) of preQ- and guanine-bound complexes as compared to the preQ-bound RNA at 295 K. The \(k_{\text{undock}}\) versus \([L]\) data in the absence or presence of Mg\(^{2+}\) as shown in Figures 2 and 3 were fit with the following single-site (noncooperative) binding Hill equation to yield \(k_{\text{undock}}\):

\[
k_{\text{undock}} = \frac{k_{\text{undock}}^0}{1 + [L]/K_{1/2}} + k_{\text{undock}}^0
\]

(2)

which describes the decrease in \(k_{\text{undock}}\) in increasing \([L]\) for the CS mechanism of folding, where \(k_{\text{undock}}^0\) and \(k_{\text{undock}}\) refer to \(k_{\text{undock}}\) at \([L] = 0\) and at \([L] = \infty\) (saturating \([L]\)), respectively.\(^{26,48}\) An additional \(k_{\text{undock}}\) term was included in the equation to account for the nonzero undocking rate even at saturating \([L]\). The value of \(k_{\text{undock}}^0\) was estimated from cross-correlation analysis of smFRET traces in the presence of 1 mM Mg\(^{2+}\) alone, where a majority of the smFRET traces were static.\(^{21}\) However, a small fraction (~20–30%) showed fast transitions close to the time resolution (~16 Hz) of the camera (Figure S5), from which we can estimate the undocking rate constant (from 20 traces) \(k_{\text{undock}} + k_{\text{undock}}\) of ~12.24 s\(^{-1}\) was observed that likely represents a lower limit for the rate constant of structural transition. Therefore, the value of \(k_{\text{undock}}^0\) was fixed at 16 s\(^{-1}\), which is the time resolution of our experiments and a lower estimate of the true value. \(k_{\text{undock}}\) was then obtained from fitting the data with eq 2, which is not sensitive to the value of \(k_{\text{undock}}^0\) The standard deviations reported on the Φ-values were obtained from carrying out error-propagation analysis using the uncertainties in the rate constants \(k_{\text{dock}}\) and \(k_{\text{undock}}\) (Tables S1 and S2) obtained from nonlinear curve fitting to eqs 1 and 2.

### ASSOCIATED CONTENT

#### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b09740.

Error propagation of smTSA, kinetic fit parameters, raw smFRET traces, cumulative dwell time distributions, and cross-correlation analyses (PDF)

#### AUTHOR INFORMATION

**Corresponding Author**

*nwalter@umich.edu*

**Notes**

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

N.G.W. thanks Prof. George Garcia for gifts of preQ\(_1\) and preQ\(_2\). This research was supported by the National Institutes of Health grant GM062357 and a subcontract on GM063162 to N.G.W.

### REFERENCES