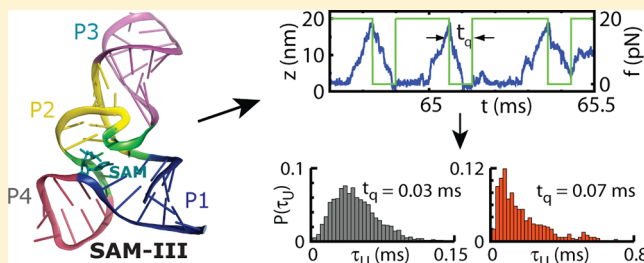


# Kinetics of Allosteric Transitions in S-adenosylmethionine Riboswitch Are Accurately Predicted from the Folding Landscape

Jong-Chin Lin and D. Thirumalai\*

Department of Chemistry and Biochemistry, Biophysics Program, Institute for Physical Sciences and Technology, University of Maryland, College Park, Maryland 20742, United States

**ABSTRACT:** Riboswitches are RNA elements that allosterically regulate gene expression by binding cellular metabolites. The SAM-III riboswitch, one of several classes that binds S-adenosylmethionine (SAM), represses translation upon binding SAM (OFF state) by encrypting the ribosome binding sequence. We have carried out simulations of the RNA by applying mechanical force ( $f$ ) to the ends of SAM-III, with and without SAM, to get quantitative insights into the  $f$ -dependent structural changes. Force–extension ( $z$ ) curves (FECs) for the apo (ON) state, obtained in simulations in which  $f$  is increased at a constant loading rate, show three intermediates, with the first one being the rupture of SAM binding region, which is greatly stabilized in the OFF state. Force-dependent free energy profiles,  $G(z, f)$ , as a function of  $z$ , obtained in equilibrium constant force simulations, reveal the intermediates observed in FECs. The predicted stability difference between the ON and OFF states using  $G(z, f)$  is in excellent agreement with experiments. Remarkably, using  $G(z, f)$ s and estimate of an effective diffusion constant at a single value of  $f$  allows us to predict the  $f$ -dependent transition rates using theory of first passage times for both the apo and holo states. To resolve the kinetics of assembly of SAM-III riboswitch in structural terms, we use force stretch–quench pulse sequences in which the force on RNA is maintained at a low ( $f_q$ ) value starting from a high value for a time period  $t_q$ . Variation of  $t_q$  over a wide range results in resolution of elusive states involved in the SAM binding pocket and leads to accurate determination of folding times down to  $f_q = 0$ . Quantitative measure of the folding kinetics, obtained from the folding landscape, allows us to propose that, in contrast to riboswitches regulating transcription, SAM-III functions under thermodynamic control provided the basal concentration of SAM exceeds a small critical value. All of the predictions are amenable to tests in single molecule pulling experiments.



## INTRODUCTION

About a decade ago it was established that riboswitches, which are RNA elements found in the untranslated region of mRNA, regulate gene expression in bacteria by binding to metabolites with exquisite specificity.<sup>1–4</sup> In the intervening years, fundamental insights into their functions have been obtained through a remarkable number of experiments ranging from X-ray crystallography enabling the determination of their structures,<sup>5</sup> biophysical techniques elucidating the mechanistic aspects, single molecule experiments,<sup>6,7</sup> and more recently theory and simulations.<sup>8–14</sup> Riboswitches consist of a conserved aptamer domain to which the metabolite binds, resulting in a conformational change and alteration in the folding pattern of the downstream expression platform leading to control of gene expression. Remarkably, there are variations in the functions of riboswitches even among structurally similar aptamer domains. For example, *add* adenine (A) riboswitch activates translation upon binding the metabolite purine, whereas the structurally similar *pbuE* adenine riboswitch activates transcription in the presence of purine. Both of these are ON riboswitches, which means that translation or transcription is activated only upon binding of the metabolite. In contrast, an OFF riboswitch (FMN for example) shuts down gene expression when the metabolite binds to the aptamer domain. Finally, some of the

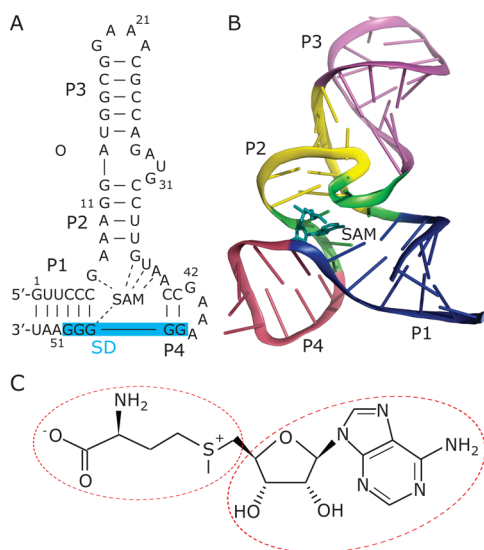
riboswitches are under kinetic control (FMN riboswitch<sup>15</sup> and *pbuE* adenine riboswitch<sup>16</sup>), whereas others (for example, *add* adenine riboswitch) are under thermodynamic control.

A first step toward a quantitative understanding of the diverse functions of riboswitches is to map their entire folding landscape so that transitions between various states involved in the conformational rearrangements can be quantitatively elucidated. Here, we consider a riboswitch that binds S-adenosylmethionine (SAM) and undergoes an allosteric transition thus enabling control of translation. At least five distinct classes of riboswitches that bind SAM or its derivative S-adenosylhomocysteine (SAH) have been identified: (i) SAM-I/SAM-IV riboswitches, carrying a common metabolite binding core,<sup>17,18</sup> (ii) SAM-II/SAM-V riboswitches,<sup>19,20</sup> which also carry a common metabolite binding core, (iii) SAM-III or SMK riboswitches,<sup>21</sup> (iv) SAM/SAH riboswitches that do not discriminate between SAM and SAH,<sup>22</sup> and (v) SAH riboswitches that selectively bind SAH, and strongly discriminate against SAM.<sup>23</sup> Both SAM-II/SAM-V and SAM-III riboswitches are unusual because the residues that control gene expression are also the binding sites for the SAM

Received: August 19, 2013

Published: October 2, 2013

ligand.<sup>24,25</sup> The SAM-III riboswitch, found in the *metK* gene (which encodes SAM synthetase) from *Lactobacillales* species, inhibits translation by sequestering the Shine–Dalgarno (SD) sequence (Figure 1), which is essential for engaging the 30S



**Figure 1.** (A) The secondary structure of SAM-III riboswitch. The blue shaded area shows the SD sequence, GGGGG. (B) The tertiary structure of SAM-III riboswitch. (C) The structure of the metabolite SAM, which in the coarse-grained simulations is represented using two interaction centers corresponding to the regions shown in red dotted lines. One of the centers corresponds to the adenosol (base and the ribose on the right) and the other represents the methionine moiety on the left.

ribosomal subunit for translation initiation.<sup>21</sup> When SAM is bound, the somewhat atypical SD sequence (GGGGG shown in the blue shaded area in Figure 1A) is sequestered by base pairing with the anti-SD (ASD) sequence, which hinders the binding of the 30S ribosomal subunits to mRNA. In the absence of SAM (Figure 1C shows the metabolite), the sequence outside the binding domain is involved to enable the riboswitch to adopt an alternative folding pattern, in which the SD sequence is exposed and free to bind the ribosomal subunit.<sup>26</sup> In this sense, SAM-III is an OFF switch for translation, in contrast to *add* A-riboswitch, which is an ON switch for translation. Regulation of gene expression, which depends on the concentration of SAM, is determined by competition between the SD-ASD pairing and loading of ribosomal subunit onto the SD sequence. In order to function as a switch, the SD sequence has to be exposed for ribosome recognition, which implies that at least part of the riboswitch structure accommodating SAM III has to unfold (Figure 1). Thus, it is important to quantitatively determine the folding landscape and the rates of conformational transitions between the ON and OFF states of the SAM-III riboswitch to understand its function.<sup>27</sup>

Single molecule pulling experiments have produced the folding landscapes of several riboswitches<sup>27</sup> including the two (*add* and *pbuE*) that bind the metabolite purine.<sup>28,29</sup> Although both these riboswitches are structurally similar their folding landscapes are different, which in part explains the remarkable discriminating capacity of the two riboswitches. These experiments and related computations<sup>8</sup> demonstrated that quantitative insights into functions of riboswitches may be

obtained by understanding the detailed folding landscapes of metabolite induced conformational change in the riboswitches. Our predictions for the free energy profile for *add* A-riboswitch were nearly quantitatively validated using laser optical tweezer experiments.<sup>29</sup> In addition, our theory that the stability of isolated helices in the three-way junction determines the order of unfolding of the paired helices in the *add* and *pbuE* A-riboswitches explained the differences in the measured sequence-dependent folding landscapes.

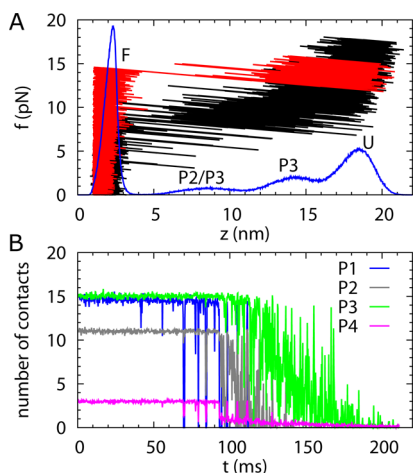
Here, we study the force-triggered unfolding and refolding of SAM-III riboswitch using a coarse-grained self-organized polymer (SOP) model using Langevin dynamics simulations.<sup>30,31</sup> The method has been successfully applied to determine the sequence-dependent folding landscapes of the purine riboswitch aptamers and other RNA structures.<sup>8,30</sup> We use similar models to make a number of predictions for future single molecule laser optical tweezer (LOT) experiments for SAM-III riboswitch. (i) The force-dependent free energy profiles as a function of extension of SAM-III show two intermediates, besides the unfolded and folded states, which are populated both in equilibrium and kinetic simulations. Binding of SAM greatly stabilizes the native state and increases the barrier to formation of alternate structures required to function as a switch. (ii) Remarkably, we find that the free energy profiles along with an estimate of a collective diffusion coefficient can be used in the theory of first passage times to predict the hopping rates between the states of SAM-III, in quantitative agreement with simulations. (iii) We implement a novel force stretch–quench protocol to resolve elusive states that may be difficult to detect in conventional (force ramp or force clamp) single molecule pulling experiments. (iv) The simulated folding landscapes and transition rates are used to produce a model for function of SAM-III riboswitch, demonstrating that the function of this riboswitch is under thermodynamic control.

## RESULTS AND DISCUSSION

### Structures of SAM-III Riboswitch in the Simulations.

The SAM-III riboswitch, with an inverted Y-shape, has four helices, P1, P2, P3, and P4 (Figure 1), in the SAM-bound ('OFF') state. The metabolite, SAM, is bound to the three-way junction formed by helices P1, P2, and P4. We calculated the complete folding landscape of the binding domain of the SAM-riboswitch using the native structure with SAM bound from the crystal structure<sup>25</sup> (PDB code: 3ESC). In order to decipher the rates of switching between the 'ON' and 'OFF' states, we also performed simulations without SAM bound. The native structure of the metabolite-free riboswitch aptamer for coarse-grained simulations was taken from the equilibrated structure generated during a 50 ns all atom molecular simulation run starting from the bound-state crystal structure (3ESC) with SAM removed. The full SAM-III riboswitch favors an alternative folding pattern with the SD sequence free to engage the ribosomal subunit in the absence of SAM. However, to function as a switch the RNA has to allosterically change its conformation to an ensemble of structures committed to bind the metabolite. Thus simulations in the presence and absence of SAM are needed to describe the molecular basis of function.

**Force-Ramp Simulations.** In the absence of the metabolite, the force–extension curve (FEC), at a constant loading rate of 96 pN/s, shows that there are two intermediate states, with extension  $z \sim 14$  and 9 nm (black curve in Figure 2A). At the unfolding force of  $\sim 9$  pN, helices P1 and P4



**Figure 2.** (A) FEC for SAM-III riboswitch in the absence of metabolite (in black) when stretched at a constant loading rate of 96 pN/s. The distribution of the extension (shown in blue) during the force ramp simulation shows multiple intermediate states: (i) *F*, the folded state; (ii) *P2/P3*, where only *P2* and *P3* are folded; (iii) *P3*, where only *P3* is folded; and (iv) *U*, unfolded state. The red curve is FEC with SAM bound. (B) Average number of contacts (over every 100 time steps at which coordinates are recorded) within different helices as a function of time for the riboswitch without SAM during a force ramp simulation. We define that two nucleotides form a contact if the distance between them is less than  $R_c = 1.3$  nm.

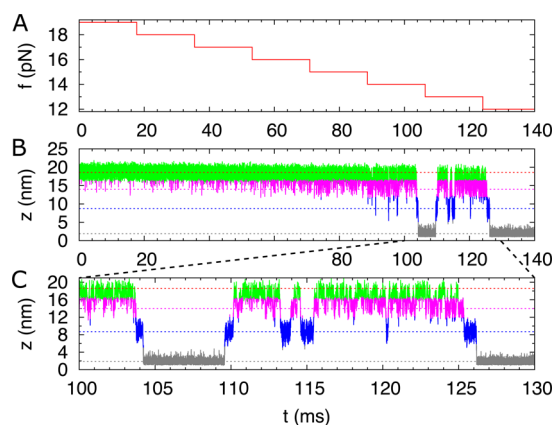
rupture in a single step, and at  $f \sim 10 - 11$  pN, a second unfolding step occurs, resulting in the unraveling of *P2*. Helix *P3* unfolds fully only when  $f$  exceeds beyond 12 pN. When SAM is bound, the riboswitch unfolds cooperatively in a single step at  $f \sim 16$  pN (Figure 2A). The distribution of histograms of extensions (blue curve in Figure 2A), shows the presence of two intermediate states ahead of global unfolding. The peak, at  $z \sim 9$  nm, corresponds to rupture of *P1* and *P4* helices (see below for additional evidence), which are the initial unfolding events. At this stage of unfolding *P2* and *P3* are still intact. The last helix to unfold is *P3*, corresponding to the peak at  $z \sim 14$  nm. Thus, upon unfolding at a constant loading rate, the hierarchical unfolding pathway of SAM-III riboswitch is  $F \rightarrow \Delta P1\Delta P4 \rightarrow \Delta P2 \rightarrow U$ , where  $\Delta P1\Delta P4$  means helices *P1* and *P4* are ruptured, and  $\Delta P2$  represents additional unfolding of *P2*. At this stage, we note that pulling at a constant loading rate does not resolve the order of rupture of *P1* and *P4*, an issue that we address below using a different protocol to alter forces.

The order of unfolding is also reflected in the unfolding dynamics, which is illustrated for a representative trajectory in Figure 2B using the time dependent changes in the number of contacts involving the helices. In particular, the intermediate states, at  $z \sim 14$  and 9 nm are identified, by tracking the number of contacts within the helices (Figure 2B). The order of unfolding, represented by rupture of contacts involving the helices, shows that *P1* and *P4* are the first to unfold, occurring nearly simultaneously. Subsequently, *P2* is destabilized, which is then followed by *P3* unfolding. Thus, both in FEC as well as in the unfolding trajectories SAM-III unfold by the same route.

After the riboswitch is completely unfolded, we reduced the force, with the same loading rate but with the opposite sign, to initiate refolding. During the refolding process, the two intermediate states with extension,  $z = 14$  and 9 nm, corresponding to *P3* and *P2/P3* folded states, respectively, are formed. Taken together, these results demonstrate that the

same intermediates are populated during the folding and unfolding processes. Therefore, we conclude that these are equilibrium intermediates in the protocol used in our pulling simulations.

**Refolding upon Stepwise Reduction in Force.** In single molecule pulling experiments different protocols can be used to quench the force from high to low values.<sup>32,33</sup> Using a stepwise reduction in force (Figure 3A), we initiated refolding to



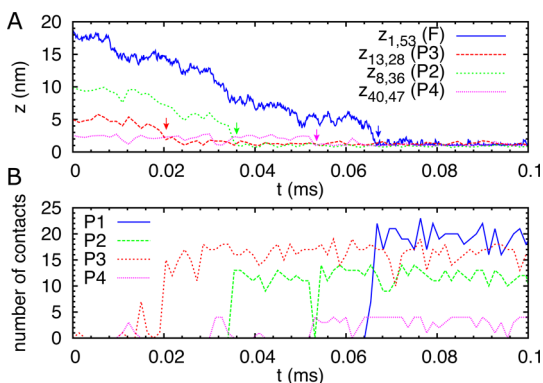
**Figure 3.** (A) Step wise reduction in force starting from an initial stretch force  $f_s$  to a final quench force  $f_q$ . The force starts at  $f_s = 19$  pN and decreases by 1 pN for every 17.7 ms until  $f_q = 12$  pN is reached. (B) End-to-end distance as a function of time for the refolding of the riboswitch in the presence of metabolite SAM upon stepwise decrease of force as shown in (A). (C) A blowup of the time traces of the extension from (B) between  $t = 100$  and 130 ms, during which the  $f$  is reduced to 13 pN. Hopping between multiple values of the extension is seen.

determine the order of folding of SAM-III riboswitch. Starting from  $f = 19$  pN, we decreased the force by 1 pN every 17.7 ms (Figure 3A), which corresponds to a loading rate of  $\sim 56$  pN/s. A sample folding trajectory, measuring the end-to-end distance (or extension) of the SAM-III riboswitch as a function of time (Figure 3B), shows that when the force is decreased to a value  $< 16$  pN,  $z(t)$  decreases to  $\sim 14$  nm, which signals folding of *P3*. Note that formation of *P3* when  $z \sim 14$  nm (Figure 3B) corresponds to the peak at the same distance in Figure 2A, reporting the histograms of  $z$  during pulling simulations at a constant loading rate. Only at  $t > 90$  ms there are signs of formation of *P2* (indicated by a small plateau at  $z \sim 9$  nm in Figure 3B,C, and a peak in  $P(z)$  at the same distance in Figure 2A). At  $f \sim 14$  pN, *P4* and *P1* fold nearly simultaneously (indicated by gray lines in  $z(t)$  in Figure 3B). Thus, even the protocol involving a stepwise reduction of force cannot fully resolve the folding of *P1* and *P4* because they form nearly concurrently.

If the value of the force is fixed at  $f = 13$  pN the riboswitch hops between the four states (*U*, *P3*, *P2/P3*, and *F*) (Figure 3C). Interestingly, Figure 3C shows that there is no direct transition to the folded state from either *U* or *P3*. In contrast, *P2/P3* state (blue color,  $z \sim (8 - 10)$  nm)) is kinetically connected to both the *F* and *U* states. Thus, it is likely that the major transition state to unfolding involves the rupture of *P3/P2* states.

We also simulated the refolding of SAM-III riboswitch by quenching the force from an initial high force to a constant low force value in a single step. When the force is quenched from 20 to 0 pN, a completely unfolded SAM-III riboswitch is fully

folded in <0.1 ms (Figure 4A). The simulations, reporting distances between specific nucleotides on the four helices, show

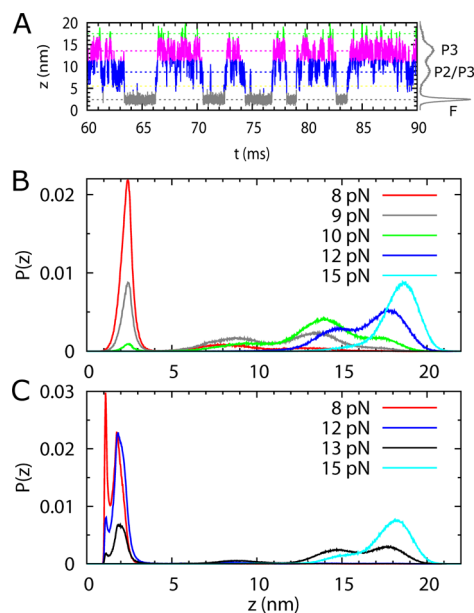


**Figure 4.** (A) A typical folding trajectory expressed using various distances  $z_{ij}$  between nucleotides  $i$  and  $j$  of the riboswitch, with bound SAM, as a function of time when the force is quenched from  $f_s = 19$  pN to  $f_f = 0$  pN. The nucleotides  $i$  and  $j$  are chosen as reporters of the individual helices and the folded state. (B) Number of contacts within different helices as a function of  $t$  during refolding upon force quench. The riboswitch folds in four clear steps: (i) P3 folds at  $t \sim 0.01$  ms, (ii) P2 at  $t \sim 0.032$  ms, (iii) P4 folds at  $t \sim 0.052$  ms, and (iv) P1 folds, resulting in the fully folded state at  $t = \tau_F \sim 0.065$  ms. Here,  $\tau_F$  is the folding time at  $f_s = 0$  pN.

a folding pathway in which P3 folds first, P2 folds second, then P4, and finally P1 folds ( $U \rightarrow P3 \rightarrow P3/P2 \rightarrow P3/P2/P4 \rightarrow F$ ), as shown in Figure 4B. Thus, all probes of refolding establish that folding and unfolding occur by the same pathway in a hierarchical manner.

**Force-Dependent Folding Landscapes.** The multiple steps in the hierarchical folding of SAM-III are reflected in the probability distributions,  $P(z)$ , of the extension,  $z$ . We calculated  $P(z)$  from the folding trajectories, reporting  $z$  as a function of  $t$  (see Figure 5A for a sample trajectory at  $f = 9$  pN). The free energy profiles were obtained from the  $P(z)$  using  $G(z_f) = -k_B T \ln P(z)$ . In the absence of SAM, when  $f < 9$  pN, the folded state (Figures 5 and 6A) is the most stable state. When  $9$  pN  $< f < 11$  pN, the most stable state becomes the P3 folded state. Binding of SAM consolidates the formation of helix P1 and P4, further stabilizing the folded state. As shown in Figure 6A, at  $f = 9$  pN, SAM binding stabilizes the folded state by  $\sim 12 k_B T$  and increases the energy barrier for leaving the folded state by  $\sim 3 k_B T$ . The distance from the folded state ( $\sim 2$  nm) to the first barrier in the absence of SAM is  $\sim 2$  nm, which indicates unzipping of 2.5 base pairs, assuming a contour length increase of 0.4 nm/nt. The position of the barrier is in the middle of helix P1, which means that the main barrier in the first step toward global unfolding of the riboswitch is the unzipping of P1. In the presence of SAM, the position of the first barrier shifts to  $\sim 5$  nm, implying that 4 base pairs of P1 next to the nucleotide G48 (Figure 1A) that has direct contacts with SAM are ruptured at the transition state. Thus, disruption of contacts with SAM becomes the key barrier in the first unfolding step.

**Hopping Rates from Folding Landscapes.** The utility of the free energy profiles,  $G(z, f)$ , is that they can be used to obtain force-dependent transition rates using the theory of mean first-passage times.<sup>8,34</sup> The rate of transition, from an initial point  $x_A$  in one state to the final point  $x_B$  in another, is given by

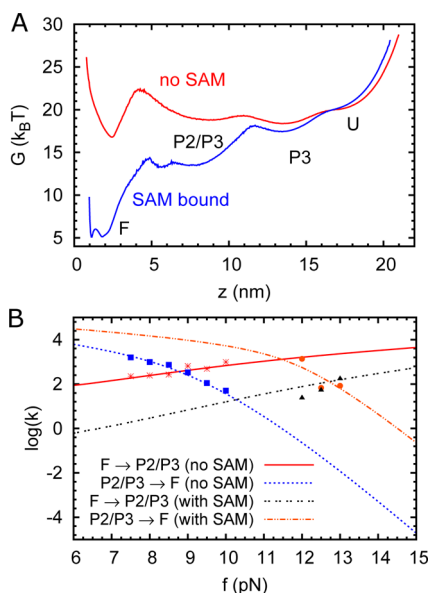


**Figure 5.** (A) Time trace of extension,  $z$ , of the SAM-III riboswitch without the metabolite at  $f = 9$  pN. (B) and (C) From time traces, such as the one in (A), we calculated the probability distributions,  $P(z)$ s of extension  $z$ . The time traces were obtained from constant force simulations with each at least 150 ms long. (B) Distributions  $P(z)$ s in the apo state of the RNA at various  $f$  values.  $P(z)$  at  $f = 9$  pN (green curve) shows four peaks, corresponding to  $F$ ,  $P2/P3$ , and  $P3$  states, and one at  $z \sim 17.5$  nm representing the unstable  $U$  state. The  $\Delta P1$  state ( $P1$  unfolded, with other helices intact), at  $z \sim 5.5$  nm, is visited transiently, with very low probability. The riboswitch stays mostly unfolded when  $f > 11$  pN. In the force range,  $9$  pN  $< f < 11$  pN, the  $P3$  folded state is the most stable. At  $f = 8$  pN the riboswitch is fully folded (red curve). The riboswitch goes from the unfolded to the folded state following the pathway,  $U \rightarrow P3 \rightarrow P2/P3 \rightarrow F$ , when force decreases from a large to a small value. (C) Same as (B) except SAM is bound to RNA showing that the intermediate states in (B) are only transiently present. In the presence of SAM, the riboswitch stays mostly unfolded when  $f > 13$  pN. As  $f$  decreases to below 13 pN, the riboswitch transits to the folded state with  $z = 2$  nm with G1 and U53 not paired (Figure 1A). At forces  $f < 8$  pN, G1 and U53 forms a base pair with  $z = 1$  nm.

$$k_{A \rightarrow B}^{-1}(f) = \bar{\tau}_{A \rightarrow B}(f) = \int_{x_A}^{x_B} dy e^{\beta G(f,y)} \frac{1}{D_{A \rightarrow B}} \int_{x_0}^y dz e^{-\beta G(f,z)} \quad (1)$$

where  $D_{A \rightarrow B}$  is the diffusion coefficient along the extension coordinate for transition between states A and B,  $\beta = 1/(k_B T)$ ,  $x_0 = -\infty$  if  $x_A < x_B$ , and  $x_0 = \infty$  if  $x_A > x_B$ . Following our earlier work,<sup>8</sup> we calculated the diffusion coefficient,  $D_{A \rightarrow B}$  by equating the transition rate calculated using eq 1 to that obtained from time traces obtained in simulations at one particular force value, which for the SAM-III riboswitch is  $f = 8$  pN.

In order to establish the efficacy of eq 1, we used  $G(f,z)$ , calculated using simulations and obtained the hopping rates  $k_{F \rightarrow P2/P3}$  and  $k_{P2/P3 \rightarrow F}$ . The results can be directly compared to the rates calculated from the folding trajectories (see Figure 5A for an example). We chose the  $P2/P3$  ( $P1$  and  $P4$  disrupted) because this state is poised to bind SAM and hence is most relevant for control of translation. Moreover, the  $P2/P3$  state (both  $P2$  and  $P3$  are folded, while  $P1$  and  $P4$  are disrupted) is kinetically connected to the  $F$  state. In the absence of the metabolite SAM, the transition rates between the folded state and the  $P2/P3$  state calculated using eq 1 are in good



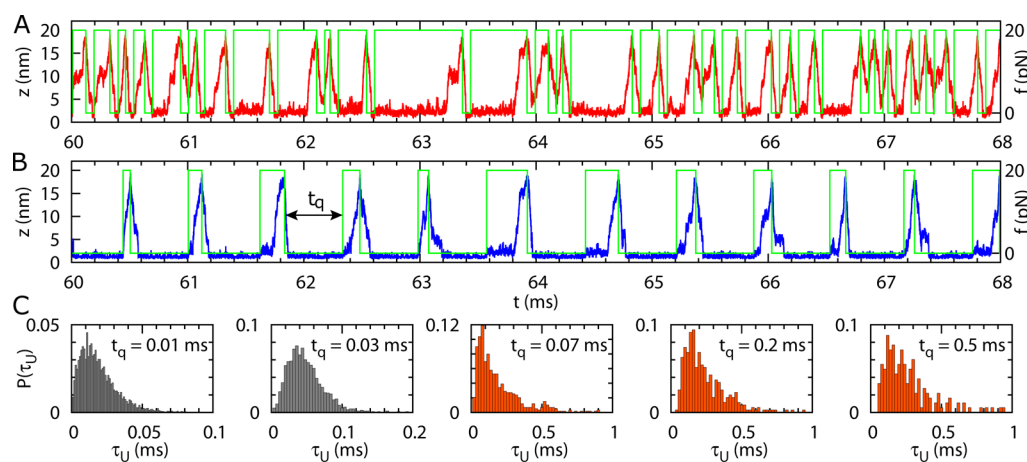
**Figure 6.** (a) Free energy profiles as a function of  $z$  for SAM-III riboswitch without (red) and with (purple) SAM at  $f = 9$  pN. Without SAM, the folded state is slightly more stable than P2/P3 and P3 states by 1–2 kcal/mol. Binding of SAM significantly increases the stability of the folded state. (b) Transition rates between the folded state ( $F$ ) and P2/P3 (both P2 and P3 are folded, but P1 and P4 disrupted) states without and with bound SAM. The lines show the rates calculated using eq 1, and the symbols are the rates obtained directly from an ensemble of folding trajectories,  $z(t)$ . The diffusion coefficient is chosen as  $D \sim 1.24 \times 10^5 \text{ nm}^2/\text{s}$  by equating the rates of  $F \rightarrow \text{P2/P3}$  transition at  $f = 8$  pN calculated by the two different methods. In the absence of SAM, the transition rate between  $F$  and P2/P3 state is  $\sim 10^2\text{--}10^3 \text{ s}^{-1}$  at  $f \sim 9$  pN. The SAM binding stabilizes the folded state and reduces the  $F \rightarrow \text{P2/P3}$  transition rate by 2 orders of magnitude.

agreement to those obtained from time traces of extension, as shown in Figure 6B. The force at which  $k_{F \rightarrow \text{P2/P3}} \approx k_{\text{P2/P3} \rightarrow F}$  without SAM, is about 9 pN, and the associated energy barrier for leaving the folded state is about  $6k_B T$ . The transition rate is about  $10^2\text{--}10^3 \text{ s}^{-1}$ . Binding of SAM stabilizes the folded state,

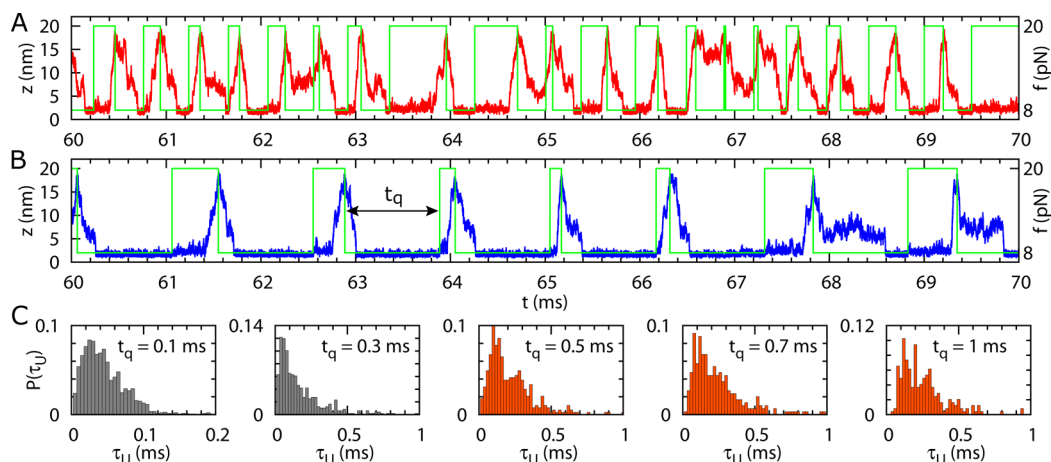
and increases the energy barrier for escaping from the folded state by  $\sim 3k_B T$ , which reduces the rate of  $F \rightarrow \text{P3}$  transition by 2 orders of magnitude at  $f = 9$  pN. The force at which  $k_{F \rightarrow \text{P2/P3}} \approx k_{\text{P2/P3} \rightarrow F}$  increases to  $\sim 13$  pN. The estimated transition rate for the  $F \rightarrow \text{P2/P3}$  transition, at zero force, is on the order of  $10 \text{ s}^{-1}$ . With the binding of SAM, the rate of unfolding P1 and P4 decreases significantly (by nearly 3 orders of magnitude) due to the deeper well of the folded state and higher energy barrier for transition from the folded state. This stabilizes the formation of helices P1 and P4 and hinders the binding of 30S ribosomal subunits to the SD sequence of the SAM-III riboswitch. Consequently, translation is inhibited.

The folding landscape of the SAM-III riboswitch shows that rupture of P3 is the last step on global unfolding and only can be disrupted when  $f > 12$  pN. This is relevant because switching between alternative folding patterns, controlled by the metabolite concentration, for a full SAM-III riboswitch only involves the unfolding of P1, P4, and P2.<sup>26</sup> Helix P3 is folded in both the ‘OFF’ and the ‘ON’ states, which leads to the conclusion that P3 is stable during translation.

**Force Stretch–Quench ( $f_s\text{--}f_q$ ) Cycles.** In the conventional use of the single molecule force spectroscopy (SMFS) (force ramp or force clamp) only those states that are populated in the accessible force range can be determined. The ability to resolve all accessible states is determined by a number of factors including the drag on the beads, handle characteristics, and instrumental time resolution. In order to decipher if hidden states are populated, it is possible to use SMFS in unconventional ways. We showed elsewhere<sup>35</sup> that force pulses during stretch–quench cycles can be used to uncover elusive states through which the biomolecule folds (see Figures 7 and 8 for schematic illustrations). We started from a fully unfolded riboswitch, prepared at a high stretching force,  $f_s$ , and quenched the force to  $f_q$ , smaller than the critical unfolding force,  $f_c$ , for a time period  $t_q$  to initiate refolding. We then increased the force to  $f_s$  ( $>f_c \approx 12$  pN for SAM-III riboswitch) rapidly, until the riboswitch was completely unfolded, and then quenched the force again. By repeating this process for varying time period,  $t_q$ , folding can be interrupted, and the states that



**Figure 7.** (A) A sample trajectory of the end-to-end distance as a function of time for the riboswitch in the presence of SAM during the stretch–quench ( $f_s\text{--}f_q$ ) cycles with  $t_q = 0.07$  ms,  $f_s = 20$  pN, and  $f_q = 0$  pN. The green line represents the force changes during the  $f_s\text{--}f_q$  cycles. (B) Same as (A) except  $t_q = 0.5$  ms. Comparison of (A) and (B) shows fingerprint of states in (A) but are hidden in (B). (C) Distributions,  $P(\tau_U)$  of unfolding times for the RNA riboswitch with SAM bound obtained during the  $f_s\text{--}f_q$  cycles (see (A) and (B)) with relaxation time period,  $t_q = 0.01, 0.03, 0.07, 0.2,$  and  $0.5$  ms. During  $t_q$  the force  $f = f_q = 0$  pN. For  $t_q$  in red  $P(\tau_U)$ s are invariant, which means that these  $t_q$  values are large enough for the RNA to fold.

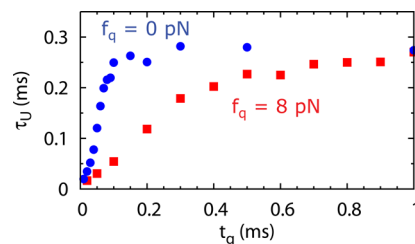


**Figure 8.** (A) A sample trajectory of the end-to-end distance as a function of time for the riboswitch in the presence of SAM during the stretch–quench ( $f_s$ – $f_q$ ) cycles with  $t_q = 0.3$  ms,  $f_s = 20$  pN, and  $f_q = 8$  pN. The green line represents the force changes during the  $f_s$ – $f_q$  cycles. (B) Same as (A) except  $t_q = 1$  ms. (C) Distributions,  $P(\tau_U)$  of unfolding times for the RNA riboswitch with SAM bound obtained during the  $f_s$ – $f_q$  cycles (see (A) and (B)) with relaxation time period,  $t_q = 0.1, 0.3, 0.5, 0.7,$  and  $1$  ms. During  $t_q$  the force  $f = f_q = 8$  pN. For  $t_q$  in red  $P(\tau_U)$ s are invariant, which means that these  $t_q$  values are large enough for the RNA to fold.

are difficult to capture in a single or stepwise quenching of force can be resolved.

We executed the protocol of varying forces by  $f_s$ – $f_q$  cycles in simulations in the presence of metabolite SAM and calculated the distributions of unfolding times for various  $t_q$  values. If  $t_q$  is much smaller than the folding time  $\tau_F$ , the riboswitch is unlikely to reach the folded state during the relaxation time period before it is stretched again. If  $t_q$  is much larger than  $\tau_F$ , the riboswitch will fold with substantial probability during the relaxation time period. In the limit  $t_q/\tau_F \gg 1$ , the distributions of unfolding times during the stretch–quench cycles reduce to the distributions of unfolding times probed in conventional experimental set up in which  $f_q \rightarrow f_s$  reduction is achieved in a single step. When  $t_q \sim \tau_F$ , the riboswitch would sometimes reach the fully folded state during the relaxation time and sometimes would not fully fold before the stretching starts (Figure 7A). Figure 7B shows that the riboswitch reaches the folded state if  $t_q \gg \tau_F$ . Thus, by varying  $t_q$  from  $t_q \ll \tau_F$  to  $t_q \gg \tau_F$ , we expect changes in the distribution,  $P(\tau_U)$ , of unfolding times when  $t_q < \tau_F$ , and a convergence in  $P(\tau_U)$  when  $t_q > \tau_F$ , when most of the riboswitches reach the folded state. More importantly, variations in  $t_q$  allow us to probe certain states of the riboswitch samples, which may be hard to detect in the normal use of SMFS.

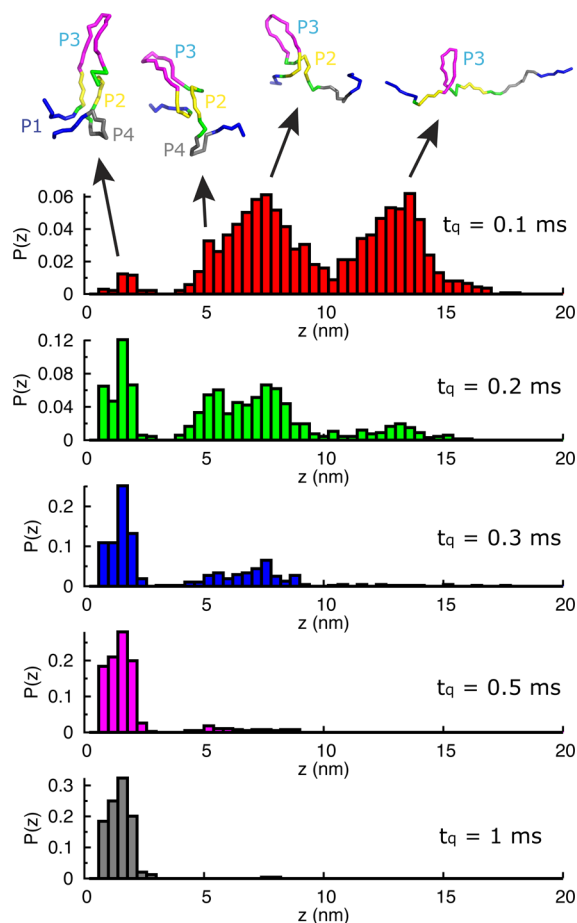
The distributions of the unfolding times, at various  $t_q$  values, in the presence of the metabolite SAM, during the stretch–quench cycles with  $f_s = 20$  pN and  $f_q = 0$  pN, are shown in Figure 7C. For relaxation time periods  $t_q > 0.07$  ms, the distributions of unfolding time become similar, which means the riboswitch is likely to reach the fully folded state during the relaxation time period. The minimum value of  $t_q$  above which  $P(\tau_U)$  becomes similar is about the same value of the folding time at  $f = f_q$ . Indeed, with  $t_q = 0.5$  ms, during which the riboswitch always reaches the folded state, the average folding time  $\tau_F = 0.063$  ms, which is consistent with the minimum  $t_q$  above which  $P(\tau_U)$  does not change significantly. The dependence of average  $\tau_U$  on  $t_q$  (Figure 9) shows that  $\tau_U$  increases sharply with  $t_q$  when  $t_q < \tau_F$ . The increase slows down when  $t_q > \tau_F$ , and the value of  $\tau_U$  reaches convergence when  $t_q > 2\tau_F$ .



**Figure 9.** Average unfolding time,  $\tau_U$ , as a function of  $t_q$  for the SAM-III riboswitch with SAM bound obtained during the  $f_s$ – $f_q$  cycles, with  $f_s = 20$  pN.

We summarize the results in Figure 8 for  $f_q \neq 0$ , which is more easily realized in LOT experiments. Figure 8A,B shows sample folding trajectories ( $z(t)$ ) as a function of  $t$ ) during the cycles with  $f_s = 20$  pN and  $f_q = 8$  pN at two  $t_q$  values. Just as in Figure 7A,B, at smaller  $t_q$  states that are not resolved at longer  $t_q$  become visible. The distributions of unfolding times become similar when  $t_q > 0.3$  ms (Figure 8C). The average folding time during the cycles with  $t_q = 1$  ms, during which the riboswitch is likely to reach the folded state, is  $\tau_F = 0.305$  ms. The dependence of  $\tau_U$  on  $t_q$  also shows that the increase of  $\tau_U$  slows down when  $t_q > 0.3$  ms. The value of  $\tau_U$  reaches a plateau when  $t_q > 0.6$  ms ( $\sim 2\tau_F$ ). The simulations using the stretch–quench cycles provide a way to estimate the folding and unfolding times of the riboswitch, over a wider range of forces, which may be difficult to access in the standard force clamp simulations.

**Resolving the Formation of P1 and P4 Helices.** The distributions of the extension,  $z$ , of the SAM-III riboswitch at the end of a quenching period during the  $f_s$ – $f_q$  cycles, with  $f_s = 20$  pN and  $f_q = 8$  pN at various  $t_q$  values are displayed in Figure 10. For large  $t_q$  ( $t_q \gg \tau_F$ ), the distribution is peaked around  $z \sim 2$  nm because the riboswitch reaches the folded state with substantial probability. For small  $t_q$ , however, several peaks appear in  $P(z)$ . The positions of the peaks correspond to the extension of distinct intermediate states. When  $t_q = 0.1$  ms,  $z(t = t_q)$  is most likely to be around 13 and 7.5 nm, corresponding to  $\Delta P1\Delta P2\Delta P4$  (helices P1, P2, and P4 are disrupted) and  $\Delta P1\Delta P4$  states, respectively. Because of the short quenching time, the probability of complete folding ( $z(t_q) \sim 2$  nm) is low.

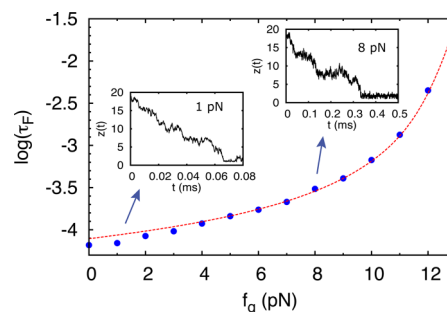


**Figure 10.** The distribution of the extension,  $z$ , of the SAM-III riboswitch at the end of a quenching period ( $t = t_q$ ) during the stretch–quench cycles, with  $f_s = 20$  pN and  $f_q = 8$  pN. When  $t_q = 0.1$  ms, which is much smaller than the folding time,  $\tau_F$ , there are several peaks in  $P(z)$ , which correspond to the states that are sampled. The position of the peaks corresponds to the most probable extensions of the intermediate states. As  $t_q$  increases, the distribution converges to around  $z(t_q) \sim 2$  nm, corresponding to the fully folded state. By varying  $t_q$  elusive states can be captured. For example, with  $\tau_q = 0.1$  ms a state with P2, P3, and P4 intact but P1 disordered is visible ( $z \sim 5$  nm) in the histogram in red.

As  $t_q$  increases to 0.2 ms, the distribution of  $z(t_q)$  around the fully folded state increases. In addition, a peak appears around  $z \sim 5$  nm, which corresponds to the  $\Delta P1$  state (P1 is unfolded) as shown in the top panel in Figure 10. The peak around the  $\Delta P1$  state, where only P1 is unfolded, appears to be transient since the peak does not appear at  $t_q < 0.1$  ms or  $t_q > 0.3$  ms. This is because there are 6 base pairs in P1 and only 2 base pairs in P4. When the force is large enough to unfold P1, it most likely unfolds P4 as well in a very short time. Conversely, when the force decreases enough to allow P4 to fold, P1 also folds rapidly following the formation of P4. It is worth pointing out that only through the use of  $f_s$ – $f_q$  cycles, we can resolve the differences in the folding of P1 and P4. In other ways of quenching or increasing the force (Figures 2–4) the formation or disruption of P1 and P4 occurs nearly simultaneously. By varying  $t_q$ , which is another experimental control, these events can be differentiated.

**Dependence of Folding Rates on  $f$ .** In order to obtain accurate estimates of the folding time,  $\tau_{U \rightarrow F}(f_q)$ , starting from the U state, we used a large stretching force  $f_s$  ( $= 30$  pN) to fully

unfold the riboswitch. Subsequently, we quenched the force to  $f_q$  in a single step and followed the dynamics till the riboswitch reaches the F state for the first time, which is the first passage time. Immediately after reaching the F state, we stretched the RNA to the U state, and maintained  $f = f_s$  to ensure that the riboswitch is fully unfolded. The force was then reduced again to  $f_q$ , and the first passage time was calculated. This process was repeated multiple times to obtain the distribution of first passage times from which the mean  $\tau_{U \rightarrow F}(f_q)$  was calculated. The blue circles in Figure 11 give the



**Figure 11.** The logarithm of the average folding time,  $\tau_F(f_q)$ , as a function of  $f_q$  for the riboswitch in the presence of SAM. The  $f_s$ – $f_q$  cycles are initiated with  $f_s = 30$  pN and  $t_s = 0.1$  ms to unfold the RNA. The dots represent results obtained directly from the time traces of the extension generated during the  $f_s$ – $f_q$  cycles. The line represents  $\tau_F(f_q)$  calculated from eq 1, with  $D_{U \rightarrow F}$  chosen as  $8.46 \times 10^4$  nm<sup>2</sup>/s. The insets show two sample folding trajectories with  $f_q = 1$  and 8 pN.

results obtained using this procedure. The value of  $\tau_{U \rightarrow F}(f_q = 0) = 0.065$  ms coincides with the result in Figure 7A obtained from the minimum  $t_q$  above which  $P(\tau_U)$  is invariant, implying that different protocols used to calculate folding times yield the same results.

In order to further establish the efficacy of eq 1 in estimating  $\tau_{U \rightarrow F}(f_q)$  we used the simulated equilibrium free energy profile at  $f = 12$  pN (Figure 6A), and calculated the profile at other values of  $f$  using

$$G(z, f) = G(z, \bar{f}) - z(f - \bar{f}) \quad (2)$$

We used eq 1 with  $G(z, f)$  (eq 2) to calculate  $\tau_{U \rightarrow F}(f_q)$  with  $D_{U \rightarrow F}$  evaluated at  $f = 9$  pN from explicit simulations. Remarkably, the mean first passage time formula with a single parameter provides an excellent estimate of the folding times over the entire range of  $f_q$ , including at  $f_q = 12$  pN, where the folding time is slow.

## CONCLUDING REMARKS

We have studied the folding and unfolding of the SAM-III riboswitch under tension in order to estimate the forces and time scales associated with unfolding of specific regions that are likely to be involved in the control of translation. The folding landscape, with multiple intermediate states, shows that helix P3 (Figure 1) is the last to unfold upon application of force, requiring an unfolding force  $\sim 2$ – $4$  pN larger than other helices in the riboswitch. This suggests that SAM-III may have evolved to control gene regulation using P3 as a scaffold, which might facilitate ribosome's access to the SD sequence for ribosomal by formation and disruption of helices P1, P4, and P2. A corollary of this inference is that mutations, which destabilize P3, should compromise the dynamic range of SAM-III function.

**Relative Stability of ON and OFF States of SAM-III.** The stable P3 helix might serve as a scaffold for folding, enabling disruption of P1 and P4 helices for binding and unbinding of SAM without globally unfolding the RNA. Binding of SAM enhances the stability of P1 and P4 and substantially increases the free energy barrier for leaving the 'OFF' state. The difference in the free energy between the 'ON' (apo) and 'OFF' (holo) states at  $f = 0$  can be estimated from the free energy profile,  $G(z)$ , at  $f = 9$  pN in Figure 6B. The free energy difference between the 'OFF' and 'ON' states at zero force is approximately given by  $\Delta G \approx \Delta G(z_{Nf}) + \Delta z_N f$ , where  $\Delta z_N$  is the shift in the location of the minima between the ON and OFF states (Figure 6B). With  $f = 9$  pN and  $\Delta z_N \approx 1$  nm, we find that  $\Delta G \approx 9.1$  kcal/mol. The estimate based on experiments is 8.6 kcal/mol,<sup>26</sup> which is surprisingly close to the theoretical prediction, especially considering that the SOP energy function was not optimized to reproduce any measurement.

**Detecting Elusive States Using  $f_s$ - $f_q$  Cycling.** The proposed strategy of using force stretch-quench ( $f_s$ - $f_q$ ) cycles, which (in principle) can be implemented using the LOT setup, provides a unique way of obtaining a fingerprint of states (see Figures 7 and 8) that may otherwise be elusive. In the case of the SAM-III riboswitch, the  $f_s$ - $f_q$  cycling simulations have enabled us to resolve the folding of P4 and P1, which occur nearly simultaneously. In addition, this method also provides a way of measuring the distribution of unfolding and folding times provided repeated  $f_s$ - $f_q$  cycling experiments on the same molecule can be performed. Of course, current experimental time resolution, achievable in LOT experiments,<sup>36,37</sup> can make it difficult to implement the  $f_s$ - $f_q$  cycling. Optimistically, the best time resolution in LOT experiments is  $\sim 100$   $\mu$ s reported in force clamp studies of leucine zipper.<sup>38</sup> The quench time,  $t_q$ , for SAM-III riboswitch is in the range of tenths of milliseconds. It should be noted that  $t_q$  depends on the system. Thus, for RNA molecules with transition times exceeding  $\sim 500$   $\mu$ s, the proposed  $f_s$ - $f_q$  cycling can be implemented. As the LOT time resolution improves we believe that the proposed way of detecting these elusive states may prove useful.

The mean unfolding (folding) time  $\tau_U \approx t_q^e$  ( $\tau_F \approx t_s^e$ ) where  $t_q^e$  ( $t_s^e$ ) is the smallest quench (stretch) time (Figures 7 and 8) beyond which the distribution of unfolding (folding) times are invariant. In force clamp methods, estimates of folding and unfolding times require that a number of transitions between the states of interest be observed. However, at high (low) forces the RNA molecule is essentially pinned in the unfolded (folded) states, which prevents getting accurate values of the folding and unfolding times. In contrast, in the novel  $f_s$ - $f_q$  cycling method the folding times are obtained from the distribution of unfolding times, which can be obtained at arbitrary forces. This can be realized using current experimental setup.

**Force-Dependent Hopping Rates.** Interestingly, the  $f$ -dependent transition rates between two distinct states can be accurately calculated using eq 1 over a limited force range, provided the force-dependent diffusion constant is known. In order to use eq 1, accurate measurement of  $G(z)$  from folding trajectories generated in a typical LOT setup is needed.<sup>39</sup> The collective effects, in eq 1, are in  $D_{A \rightarrow B}$ , which in principle can be computed using simulations. However, the present work shows that, over a limited  $f$  range,  $D_{A \rightarrow B}$  does not depend sensitively on  $f$  nor on the shapes of  $G(z, f)$ . For a quick estimate of  $D_{A \rightarrow B}$ , one can use dimensional arguments to obtain  $D_{A \rightarrow B} \approx a^2/\tau_0$ ,

where  $a \approx 0.7$  nm (approximate distance between phosphates in RNA) and  $\tau_0 \approx 10^{-6}$  s, the value of prefactor in RNA folding.<sup>40</sup> This produces an estimate of  $4.9 \times 10^5$  nm<sup>2</sup>/s, which is within 1 order of magnitude of the values used to generate the results in Figures 6B and 11 (see the captions to these figures). We note in passing that the success of eq 1 in making reliable predictions for transition rates over a small range of forces does not imply that  $z$  is a good reaction coordinate (for additional discussion on this point see Lin et al.).<sup>10</sup> Rather, it only means that accurate measurements of  $G(z)$  can be used to obtain hopping rates provided the many dimensional aspects hidden in  $D_{A \rightarrow B}$  (or  $\tau_0$ ) can be explicitly calculated.

**Implications for Function.** It appears that the kinetic processes in riboswitches that control transcription may be different from those that regulate translation. In the former case, the ability to function as a switch efficiently depends on an interplay of a number of time scales.<sup>15,41,42</sup> Nominally, they are the bimolecular binding rate ( $k_b$ ), the folding times of the aptamer and the time scales to switch and adopt alternate conformations with the downstream expression platform and the rate of transcription. In the 'OFF' switches (transcription is terminated upon metabolite binding), a decision to terminate or complete transcription has to be made before the terminator is synthesized, which puts bounds on the transcription rate,  $k_b$ , and the aptamer folding rates. For typical values of these parameters in both FMN<sup>15</sup> and *pbuE* A-riboswitches,<sup>16</sup> efficient function mandates that the riboswitches be under kinetic control, which implies that the 'OFF' and 'ON' states are not in equilibrium.

The SAM-III function, which controls translation, is different, as first demonstrated by Smith et al.<sup>43</sup> The major time scales that control the function of SAM-III riboswitch, and presumably those that regulate translation in general, are  $k_b$ ,  $k_{dis}$ , and  $k_{mRNA}$ , where  $k_{dis}$  is the rate of disassociation of SAM from the complex with the riboswitch, and  $k_{mRNA}$  is the degradation rate of mRNA. Thus, the only clear bound on the function of SAM-III is that both  $k_{dis}$  and  $k_b[M]$  ( $[M]$  is the concentration of SAM) be much greater than  $k_{mRNA}$ . These conditions imply that the SAM concentration  $[M] \gg k_{mRNA}/k_b$ . Using the values of  $k_b \sim 0.11$   $\mu$ M<sup>-1</sup>s<sup>-1</sup> and  $k_{dis} \approx 0.089$  s<sup>-1</sup>,<sup>43</sup> we obtain that the SAM concentration has to exceed about 50 nM where we have used the estimate for  $k_{mRNA} \approx 3$  min<sup>-1</sup>. It is worth pointing out that our estimates of folding and unfolding times (needed to accommodate SAM binding) based on simulations at low forces are much less than  $k_{mRNA}$ , the inverse of which sets the longest time for translational control. Therefore, the transition between the OFF and ON states can occur multiple times before mRNA is degraded, which gives additional credence to the argument that the function of SAM-III is under thermodynamic control. It is worth noting that these arguments and the ones used in previous studies<sup>43</sup> are tentative because it is assumed that transcription and translation are not coupled. There is evidence that in bacteria transcription and translational speeds are similar,<sup>44</sup> which might complicate the model. However, because  $k_{mRNA}$  is small relative to other rates, it is still possible that riboswitches regulating translation are more likely to be under thermodynamic even if the assumption of decoupling between transcription and translation is relaxed.

## METHODS

**Self-Organized Polymer Model.** We represent the riboswitch using the self-organized polymer (SOP) model that has been



successfully used in a number of studies.<sup>8,30,31,45,46</sup> In the SOP model, each nucleotide is represented as a single interaction site. We model the metabolite SAM as two sites with one representing the adenosyl part (A), and the other corresponding to the methionine part including sulfur (M) (see Figure 1C).

The SOP energy of the RNA in the presence of bound SAM is  $V_T = V_{\text{RIB}} + V_{\text{RIB-SAM}}$ , where  $V_{\text{RIB}} = V_{\text{FENE}} + V_{\text{NB}}$  with

$$V_{\text{FENE}} = - \sum_{j=1}^{N-1} \frac{k}{2} R_0^2 \ln \left( 1 - \frac{(r_{i,i+1} - r_{i,i+1}^0)^2}{R_0^2} \right) \quad (3)$$

$$V_{\text{NB}} = \sum_{i=1}^{N-3} \sum_{j=i+3}^N \varepsilon_A \left[ \left( \frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{ij}^0}{r_{ij}} \right)^6 \right] \Delta_{ij} + \sum_{i=1}^{N-2} \varepsilon_i \left( \frac{\sigma^*}{r_{i,i+2}} \right)^6 + \sum_{i=1}^{N-3} \sum_{j=i+3}^N \varepsilon_j \left( \frac{\sigma}{r_{ij}} \right)^6 (1 - \Delta_{ij}) \quad (4)$$

In eq 3 the chain connectivity term is  $V_{\text{FENE}}$  with  $k = 2000$  kcal/(mol·nm<sup>2</sup>),  $R_0 = 0.2$  nm,  $r_{i,i+1}$  is the distance between two adjacent nucleotides  $i$  and  $i + 1$ , and  $r_{i,i+1}^0$  is the corresponding distance in the native structure. The term,  $V_{\text{NB}}$ , in eq 3 describes the nonbonded interactions, which include the stabilizing forces between the nucleotides that are in contact in the native state. The interactions between the nucleotides that form non-native contacts are taken to be repulsive. Two nucleotides  $i$  and  $j$  are in native contact,  $\Delta_{ij} = 1$ , when the distance between them in the native structure is within a cutoff distance,  $R_c = 1.3$  nm, for  $|i - j| > 2$ . If the distance is larger than  $R_c$ ,  $\Delta_{ij} = 0$ .

We use two values for the parameter  $\varepsilon_h$  depending upon whether the interaction between two nucleotides in native contact are clarified as secondary or a tertiary interaction. If the two nucleotides are within a hairpin or helix, then  $\varepsilon_h = \varepsilon_s$ , otherwise,  $\varepsilon_h = \varepsilon_t$ . The largely hierarchical nature of RNA folding [16] suggests that the strength of the secondary interaction is greater than the tertiary interaction. In our simulations, we set  $\varepsilon_s = 0.7$  kcal/mol, and  $\varepsilon_t/\varepsilon_s = 1/2$ , which capture the balance between secondary and tertiary interactions.<sup>8</sup> The strength of the repulsive interaction for non-native contacts is parametrized by  $\varepsilon_p$ , which is set to 1.4 kcal/mol. We choose  $\sigma = 0.7$  nm, and  $\sigma^* = 0.35$  nm (eq 4) for  $i, i + 2$  pairs to prevent the flattening of the helical structure when the overall repulsion is large.

The interaction between SAM and the riboswitch,  $V_{\text{RIB-SAM}}$  is taken to be

$$V_{\text{RIB-SAM}} = \sum_{i=1}^N \sum_{j=A,S} \varepsilon_{\text{SAM}} \left[ \left( \frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{ij}^0}{r_{ij}} \right)^6 \right] \Delta_{ij} + \sum_{i=1}^N \sum_{j=A,S} \varepsilon_i \left( \frac{\sigma}{r_{ij}} \right)^6 (1 - \Delta_{ij}) \quad (5)$$

The parameter  $\varepsilon_{\text{SAM}}$  is set to  $\varepsilon_s$  for the interaction between SAM and the nucleotides that are in contact with SAM. In the native structure, there are 23 nucleotides that are in contact with either the adenosyl or the methionine part of SAM. To prevent SAM from drifting away from the riboswitch during the simulation, a bond is added between the site representing the adenosyl part of SAM and G7 of the riboswitch (Figure 1).

**Simulations.** The dynamics of the system is described using the Langevin equation in the overdamped limit or Brownian dynamics. The equation of motion for a nucleotide  $i$  is given by

$$\gamma m_i \frac{dx_i}{dt} = - \frac{\partial V_i}{\partial r_i} + F_i(t) \quad (6)$$

where  $\gamma$  is the friction coefficient, and  $m_i$  is the mass of nucleotide  $i$ .  $F_i(t)$ , the random force, satisfies  $\langle F_i(t) \rangle = 0$  and  $\langle F_i(t) F_i(t') \rangle = 2k_B T \gamma m_i \delta(t - t')$ , where the averages are over an ensemble of realizations or trajectories.

The integration step in the simulation is  $\Delta \tau_H = (\gamma \varepsilon_h) / (k_B T) h \tau_L$ , where the typical value for  $\tau_L$  for nucleotides is 4 ps, and the integration step size is  $h = 0.05 \tau_L$ . For the overdamped limit, we use  $\gamma = 100 \tau_L^{-1}$ , which approximately corresponds to the friction coefficient for a nucleotide in water.<sup>47</sup> For  $\varepsilon_h = 0.7$  kcal/mol, this results in an integration time step of about 23 ps.

In the force clamp (or force jump) simulations, an external force is applied directly to the 5'-end of the riboswitch, while the 3'-end is fixed. In the force ramp simulations, the 5'-end is attached to a spring with spring constant  $k_s = 0.15$  pN/nm, while the spring is pulled with a constant speed  $v_s = 6.4 \times 10^2$  nm/s, which gives a loading rate of  $r_f = k_s v_s = 96$  pN/s. Such a value of  $r_f$  can be realized in LOT experiments but is typically about a factor of (10–20) larger than usually employed in such experiments.

## AUTHOR INFORMATION

### Corresponding Author

thirum@umd.edu

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Michael Hinczewski and Pavel Zhuravlev for useful discussions and comments on the manuscript. This work was supported by a grant from the National Institutes of Health (GM 089685).

## REFERENCES

- (1) Winkler, W. C.; Breaker, R. R. *Annu. Rev. Microbiol.* **2005**, *59*, 487–517.
- (2) Cheah, M. T.; Wachter, A.; Sudarsan, N.; Breaker, R. R. *Nature* **2007**, *447*, 497–500.
- (3) Breaker, R. R. *Science* **2008**, *319*, 1795–1797.
- (4) Serganov, A.; Nudler, E. *Cell* **2013**, *152*, 17–24.
- (5) Montange, R. K.; Batey, R. T. *Ann. Rev. Biophys.* **2008**, *37*, 117–133.
- (6) Fiegand, L. R.; Garst, A. D.; Batey, R. T.; Nesbitt, D. J. *Biochemistry* **2012**, *51*, 9223–9233.
- (7) Larson, M. H.; Landick, R.; Block, S. M. *Mol. Cell* **2011**, *41*, 249–262.
- (8) Lin, J. C.; Thirumalai, D. *J. Am. Chem. Soc.* **2008**, *130*, 14080–14081.
- (9) Whitford, P. C.; Schug, A.; Saunders, J.; Hennelly, S. P.; Onuchic, J. N.; Sanbonmatsu, K. Y. *Biophys. J.* **2009**, *96*, L7–L9.
- (10) Lin, J. C.; Hyeon, C.; Thirumalai, D. *J. Phys. Chem. Lett.* **2012**, *3*, 3616–3625.
- (11) Allner, O.; Nilsson, L.; Villa, A. *RNA* **2013**, *19*, 916–926.
- (12) Quarta, G.; Sin, K.; Schlick, T. *PLoS Comp. Biol.* **2012**, *8*, e1002368.
- (13) Feng, J.; Walter, N. G.; Brooks, C. L., III. *J. Am. Chem. Soc.* **2011**, *133*, 4196–4199.
- (14) Priyakumar, U. D.; MacKerell, A. D., Jr. *J. Mol. Biol.* **2010**, *396*, 1422–1438.
- (15) Wickiser, J. K.; Winkler, W. C.; Breaker, R. R.; Crothers, D. M. *Mol. Cell* **2005**, *18*, 49.
- (16) Frieda, K. L.; Block, S. M. *Science* **2012**, *338*, 397–400.
- (17) Winkler, W. C.; Nahvi, A.; Sudarsan, N.; Barrick, J. E.; Breaker, R. R. *Nat. Struct. Biol.* **2003**, *10*, 701–707.
- (18) Weinberg, Z.; Rugulski, E. E.; Hammond, M. C.; Barrick, J. E.; Yao, Z.; Ruzzo, W. L.; Breaker, R. R. *RNA* **2008**, *14*, 822–828.
- (19) Corbino, K. A.; Barrick, J. E.; Lim, J.; Weiz, R.; Tucker, B. J.; Puskarz, I.; Mandal, M.; Rudnick, N. D.; Breaker, R. R. *Genome Biol.* **2005**, *6*, R70.
- (20) Poiata, E.; Meyer, M. M.; Ames, T. D.; Breaker, R. R. *RNA* **2009**, *15*, 2046–2056.
- (21) Fuchs, R. T.; Grundy, F. J.; Henkins, T. M. *Nat. Struct. Mol. Biol.* **2006**, *13*, 226–233.

- (22) Weinberg, Z.; Wang, J. X.; Bogue, J.; Yang, J.; Corbino, K.; Moy, R.; Breaker, R. R. *Genome Biol.* **2010**, *11*, R31.
- (23) Wang, J. X.; Lee, E. R.; Morales, D. R.; Lim, J.; Breaker, R. R. *Mol. Cell* **2008**, *28*, 691–702.
- (24) Gilbert, S. D.; Rambo, R. P.; Tyne, D. V.; Batey, R. T. *Nat. Struct. Mol. Biol.* **2008**, *15*, 177–182.
- (25) Lu, C.; Smith, A. M.; Fuchs, R. T.; Ding, R.; Rajashankar, K.; Henkins, T. M.; Ke, A. *Nat. Struct. Mol. Biol.* **2008**, *15*, 1076–1083.
- (26) Lu, C.; Smith, A. M.; Ding, F.; Chowdhury, A.; Henkin, T. M.; Ke, A. *J. Mol. Biol.* **2011**, *409*, 786–799.
- (27) Anthony, P. C.; Perez, C. F.; Garcia-Garcia, C.; Block, S. M. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 1485–1489.
- (28) Greenleaf, W. J.; Frieda, K. L.; Foster, D. A. N.; Woodside, M. T.; Block, S. M. *Science* **2008**, *319*, 630–633.
- (29) Neupane, K.; Yu, H.; Foster, D. A. N.; Wang, F.; Woodside, M. T. *Nuc. Acids Res.* **2011**, *39*, 7677–7687.
- (30) Hyeon, C.; Thirumalai, D. *Biophys. J.* **2007**, *92*, 731–743.
- (31) Hyeon, C.; Dima, R. I.; Thirumalai, D. *Structure* **2006**, *14*, 1633–1645.
- (32) Li, P. T. X.; Collin, D.; Smith, S. B.; Bustamante, C.; Tinoco, I., Jr. *Biophys. J.* **2006**, *90*, 250–260.
- (33) Garcia-Manyes, S.; Dougan, L.; Badilla, C. L.; Brujic, J.; Fernández, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10534–10539.
- (34) Hyeon, C.; Morrison, G.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9604–9609.
- (35) Barsegov, V.; Thirumalai, D. *Phys. Rev. Lett.* **2005**, *95*, 168302.
- (36) Manosas, M.; Wen, J.-D.; Li, P. T. X.; Smith, S. B.; Bustamante, C.; I. Tinoco, J.; Ritort, F. *Biophys. J.* **2007**, *92*, 3010–3021.
- (37) Wen, J.-D.; Manosas, M.; Li, P. T. X.; Smith, S. B.; Bustamante, C.; I. Tinoco, J. *Biophys. J.* **2007**, *92*, 2996–3009.
- (38) Gebhardt, J. C. M.; Bornschlöggl, T.; Rief, M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 2013–2018.
- (39) Hinczewski, M.; Gebhardt, J. C. M.; Rief, M.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 4500–4505.
- (40) Hyeon, C.; Thirumalai, D. *Biophys. J.* **2012**, *102*, L11–L13.
- (41) Lin, J. C.; Thirumalai, D. *Biophys. J.* **2012**, *103*, 2320.
- (42) Beisel, C. L.; Smolke, C. D. *PLoS Comp. Biol.* **2009**, *5*, e1000363.
- (43) Smith, A. M.; Fuchs, R. T.; Grundy, F. J.; Henkin, T. M. *Mol. Microbiol.* **2010**, *78*, 1393–1402.
- (44) Proshkin, S.; Rahmourni, A. R.; Mironov, A.; Nudler, E. *Science* **2010**, *328*, 504–508.
- (45) Mickler, M.; Dima, R. I.; Dietz, H.; Hyeon, C.; Thirumalai, D.; Rief, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20268–20273.
- (46) Hyeon, C.; Lorimer, G. H.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18939–18944.
- (47) Veitshans, T.; Klimov, D.; Thirumalai, D. *Folding De.* **1997**, *2*, 1–22.