

An anionic ligand snap-locks a long-range interaction in a magnesium-folded riboswitch

Rajeev Yadav ^{1,2}, Julia R. Widom ^{1,3}, Adrien Chauvier ¹ & Nils G. Walter ¹✉

The archetypical transcriptional *crcB* fluoride riboswitch from *Bacillus cereus* is an intricately structured non-coding RNA element enhancing gene expression in response to toxic levels of fluoride. Here, we used single molecule FRET to uncover three dynamically interconverting conformations appearing along the transcription process: two distinct undocked states and one pseudoknotted docked state. We find that the fluoride anion specifically snap-locks the magnesium-induced, dynamically docked state. The long-range, nesting, single base pair A40-U48 acts as the main linchpin, rather than the multiple base pairs comprising the pseudoknot. We observe that the proximally paused RNA polymerase further fine-tunes the free energy to promote riboswitch docking. Finally, we show that fluoride binding at short transcript lengths is an early step toward partitioning folding into the docked conformation. These results reveal how the anionic fluoride ion cooperates with the magnesium-associated RNA to govern regulation of downstream genes needed for fluoride detoxification of the cell.

¹Single Molecule Analysis Group, Department of Chemistry and Center for RNA Biomedicine, University of Michigan, Ann Arbor, MI 48109, USA. ²Present address: Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA. ³Present address: Department of Chemistry and Biochemistry, University of Oregon, Eugene, OR 97403, USA. ✉email: nwalter@umich.edu

Riboswitches are structured noncoding RNA elements typically found in the 5' untranslated regions of bacterial messenger RNAs^{1–5}. They regulate gene expression, most often at the level of transcription and translation, through the binding of specific ligands that induce changes in riboswitch structure. A typical riboswitch comprises a highly structured aptamer (receptor) domain and a downstream regulatory domain (expression platform); binding of the ligand to the aptamer domain induces a conformational switch in the expression platform that directs downstream gene expression. So far over 40 classes of riboswitches have been identified, sensing diverse ligands including metabolites, enzyme cofactors, and metal ions^{3,6–9}. The critical role of riboswitches in bacterial gene regulation makes them an attractive target for the development of new antibiotic therapies¹⁰.

Fluoride riboswitches, also known as *crcB* motif RNAs, are a family of transcriptional riboswitches present in many bacteria and archaea¹¹. On recognition of detrimental levels of fluoride (F⁻), they activate the expression of genes that encode fluoride-sensitive transporters and enzymes such as enolase and formate-hydrogenlyase to detoxify the cell^{11,12}. We focus on the archetypical mesophilic *Bacillus cereus* (*B. cereus*) fluoride riboswitch, which contains an intrinsic terminator that becomes disrupted upon F⁻ binding (Fig. 1a). The crystal structure of the ligand-bound thermophilic aptamer of the *Thermotoga petrophila* fluoride riboswitch shows a pseudoknot-structured docked conformation stabilized by long-range interactions, with the F⁻ binding pocket surrounded by three Mg²⁺ ions coordinated with backbone phosphate groups and water molecules¹³. NMR spectroscopy of the *B. cereus* fluoride riboswitch showed that it adopts a similar tertiary structure in solution both in ligand-bound and ligand-free form, where the ligand allosterically induces the formation of a reverse Hoogsteen A40-U48 base pair at the interface of the aptamer and expression platform¹⁴. These structural studies were performed at equilibrium, but RNA must fold during transcription where it might function before it is fully transcribed, with the RNA polymerase (RNAP) potentially pausing to assist^{6,15,16}.

Various methods have been used to understand co-transcriptional folding of RNA including single-molecule fluorescence resonance energy transfer (smFRET);^{17–19} optical-trapping;⁶ and selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-seq)^{20–22}. Applying SHAPE-seq to the *B. cereus* fluoride riboswitch with RNAP stalled at each nucleotide position, in particular, revealed ligand-dependent delays in terminator formation and identified the point at which the outcome of transcription is determined²⁰. Despite these advances, missing is a dynamic conformational analysis of global RNA structure at functionally relevant transcription points, as signified by either RNA folding topology or intrinsic RNAP pausing, as potential key forces in the mechanism of ligand-guided gene regulation²³.

Here, we used smFRET to determine the conformational dynamics of the *B. cereus* fluoride riboswitch in dependence of both Mg²⁺ and F⁻, first in isolation, then in the presence of bacterial RNAP. We find that the riboswitch in the presence of Mg²⁺ folds into a pseudoknotted docked conformation, previously observed by NMR spectroscopy¹⁴, that still transiently visits two other, less intricately folded states; upon addition of F⁻, the docked structure becomes stably snap-locked by engaging the long-range linchpin A40-U48 base pair. The proximity of RNAP in a downstream elongation complex facilitates the folding of the aptamer into this stably docked conformation. Overall, our results reveal the importance of dynamic RNA docking as a path toward conformational capture by an anionic ligand, aided by RNAP.

Results

smFRET analysis reveals three conformational states of the fluoride riboswitch. In the present study, we used smFRET to investigate the fluoride riboswitch in isolation (Fig. 1b) and as a component of transcription elongation complexes (ECs), comparing the equilibrium populations of accessible conformational states and their kinetic parameters. Given that riboswitch properties and interactions within the transcription machinery will evolve as additional RNA sequence is synthesized in 5'-to-3' direction¹⁶, we prepared a series of three riboswitch variants—termed RNA48, RNA58 and RNA64 in reference to their lengths—by splint ligation. The ligated 5' RNA segment contains both donor and acceptor fluorophores, placed to interrogate aptamer pseudoknot formation by smFRET, whereas the 3' segment entails nucleotides that can be surface captured and mimic the 3' most 7-10 nucleotides of the transcript in the RNA/DNA transcription hybrid not available for folding (Fig. 1c). The variant of shortest length (RNA48) represents the RNA produced at a ligand-dependent pause position that we observed using in vitro transcription assays (Supplementary Fig. 1). By comparison, in RNA58 the aptamer is fully synthesized and can fold into the pseudoknotted conformation, however, with its 3' segment still residing in the exit channel of RNAP²⁰. Finally, in RNA64 the aptamer domain has completely emerged from the exit channel and is free to fold. We first characterized this latter RNA as a reference baseline.

To address the effects of both Mg²⁺ and F⁻ ions on the conformation of the complete aptamer domain, we investigated the structural dynamics of RNA64 upon immobilization on polyethylene glycol (PEG)-passivated, streptavidin-coated quartz slides via a biotinylated locked nucleic acid (LNA) capture probe (CP, covering the short anchor sequence downstream shown in orange in Fig. 1b). We used smFRET to monitor changes in the positions of donor and acceptor fluorophores, with the donor placed at position 33 and the acceptor coupled to the 5' end of the riboswitch. Based on hidden Markov modeling (HMM), we found the traces to exhibit three FRET states, of low- (~0.4), mid- (~0.6–0.7) and high-FRET value (~0.8–0.85; Fig. 1d, e, and Supplementary Fig. 2). Upon addition of Mg²⁺ and F⁻, most molecules become stabilized in the mid-FRET state (Fig. 1d, right; Fig. 1e shows a trace where this occurs during observation), with an estimated inter-fluorophore distance of 46–50 Å that is very close to the distance observed in the crystal structure (50 Å; based on PDB ID 4ENC)¹³. These observations are consistent with the notion that the mid-FRET state represents the previously characterized, compact “docked” conformation of the fluoride riboswitch, henceforth referred to as “D” state.

To further characterize the three states we distinguished, we performed a control experiment wherein we blocked the formation of the main pseudoknot (PK1; we note that several different nomenclatures are in use - we here follow reference²⁰) using an LNA oligonucleotide complementary to nucleotides A39-U48 of the riboswitch (Supplementary Fig. 3). Independent of Mg²⁺ and F⁻, we observed only the high-FRET state, supporting the notion that, despite representing a structure with proximal fluorophores, this state lacks the pseudoknot and thus docking; we therefore refer to it as “prefolded” (P). By comparison, the low-FRET state of RNA64 is much less compact (as evidenced by the distal fluorophores) and will be denoted as unfolded (U). In the following, we further test these assignments.

Increasing magnesium induces folding into two distinct D states, first dynamic, then static. To determine the populations in the three conformational states under different conditions, we generated FRET histograms from 100–300 molecules each and

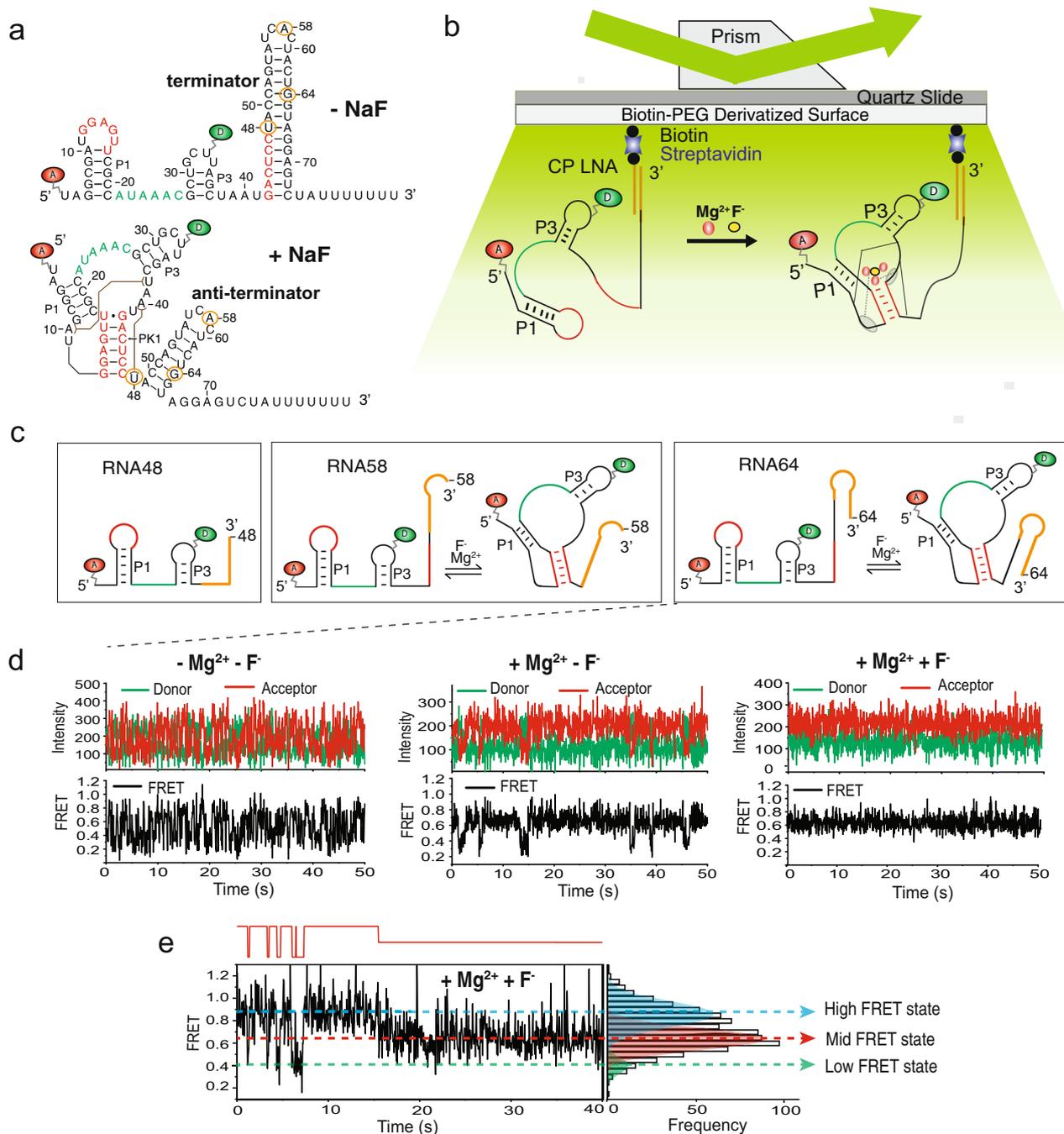
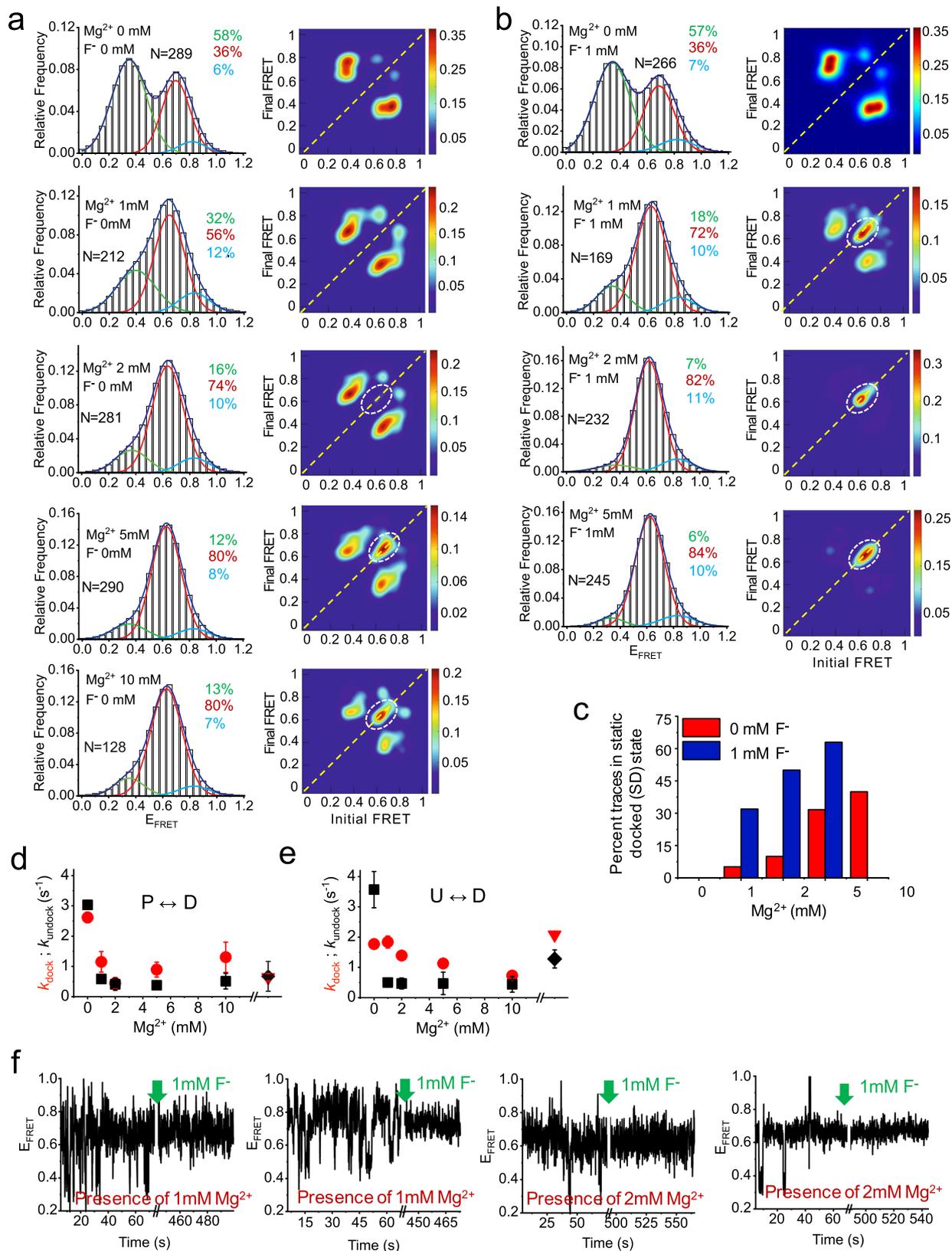


Fig. 1 smFRET analysis of the *Bacillus cereus* *crcB* fluoride riboswitch. **a** Structure of the fluoride riboswitch in the absence and presence of fluoride. The circled nucleotides show the 3' ends of the constructs used for smFRET. The positions of donor and acceptor are shown by green and red ovals labeled D and A, respectively. **b** Experimental smFRET setup. The RNA is immobilized through a biotinylated LNA capture probe. **c** The different lengths of RNA used for smFRET. The orange sequence at the 3' end of the riboswitch is complementary to the CP LNA (or bubble DNA, in experiments performed with RNAP as shown in Fig. 4). The red and green segments are a guide to the eye to compare the riboswitch structure in the absence and presence of F⁻, analogous to the color scheme used in panel a. **d** Representative smFRET traces for RNA64 in the absence of Mg²⁺ and F⁻ (left), in the presence of only Mg²⁺ (middle), and in the presence of both Mg²⁺ and F⁻ (right). **e** Example of an RNA64 smFRET trace that transitions between all three FRET states (observed in 5–10% of traces in the presence of Mg²⁺ only and Mg²⁺ plus F⁻). The HMM fit is shown in red above the trace and the FRET distribution of the trace, fitted with three Gaussian peaks, is shown on the right.

fitted them using a sum of multiple Gaussian peaks (Fig. 2a, b). As expected from our qualitative inspection, the FRET histograms were best fitted with three peaks centered at $E_{\text{FRET}} \approx 0.35, 0.69,$ and 0.83 in the absence of Mg²⁺, and similarly $E_{\text{FRET}} \approx 0.35, 0.63,$ and 0.85 in the presence of 5 mM Mg²⁺. With Mg²⁺, however, the mid-FRET D state becomes more populated, increasing from ~36% in the absence of Mg²⁺ to 56% in 1 mM Mg²⁺, and

eventually saturating at 80% at 5 mM Mg²⁺ (Fig. 2a). This Mg²⁺-dependence is fitted well with a Hill equation to yield a half-saturation point of $K_{1/2} \sim 1$ mM and a cooperativity coefficient of $n = 2.1$ (Supplementary Fig. 4). Taken together, these observations provide further evidence that the mid-FRET is the Mg²⁺-dependent docked D state and requires multiple Mg²⁺ ions to form.



Next, we generated transition occupancy density plots (TODPs), which show as a heat map the fraction of single-molecule traces that exhibit at least once any given initial-to-final FRET transition²⁴. In the presence of nonphysiologically high (10 mM) Mg²⁺, a sub-population of static traces in the mid-FRET (~0.63) state becomes dominant, evidenced on the diagonal

in the TODP (Fig. 2a, b). At a more physiological 1 mM Mg²⁺, only 4% of traces reside in this static D state, rising to 30% and 38% in the presence of 5 mM and 10 mM Mg²⁺, respectively (Fig. 2c). These data indicate that the fluoride riboswitch transiently interconverts between docked and undocked conformations in the absence of Mg²⁺, and increasingly adopts a

Fig. 2 Effect of Mg^{2+} and F^- on the folding of the riboswitch (RNA64). FRET histograms and corresponding TODPs in the absence (a) and presence (b) of F^- . Histograms are well fitted with three Gaussian peaks, shown in green, red and cyan for the low-, mid- and high-FRET states, respectively, with the cumulative fit shown in blue. The percent population of all fitted peaks are shown in respective colors in each histogram panel, and the number of molecules that were analyzed is indicated by “N”. TODPs represent dynamic traces as ‘off-diagonal’ and static traces as ‘on-diagonal’ features, where the color scale shows the prevalence of each population. **c** Percent of all traces that remain statically in the mid-FRET (docked) state as a function of Mg^{2+} concentration in the absence (red) and presence (blue) of 1 mM F^- . **d, e** Kinetics of low- to mid-FRET and high- to mid-FRET transitions as a function of Mg^{2+} concentration. P, U, D stand for prefolded (high-FRET), unfolded (low-FRET) and docked (mid-FRET) states, respectively. The black diamond and red triangle symbols represent rates in 2 mM Mg^{2+} and 1 mM F^- . The error bars are presented as \pm standard deviation for $n = 2$ independent experimental datasets. **f** Representative smFRET traces during real-time ligand jump experiments at 1 mM Mg^{2+} (left two traces) and at 2 mM Mg^{2+} (right two traces) before and after the addition of 1 mM F^- . Upon addition of F^- molecules transition from the DD to the SD conformation. The time of F^- addition is indicated by the green arrow.

static form of the D state as the Mg^{2+} concentration is increased. The observed dwell time in this static mid-FRET state is limited by photobleaching, which ranges from 5 s to 80 s with a mean time of ~ 20 s (Supplementary Fig. 5), representing a lower limit of the static D state lifetime. We henceforth call this the stably docked (SD) conformation of the fluoride riboswitch.

Along with the SD population, a substantial portion of molecules still dynamically transition among FRET states as evidenced by the off-diagonal contours in the TODPs (Fig. 2a). We measured rate constants for docking and undocking from both the prefolded (P, high-FRET) and unfolded (U, low-FRET) to the docked (D, mid-FRET) state (Fig. 2d, e, and Supplementary Table 1). For U to D transitions, the docking and undocking rate constants are $k_{\text{dock}} \approx 1.8 \text{ s}^{-1}$ and $k_{\text{undock}} \approx 3.6 \text{ s}^{-1}$ in the absence of Mg^{2+} , slowing to 1.4 s^{-1} and 0.5 s^{-1} , respectively, in the presence of 2 mM Mg^{2+} . The more significant reduction in the k_{undock} rate constant upon Mg^{2+} addition shows that the metal ion also promotes this dynamic docked (or DD) conformation. At Mg^{2+} concentrations higher than 1 mM these rate constants remain unchanged (Fig. 2e), suggesting that—in contrast to the SD state population—they plateau at physiological Mg^{2+} concentrations of ~ 1 mM. That is, physiological Mg^{2+} concentrations induce the riboswitch into the dynamic docking (DD conformation), whereas higher Mg^{2+} concentrations access a distinct, locked SD conformation.

Fluoride snap-locks the Mg^{2+} -bound riboswitch from the DD into the SD state by conformational selection. We next asked how the conformation of the riboswitch changes in response to ligand binding. We did not observe any change in the FRET histograms or TODPs when 1 mM F^- was added in the absence of Mg^{2+} , indicating that the anionic ligand does not appreciably bind the negatively charged riboswitch (Fig. 2b). In contrast, in the presence of 1 mM and 2 mM Mg^{2+} the addition of 1 mM F^- leads to a drastic (\sim sixfold and \sim fivefold over zero F^- , respectively) increase in the population of the SD conformation (Fig. 2b, c). Our findings indicate that F^- locks the Mg^{2+} -induced DD conformation into the SD conformation more potently than an increase in Mg^{2+} concentration.

This becomes particularly evident in a ligand-jump experiment²⁵ wherein we observe the same set of molecules before and after the addition of F^- . We find that most molecules respond to the F^- jump in both 1 mM and 2 mM Mg^{2+} , with DD traces converting directly into the SD state (Fig. 2f). These observations support the notion that the DD state transiently samples a correctly folded docked conformation, wherein the three crystallographically observed¹³ Mg^{2+} ions create a transient, positively charged pocket in the highly negatively-charged RNA. F^- then binds to this pocket and “snap-locks” the docked conformation into place, resulting in the observed high population of SD molecules. In light of the two alternate mechanisms of molecular recognition known as conformational selection (CS,

“folding first”) or induced-fit (IF, “binding first”)^{26,27}, our data suggest that F^- only binds to the Mg^{2+} -stabilized docked state of the riboswitch and thus follows the CS mechanism.

The linchpin base-pair A40-U48 is required for the SD state.

The highly conserved reverse Hoogsteen A40-U48 base pair (Fig. 1a) was previously observed by NMR and X-ray crystallography to cap the PK1 helix in the presence of F^- and therefore proposed to steer transcription toward anti-termination^{13,14} (note that the latter reference labels it as A37-U45). We therefore asked whether this interaction is important for forming the SD state by disrupting it with a site-specific base mutation, U48A. In the absence of both Mg^{2+} and F^- , the U48A mutant riboswitch traces are highly dynamic. The population of the low-FRET U state dramatically decreased by \sim threefold in the U48A mutant compared to the U48 wild-type, and we did not observe a distinct P state, perhaps because the DD state was shifted slightly higher in E_{FRET} value, resulting in overlap (Fig. 3a). In the presence of 2 mM Mg^{2+} , the U48A mutant resembles the wild-type, with a minor (10%) U state at $E_{\text{FRET}} \approx 0.35$ and a major (90%) DD state at $E_{\text{FRET}} \approx 0.7$ (Fig. 3b). Most strikingly, upon addition of F^- there was no evidence of on-diagonal SD traces in the TODP (Fig. 3), suggesting that the linchpin base-pair A40-U48 is essential for fluoride-mediated stabilization of the SD state.

The observed low population of the U state in the absence of Mg^{2+} and F^- is consistent with the results of our measurements with bases A39 through U48 blocked using a complementary LNA (Supplementary Fig. 3). Like the mutation, this blockage precludes the formation of the A40-U48 base-pair, but it additionally precludes pseudoknot formation and yields the P state almost exclusively. Nearly identical behavior is observed for RNA48, in which the 3' arm of the pseudoknot is sequestered, again precluding docking and formation of the A40-U48 base-pair. Thus, any disruption to the A40-U48 base-pair significantly disfavors the U state, whether or not the disruption also impacts the pseudoknot. If the pseudoknot is additionally disrupted, only the P state is significantly populated (RNA48 and RNA64 with blocking LNA), whereas the mutation alone allows both the D and P states to remain populated. The SD state is accessible only when both the pseudoknot and the A40-U48 base-pair can form, as in wild-type RNA64 (Fig. 2).

RNA polymerase stabilizes riboswitch docking. It is well known that RNA folding is highly sensitive to many cell-specific factors including macromolecular crowding^{28,29}, ionic conditions³⁰ and proximal RNAP²³. In particular, upon incorporation into a transcription complex containing DNA template and RNAP, the folding of the riboswitch could be altered by factors, such as macromolecular crowding, specific RNA-protein interactions and electrostatic interactions with the DNA template. To address this, we recorded smFRET traces with the riboswitch immobilized

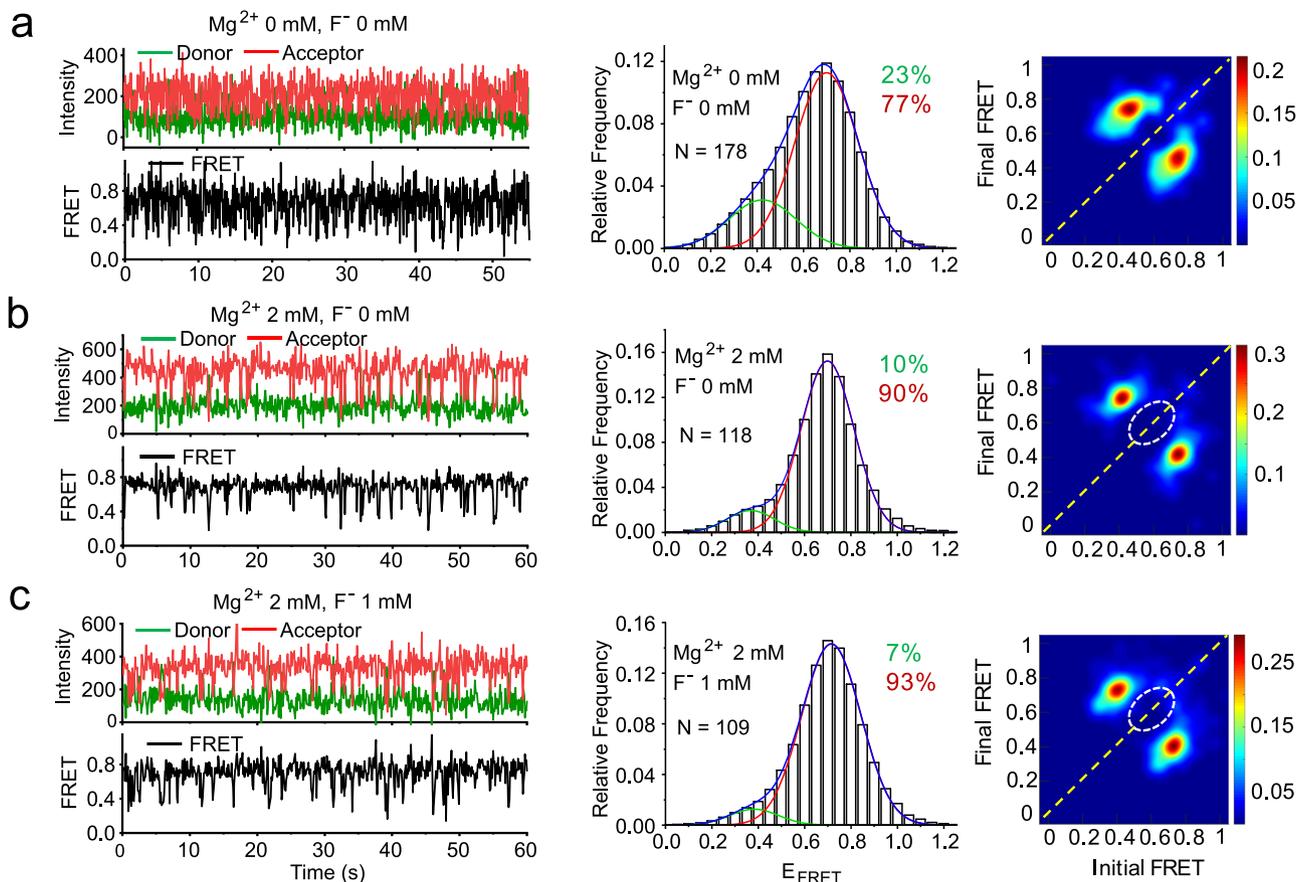


Fig. 3 U48A mutation in RNA64 leads to loss of the SD state. Representative smFRET traces, histograms and TODPs for the U48A mutant (**a**) in the absence of Mg^{2+} and F^- ; (**b**) in the presence of 2 mM Mg^{2+} alone; and (**c**) in the presence of both 2 mM Mg^{2+} and 1 mM F^- . The percent population of all fitted peaks are shown in respective colors in each histogram panel. The dashed white circle in the TODP highlights the absence of the SD state. The number of molecules that were analyzed is indicated by “N”.

through biotinylated RNAP on PEG-passivated, streptavidin-coated quartz slides (Fig. 4a), which enabled observation of the riboswitch in an active, yet halted elongation complex (EC)^{23,31,32}. Since our F^- riboswitch is derived from *Bacillus cereus*, we additionally performed transcription assays using RNA polymerase from the closely related species *Bacillus subtilis*, as well as the more distantly related *E. coli*. Both RNAPs exhibited pausing at U48 and fluoride-induced antitermination, although their behavior differed at locations further downstream than those investigated in this study (Supplementary Fig. 1). As a result, we elected to use the better studied *E. coli* RNAP in our experiments, which also allowed us to build on previous work²³ that performed smFRET on ECs containing *E. coli* RNAP. While the two RNAPs may affect riboswitch behavior in somewhat different ways, we expect that the fundamental mechanism of ligand sensing that we have uncovered will be conserved as are the bacterial RNAPs.

Similar to the isolated RNA64, EC64 also shows three FRET states with $E_{FRET} \approx 0.40, 0.60$ and 0.85 . Both dynamic and static traces were observed as shown by off-diagonal and on-diagonal features in TODPs, respectively (Fig. 4). This suggests that in the presence of RNAP and the DNA template, the riboswitch retains the previously observed prefolded P, docked D and unfolded U states. For low- to mid-FRET transitions the observed rate constants are $k_{dock} \approx 3.5 s^{-1}$ and $k_{undock} \approx 0.5 s^{-1}$ in the presence of 2 mM Mg^{2+} ; note that the docking rate is ~ 3 -fold higher than that observed for isolated RNA64 (Fig. 4e). Similar results were observed for high- to mid-FRET transitions, where the docking rate is increased by ~ 3.5 -fold in EC64 (Supplementary Fig. 6).

Moreover, we found that the SD conformation forms more readily for EC64 than RNA64. In the presence of 2 mM Mg^{2+} , the population of the SD traces is 28% for EC64, which is ~ 3 -fold higher than that observed for RNA64 (Figs. 2a, 4b–d). Upon addition of 1 mM F^- , the prevalence of SD traces further increased to 69%, or ~ 1.3 -fold higher than that observed for RNA64 (Fig. 4d). We could not obtain reliable kinetic parameters for EC64 in the presence of F^- due to the dominance of SD states and thus the low prevalence of dynamic transitions. The increased number of SD traces in EC64 indicates that the transcription machinery thermodynamically stabilizes the docked conformation of the riboswitch (Fig. 4f). Still, the overall folding behavior observed for EC64 is the same as observed for RNA64 in the absence of RNAP, with the DD conformation predominant in the absence of F^- and the SD conformation becoming predominant in the presence of F^- (compare Figs. 2a and 4b–d).

To estimate the favorable effect of the transcription machinery on folding of the F^- riboswitch more quantitatively, we computed the free energy change upon docking and performed single-molecule transition-state analysis (smTSA, see Methods)^{30,33–35}. In the presence of only 2 mM Mg^{2+} , we determined that in EC64 the docked conformation is stabilized by ~ 0.61 kcal/mol compared to RNA64 (Fig. 4f), consistent with previous studies of molecular crowding^{28,29}. Furthermore, smTSA revealed that the difference in the free energy of activation for docking in an EC relative to the isolated RNA, $\Delta\Delta G_{dock}^{++}$, is very similar at 0.58 kcal/mol at 2 mM Mg^{2+} (Fig. 4f), suggesting that the transition state is late with essentially all docking interactions already established, including those involving Mg^{2+} ions. Thus, we propose that the crowding

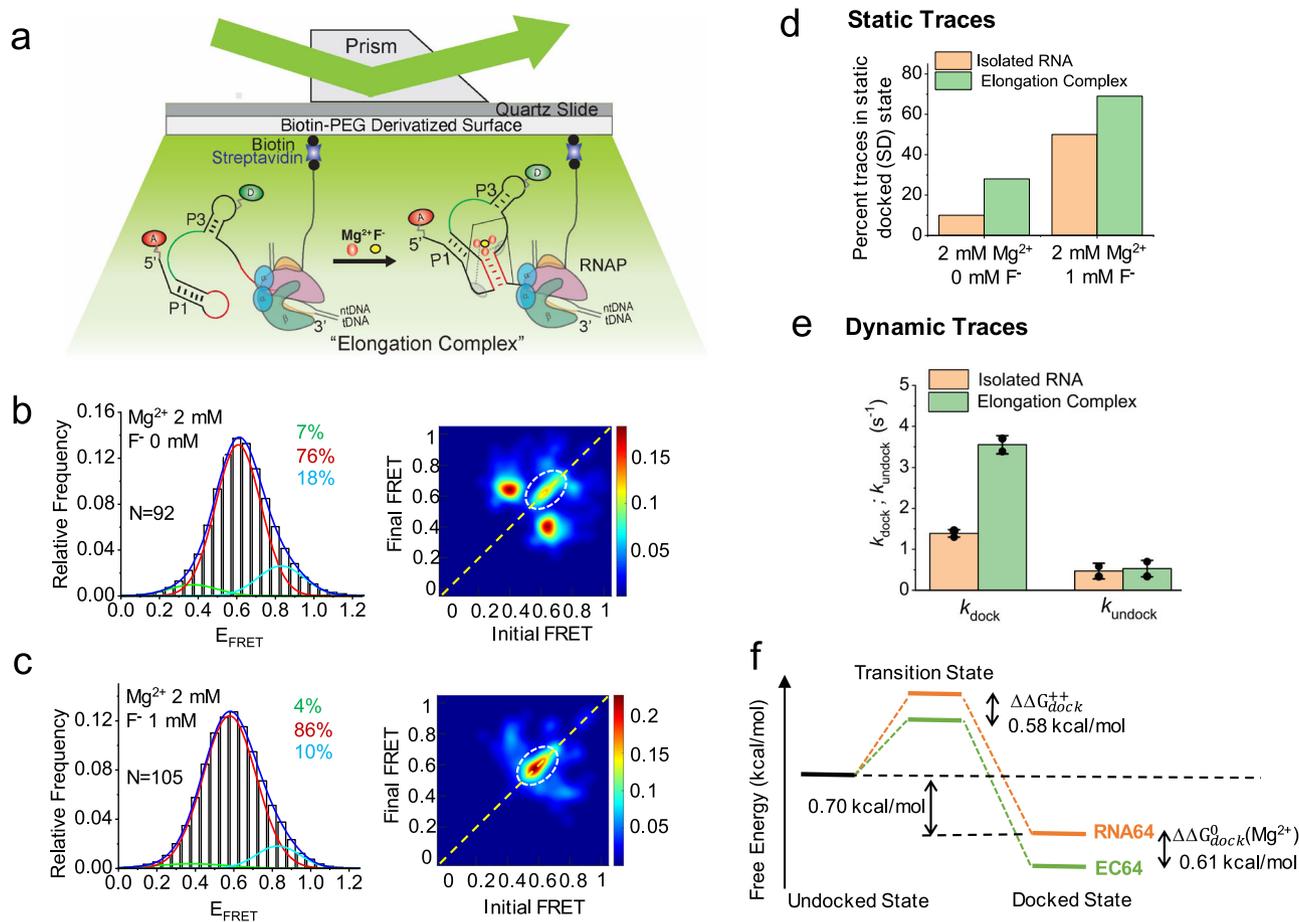


Fig. 4 The presence of RNA polymerase favors the docked conformation of RNA64. **(a)** smFRET experimental setup for probing elongation complexes. RNA is immobilized through biotinylated *E. coli* RNAP. **(b, c)** FRET histograms and corresponding TODPs in the presence of only 2 mM Mg²⁺ **(b)** and in the presence of both 2 mM Mg²⁺ and 1 mM F⁻ **(c)**. Histograms are well fitted with three Gaussian peaks, shown in green, red and cyan for the low-, mid-, and high-FRET states, respectively, with the cumulative fit shown in blue. The percent population of all fitted peaks are shown in respective colors in each FRET histogram panel, and the number of molecules that were analyzed is indicated by "N". A dashed white circle highlights the presence of the SD conformation in both TODPs. **(d)** Percent of all traces that remain statically in the mid-FRET (docked) state for isolated RNA (orange) and ECs (green) both in the presence of 2 mM Mg²⁺ only (left) and in the presence of both 2 mM Mg²⁺ and 1 mM F⁻ (right). **(e)** Kinetic parameters for the unfolded-to-folded docked (low- to mid-FRET) transition in the presence of 2 mM Mg²⁺. The error bars are presented as \pm standard deviation for $n = 2$ independent datasets (with the values as $k_{\text{dock}}: 1.4 \pm 0.1 \text{ s}^{-1}$ and $k_{\text{undock}}: 0.5 \pm 0.2 \text{ s}^{-1}$ for isolated RNA and $k_{\text{dock}}: 3.6 \pm 0.2 \text{ s}^{-1}$ and $k_{\text{undock}}: 0.5 \pm 0.2 \text{ s}^{-1}$ for EC). Reliable kinetic parameters could not be obtained in the presence of F⁻ due to the low prevalence of dynamic traces. **(f)** Free energy diagram for RNA64 (orange) and EC64 (green) in the presence of 2 mM Mg²⁺. The relative free energies of the undocked and docked states were obtained from the dwell times in the docked and unfolded states, while the relative barrier heights were obtained using single-molecule transition state analysis (Methods).

near the surface of RNAP provides additional stabilizing interactions in the TS that accelerate the Mg²⁺-induced folding into the docked riboswitch conformation. (Of note, we did not perform TSA in the presence of F⁻ due to the small number of dynamic traces observed under those conditions).

RNA polymerase plays a key role in the folding of riboswitch at early stages of transcription. We next investigated the effects of the proximal RNAP on the folding of shorter riboswitch RNAs of lengths 58 and 48 (Fig. 5 and Supplementary Figs. 7, 8). For the isolated RNA58 as well as RNAP-containing EC58, the FRET histograms were best fitted with two Gaussian peaks representing high and mid-FRET states (Fig. 5d and Supplementary Figs. 7, 8). We again assigned the mid-FRET state to the docked conformation. For RNA58 almost all the traces were static in nature (Supplementary Fig. 7), while both dynamic and static traces were observed for EC58 (Fig. 5d). Of note, in EC58 we observed a shift from the P (high-FRET) state to the SD (static mid-FRET) state upon addition

of Mg²⁺, with further addition of F⁻ stabilizing the SD state even more. This stands in contrast to EC64, in which Mg²⁺ and F⁻ induced a shift from the U state to the D state (Fig. 4).

We also performed smFRET experiments using an RNA variant, in complexes RNA58n and EC58n, in which the donor labeling position was altered (labeled at A39 instead of U33; Supplementary Note 1 and Supplementary Fig. 9). For this RNA, we disrupted the pseudoknot by two methods—by making a U45A/C46U double mutation, and by blocking the 5' arm of the pseudoknot with an LNA oligonucleotide. In both cases, the population was shifted from the mid-FRET to the high-FRET state, indicating that the mid-FRET state here also represents a docked conformation that is depleted when pseudoknot formation is disfavored (Supplementary Fig. 10). Again, we observed that the F⁻ ligand favors the docked conformation in EC58n (Supplementary Note 1 and Supplementary Figs. 9–14). These results indicate that the transcription machinery facilitates folding of the riboswitch into the docked conformation, and that the P state is prevalent in the absence of Mg²⁺ and F⁻.

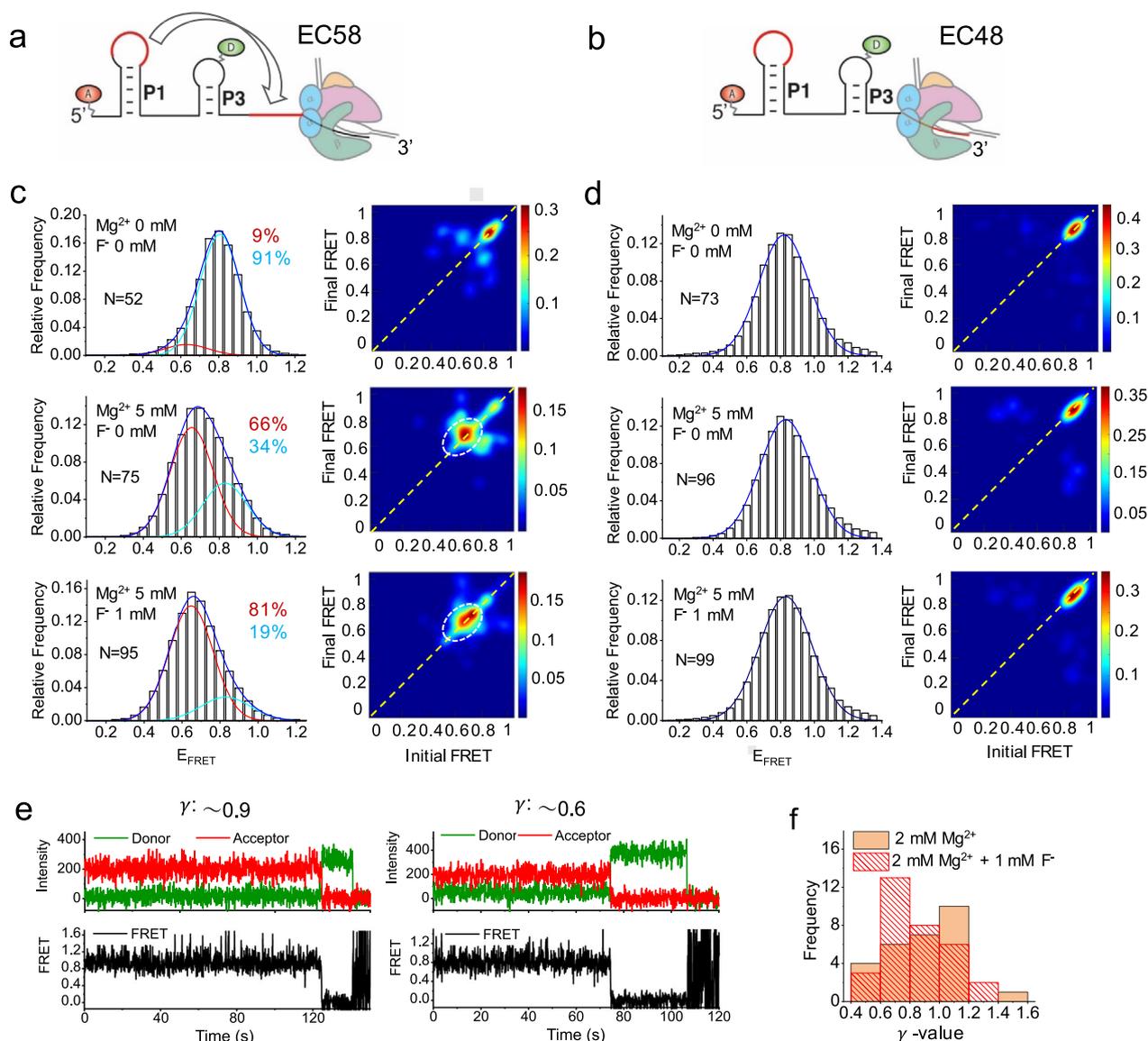


Fig. 5 Effect of Mg^{2+} and F^- on EC58 and EC48. **(a, b)** Schematics of the fluoridic riboswitch in an elongation complex at transcript lengths of 58 and 48 nucleotides, respectively. **(c, d)** FRET histograms and corresponding TODPs in the absence and presence of Mg^{2+} and F^- for EC58 and EC48, respectively. The FRET histograms are well fitted with two Gaussian peaks for EC58 shown in red and cyan for the mid- and high-FRET states, respectively, with the cumulative fit shown in blue; and a single Gaussian peak for EC48 shown in blue. The percent population of fitted peaks are shown in respective colors in each histogram panel, and the number of molecules that were analyzed is indicated by “N”. A dashed white circle highlights the presence of the SD conformation in TODPs for EC58. **(e)** Representative time traces for EC48 showing similar E_{FRET} values but different γ values, which were determined using fluorescence intensities of donor (I_D) and acceptor (I_A) before (pre) and after (post) photobleaching of the acceptor as $\gamma = (I_{A,Pre} - I_{A,Post}) / (I_{D,Post} - I_{D,Pre})$ ^{38,39}. **(f)** Histogram of the γ -parameter for EC48 in the absence and presence of F^- as indicated.

Interestingly, we observed very slow dynamics, on the time-scale of tens of seconds, for both the EC58 and EC58n (Supplementary Figs. 8 and 11). The intrinsic dynamics of pseudoknot formation is fast, as observed for EC64, in which the complete aptamer is accessible for folding. We hypothesized that the slow dynamics observed in EC58 and EC58n are caused by another parameter, possibly interaction of the riboswitch with RNAP. To test this, we performed further investigations including protein-induced fluorescence enhancement (PIFE)^{36,37} and single-molecule fluorescence lifetime measurements for the construct EC58n labeled with a single fluorophore at A39 (Supplementary Note 2 and Supplementary Fig. 14). We observed PIFE in both the absence and the presence of F^- , indicating the existence of fluctuating RNA-RNAP interactions (Supplementary Fig. 14b). Furthermore, fluorescence lifetime measurements

showed that addition of F^- causes the RNA to undergo a conformational change that increases the mobility of the fluorophore (Supplementary Fig. 14c–e). This change may result from the breaking of RNA-RNAP interactions that confine the fluorophore in the absence of F^- .

In case of the shortest design EC48, representing a ligand-dependent transcription pause position (Supplementary Fig. 1), almost all of the population (>98%) remains in the high-FRET state in the absence and presence of Mg^{2+} and F^- (Fig. 5c). This can be rationalized given that EC48 cannot adopt the docked conformation because the 3' segment of the aptamer is still sequestered within the RNA:DNA hybrid. In our in vitro transcription assay, we observed transcriptional pauses at even shorter RNA lengths (U30, U34, and U41; Supplementary Fig. 1) located in the aptamer domain that might extend the time

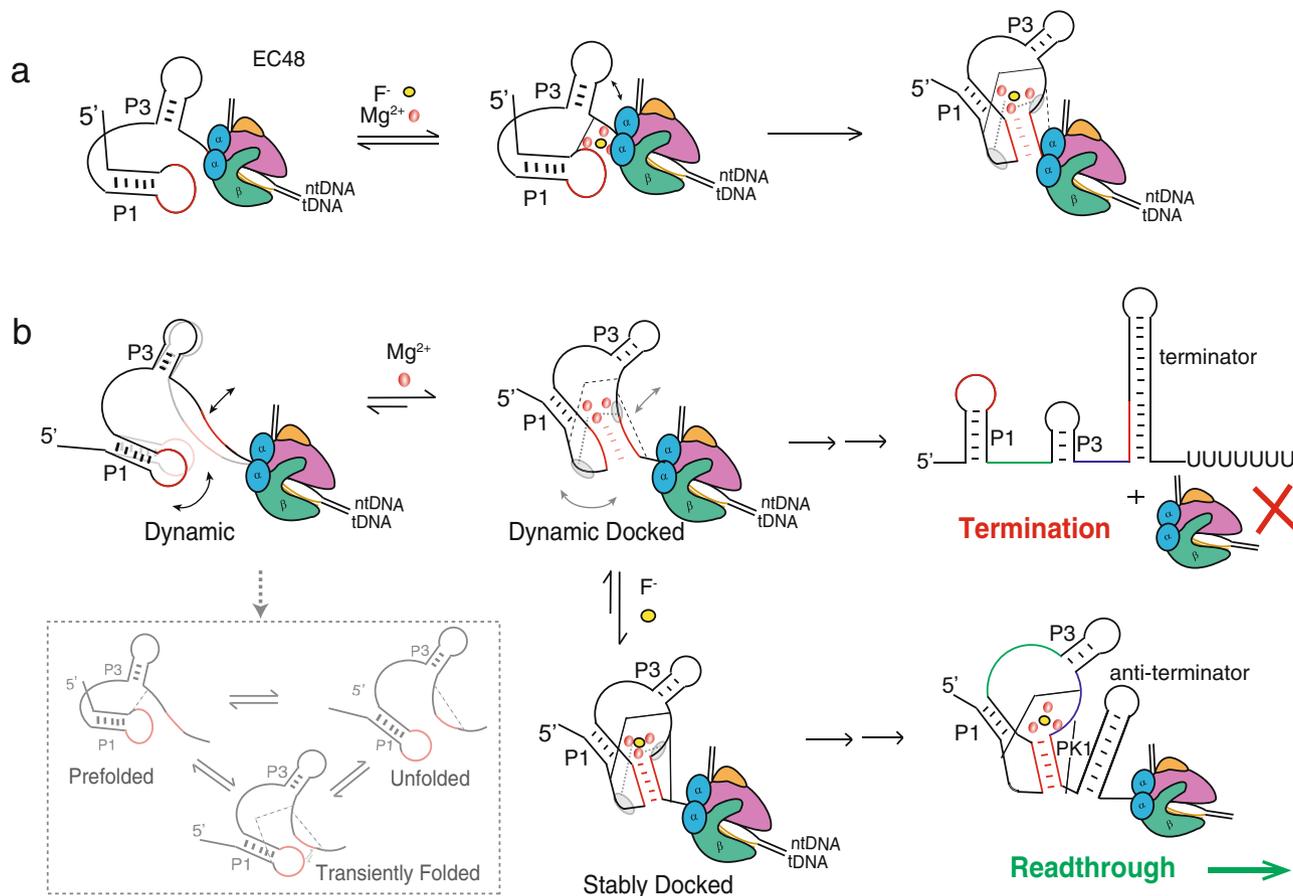


Fig. 6 Models for ligand-dependent transcription regulation by the fluoride riboswitch. **(a)** A proposed local conformational change at early transcript length, where F^- binding pre-organizes the aptamer by poising the 5' segment of the pseudoknot in close proximity to the RNAP exit channel, facilitating pseudoknot formation once the 3' segment is transcribed. **(b)** If F^- binds late, the aptamer is highly dynamic in nature in the absence of Mg^{2+} , transitioning among the prefolded, folded and unfolded states. The unfolded state is prevalent when the aptamer has been transcribed completely and the A40-U48 interaction can occur (RNA64). In the presence of physiological Mg^{2+} concentrations, the riboswitch dynamically samples a docked conformation similar to the ligand-bound conformation (“dynamic docked”). The binding of F^- stabilizes this conformation (to become “stably docked”), mediated by the linchpin long-range A40-U48 interaction.

available for the folding of the P1 hairpin. In agreement with the SHAPE-seq study²⁰, these pauses are independent of the concentration of F^- (Supplementary Fig. 1 and Supplementary Note 3). By contrast, the ligand-dependent transcriptional pause at position U48 is proximal to the 3' segment of the pseudoknot, and may provide time for the Mg^{2+} and F^- ligands to bind. Thus, the observed effect of F^- on the U48 pause half-life must occur through a local F^- -induced change that is not detectable through the smFRET analysis presented so far.

To further investigate the possibility of a local F^- -induced change in EC48, we measured gamma (γ) parameters from donor and acceptor time trajectories (Methods), where the deviation of the γ -value from a value of 1 is evidence of a change in the local environment around the fluorophore that could be due to interactions with protein or nucleic acids (Fig. 5e)^{38,39}. For EC48, a broad distribution of γ -values was observed for individual molecules in the presence of 2 mM Mg^{2+} , ranging from 0.4 to 1.6 with a maximum close to 1; upon addition of F^- , the maximum population shifted to ~ 0.6 (Fig. 5f, time traces with different γ -values are shown in Fig. 5e). This deviation in γ supports the notion of an interaction of the L3 loop (where the donor fluorophore is located) with RNAP. The binding of F^- may bring L3 into closer proximity with the polymerase, facilitating the formation of the pseudoknot as additional RNA is transcribed (Fig. 6a).

Discussion

We have explored in detail the conformational dynamics of the *B. cereus* *crcB* fluoride riboswitch at the single-molecule level. First, we investigated the folding dynamics of the aptamer and how it changes as a function of Mg^{2+} and F^- concentrations. On this foundation, we further explored the involvement of transcription machinery, specifically bacterial RNA polymerase, which induces the riboswitch to fold into the docked conformation, and explored other possible factors involved in the folding process. Using smFRET analysis, we resolved three conformational states: prefolded, unfolded, and docked, and determined their fractional populations and conformational dynamics including inter-conversion rates. We observed that the prefolded and unfolded states are prevalent before and after the aptamer has been fully transcribed, respectively, and each of these states is able to transition to the docked conformation in the presence of F^- ligand. Thus, the prefolded state may play a role as an intermediate on the pathway to the docked conformation at the early stages of transcription (Fig. 6). We also observed ligand-dependent switching behavior even at short transcript lengths, finding evidence that a local conformational change induced by F^- suppresses transcriptional pausing at U48.

The *B. cereus* fluoride riboswitch has previously been characterized by chemical-exchange saturation transfer (CEST) NMR spectroscopy¹⁴. In that study, highly localized fleeting dynamics

were observed at the interface of the aptamer domain and the expression platform. The ligand-free aptamer was found to transiently access an excited conformational state and ligand binding allosterically suppressed that state, while the riboswitch adopted a tertiary structure in the presence of only Mg^{2+} similar to that in the presence of both Mg^{2+} and F^{-} ¹⁴. Using smFRET, we were able to identify key differences between the global behavior of the riboswitch in the presence of only Mg^{2+} and in the presence of Mg^{2+} and F^{-} . Specifically, the DD state is prevalent in the presence of only Mg^{2+} , getting “snap-locked” into the SD state upon F^{-} binding. The CEST-NMR study showed that the A40-U48 “linchpin” base-pair is important for aptamer folding and transcription regulation¹⁴. Here we observed that the A40-U48 base-pair plays a significant role in stabilizing the docked state in the presence of F^{-} , with a single mutation breaking this base-pair and completely abolishing the SD state. This stands in contrast to the pseudoknot-structured preQ₁, ZTP and SAM II sensing riboswitches, where the docked conformation is mainly stabilized through the formation of the pseudoknot helix⁵. This distinction may be a unique consequence of the anionic F^{-} ligand binding to a polyanionic RNA.

Transcriptionally acting riboswitches are thought to be kinetically controlled and may be required to sense their ligands at the early stages of transcription in order to initiate the folding process before RNAP passes the intrinsic terminator hairpin⁴⁰. SHAPE-seq indicated that the aptamer folds unstably in the absence of ligand, whereas in the presence of ligand it adopts a long-lived folded conformation²⁰. This is consistent with our observation of a DD conformation in the presence of Mg^{2+} that transitions to an SD conformation upon addition of ligand. Given that each individual transcript makes its own “decision” between termination and read-through, single-molecule methods are particularly well-suited to analyze the key factors contributing to that decision. Considering the average transcription rate of 20–25 nt/sec in vivo⁴¹, there is a brief window of ~1–2 s for ligand binding between the departure of RNAP from the U48 pause and its arrival at the poly-U termination sequence pause. Of note, the prevalence of high stability SD traces (lifetime in the docked state at least 20 s) increases from 10% (RNA64) and 25% (EC64) without ligand to 50 and 70% with ligand, respectively. This indicates that a subpopulation of nascent RNAs will adopt a docked state at an early stage of transcription, and that they will persist in that state until the entire anti-terminator is formed. Additionally, the remaining (DD) population exhibits a dwell time in the docked state (τ_{docked}) of ~1–2 s, meaning that termination or readthrough are both possible outcomes for the subpopulation of nascent RNAs that adopt this conformation. Thus, at least 70% of transcripts (the entire SD subpopulation + a fraction of the DD subpopulation) are expected to undergo read-through rather than termination in the presence of F^{-} . This is consistent with previously reported in vitro transcription experiments that showed ~70–100% read-through depending on the rate of transcription¹⁴. Thus, the early binding of ligand during transcription would have a high impact on the downstream gene expression. Indeed, we observed that the γ -value of EC48 and the half-life of the pause at U48 both depend on F^{-} (Fig. 5 and Supplementary Fig. 1), suggesting that F^{-} may bind at an early stage of transcription, in contrast to the findings of the prior SHAPE-seq study²⁰.

It is well known that Mg^{2+} promotes many riboswitches to fold into partially compact or docked-like conformations that become the basis for CS-mediated ligand binding^{25,27,30}. A small change in the Mg^{2+} concentration may alter the nature of ligand-binding mechanism. For example, the preQ₁ riboswitch from *Thermotoga maritima* follows an IF mechanism in the absence of Mg^{2+} , while the addition of only 10 μ M Mg^{2+} shifts to the

mechanism to CS³⁰. Similarly, small differences in aptamer sequence can significantly impact the ligand-mediated folding mechanism^{27,42}. Here, we propose a mechanism (Fig. 6b) in which ligand binding by the aptamer domain follows a CS process, where physiological Mg^{2+} concentrations lead the riboswitch to dynamically sample a conformation similar to the F^{-} -bound conformation (DD). F^{-} binding then stabilizes this conformation, leading to the observation of a long-lived docked state (SD).

The anionic ligand F^{-} does not directly interact with any specific nucleotide in the riboswitch but instead bridges a cluster of three Mg^{2+} ions in contact with the backbone phosphates near the PK1 helix¹³ (Fig. 6b). In contrast, riboswitches that sense cations, such as Mg^{2+} and Mn^{2+} bind their ligands largely through direct interactions with the phosphate backbone. Despite these differences, the Mn^{2+} riboswitch appears to follow a similar folding pathway as the one we uncovered for the F^{-} riboswitch. In the presence of Mg^{2+} alone, the Mn^{2+} riboswitch transiently samples a docked conformation that converts to a stably docked conformation upon addition of submillimolar Mn^{2+} ²⁵. The similarity extends to the fact that a Mg^{2+} ion binds to the RNA very close to the riboswitch ligand, whereupon either Mn^{2+} or F^{-} binding can serve as a linchpin for the docking interaction.

Our results in elongation complexes with a proximal RNAP demonstrate that the transcription machinery facilitates folding of the riboswitch into the docked conformation (Fig. 4f). This behavior is analogous to that seen in a previous study, where RNAP was found to stabilize the docked pseudoknot conformation of the preQ₁ riboswitch²³, suggesting that stabilization of RNA folding by RNAP may be a more general phenomenon. In addition to molecular crowding, electrostatic interactions^{23,43} between the RNA and protein and slow solvation^{34,44} around proteins and RNAs may help lower the transition state energy as observed here (Fig. 4f).

Taken together, our findings provide an understanding of the ligand-sensing mechanism of the F^{-} riboswitch and the coupling between its conformational dynamics and the transcription machinery that may lay the basis for the development of novel antibiotics. We anticipate that the mechanism uncovered here may underlie the co-transcriptional folding pathways and regulatory functions of other riboswitches and RNAs.

Methods

Protein preparation. *E. coli* core RNAP containing an ‘AviTag’ biotinylation tag (GLNDFEAKQKIEWH) on the C terminus of the β' subunit was expressed from plasmid pIA1202 in BLR(DE3) cells⁴⁵. The cells were grown at 37 °C in LB media supplemented with 100 μ g/mL carbenicillin until OD₆₀₀ reached around 0.52. They were induced with 0.4 mM IPTG and, at the same time, 20 μ M biotin was introduced into the media. The cells were harvested 4 h postinduction. The protein was purified using a Ni-NTA agarose column, heparin column and finally a Mono Q ion-exchange column, followed by overnight dialysis into storage buffer (10 mM Tris-HCl, pH 7.5, 50% glycerol, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT) as described previously²³ and stored at -80 °C.

Preparation of RNA constructs for smFRET. Different lengths of riboswitch were prepared by splinted ligation of two (dye-labeled and nonlabeled) RNA oligonucleotides. Dye-labeled RNA oligonucleotides and nonlabeled RNA oligonucleotides bearing a 5'-phosphate were purchased from Dharmacon (GE Healthcare Life Sciences) and Integrated DNA Technologies, respectively. We used two different dye-labeled RNAs; the first contains a Cy5-9S label at U3 and a 5'-aminohexylacrylamino-uridine (5-LC-N-U) at position U33, while other one contains a Dy547 label at the 2' position of A39 and a 5'N3 at position U3; these constructs were labeled using a Cy3 and Cy5 mono-reactive dye pack (GE Healthcare Life Sciences), respectively. The RNA was isolated by ethanol precipitation.

For ligation, 300 pmol each of labeled RNA oligonucleotide and 5'-phosphorylated RNA oligonucleotide were annealed to the complementary DNA or LNA splint (400 pmol) in a 20 μ L of ligation buffer (150 mM NaCl, 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA) by heating at 90 °C for 2 min and followed by cooling at 37 °C for 10 min and 22 °C for 10 min. This annealing mixture was diluted to 50 μ L in T4 RNA ligase 2 buffer and 40 units of T4 RNA ligase 2 (New England Biolabs) were added. The reaction mixture was incubated at 37 °C for 3 h

and the reaction was stopped by the addition of gel loading buffer (95% formamide, 1 mM EDTA). The RNA was purified by denaturing urea, 8–9% polyacrylamide gel electrophoresis (PAGE). The ligated RNA was electro-eluted from the gel (BioRad Model 422) and isolated by ethanol precipitation. After drying, the RNA pellet was dissolved in water and the concentration was determined using a NanoDrop spectrophotometer. During the labeling and ligation procedures, the samples having dye-labeled construct were protected from ambient light by aluminum foil.

Sample preparation and smFRET experiments. We have performed smFRET experiments in the presence and absence of RNAP, where immobilization of sample is achieved using biotinylated RNAP and biotinylated capture probe (CP)-locked nucleic acids (LNA), respectively. The 3' end of each riboswitch variant consisted of 10 nucleotides complementary to the template strand of the DNA transcription bubble or LNA sequence. For studying elongation complexes, we first prepared an artificial DNA bubble as previously described^{23,46}. In brief, the sequence of nontemplate DNA (ntDNA) and template DNA (tDNA) strands (Table 2 S) were designed with 11 noncomplementary nucleotides in the middle of the bubble. Ten nucleotides of this noncomplementary region on the tDNA strand are complementary to those appended onto the 3' end of RNA. The sequence of the CP was designed to be complementary to the same 10 nucleotides, and LNA bases were included to achieve a melting temperature T_m of more than 70 °C. The sequence of the 10 nucleotides that form the RNA:LNA or RNA:DNA hybrid was not based on the expression platform, and was identical in the RNA constructs of different lengths with the exception of three bases that had to be changed to accommodate the portion of the aptamer oligonucleotide that became part of the hybrid in the shortest design, RNA48.

For experiments in ECs, the RNA of interest was combined at a final concentration of 0.5 μM with tDNA and ntDNA each at 1 μM in imaging buffer (50 mM Tris-HCl, pH 7.5, 100 mM KCl). The mixture was annealed by incubating at 90 °C for 2 min, then 37 °C for 10 min and then RT for 10 min. This annealing mixture was diluted in imaging buffer supplemented with 1 mM MgCl_2 to a final concentration of 50 nM and 0.37–0.42 μM *E. coli* RNAP was added; this mixture was incubated for 15 min at 37 °C and then kept on ice. This reaction mixture was diluted in imaging buffer to a final concentration of 100–200 pM RNA in and flowed into the imaging chamber (which had previously been passivated with a 10:1 ratio of PEG to biotinylated PEG, then incubated for 10–15 min with a 0.2 mg/mL solution of streptavidin). For experiments on isolated RNA, 0.5 μM RNA was combined with 5–7.5 μM of CP-LNA in the imaging buffer. The mixture was annealed by incubating at 90 °C for 2 min, then 37 °C for 10 min and then RT for 10 min. The annealing mixture was diluted in imaging buffer to a concentration of 20–100 pM RNA and flowed onto the imaging chamber. After incubation for 2–10 min, unbound RNA molecules were removed by extensive flushing of the imaging chamber with imaging buffer.

Unless otherwise noted, all experiments were carried out in 50 mM Tris-HCl, pH 7.5, 100 mM KCl, supplemented with varying concentrations of MgCl_2 and fluoride as indicated in the text and figures. An enzymatic oxygen scavenging system (OSS) was also included consisting of 165 U/mL glucose oxidase from *Aspergillus niger*, 2170 U/mL catalase from *Corynebacterium glutamicum*, 44 mM glucose and 5 mM trolox to minimize photo-bleaching and photo-blinking of dye molecules⁴⁷. This oxygen scavenger was chosen based on previously reported measurements of RNAP elongation activity in buffers containing various OSS components²³. The imaging chamber was incubated with each buffer for 10 min prior to data collection. All smFRET movies were recorded at room temperature (22 °C) on a prism-based TIRF microscope using a CCD (I-Pentamax, Princeton Instruments) or sCMOS camera (Hamamatsu ORCA-Flash4.0 V3) with the exposure time of 50 ms, unless otherwise specified. A continuous laser source of 532 nm was used to excite Cy3 or DY547 and the emission from both Cy3/DY547 (donor) and Cy5 (acceptor) were recorded simultaneously on the same camera. The sample was directly excited using a 640 nm laser at the beginning and the end of each movie to verify the presence of Cy5. Labview was used for all the commands and data acquisition.

smFRET data analysis. All of the smFRET data analysis was performed using custom Matlab scripts (The Math Works). For each dataset, donor-acceptor pairs were identified and localized, and the intensity of each fluorophore was tracked throughout the movie. For further analysis, we manually selected the traces that adhere to the following criteria: Single-step photobleaching, a total (donor + acceptor) fluorescence intensity of >250 (arbitrary units), a fluorescence duration of at least 100 frames (prior to photobleaching) and, if transitions were observed, anticorrelation between donor and acceptor signals. These selection criteria ensure that the traces that get selected originate from single donor-acceptor pairs, not from aggregates or background impurities. All selected traces were processed to subtract background and correct for cross-talk and fluorescence bleed-through. The FRET efficiency (E_{FRET}) was calculated by using background-corrected fluorescence intensities of donor (I_D) and acceptor (I_A) and the equation $E_{\text{FRET}} = I_A / (I_A + I_D)$. FRET histograms were created by sorting the normalizing E_{FRET} and taking the same weightage from each trace to avoid the population inequality due to the different time length of traces that occurs due to photo-bleaching. OriginLab was used for plotting and fitting of histograms. The approximate distances (r) between the fluorophores were calculated using the mean

FRET (E) values and an R_0 value of 54 Å for the Cy3–Cy5 the dye pair using the following equation⁴⁸:

$$E = R_0^6 / (R_0^6 + r^6) \quad (1)$$

HMM and kinetic analysis. For kinetic analysis, the QuB software was used as described previously^{24,49}, where the traces were idealized using hidden Markov modeling (HMM) with a three-state model corresponding to two undocked states (low and high FRET) and a docked state (mid-FRET) using the segmental k-means algorithm. The idealized smFRET traces were further analyzed using custom Matlab scripts that identified transitions, created TODPs and exported dwell-times. TODPs are heat maps that show the fraction of traces that exhibit a given type of transition at least once²⁴, with all traces weighted equally. In brief, regardless of the number of transitions in a dynamic trace, only a single transition is counted to construct the TOPD, which avoids the over-representation of molecules exhibiting many fast transitions. A typical TODP includes both dynamic traces (fast transitions along with slow or rare transitions) and static traces. The sub-populations of different FRET transitions from dynamic traces show up as off-diagonal contours in the TODP heat map, whereas on-diagonal contours represent the sub-population of static traces. For the dynamic traces, dwell times of each state were extracted and converted to cumulative dwell-time histograms. Lifetimes in the undocked ($\tau_{\text{prefolded}}$ and τ_{unfolded}) and docked (τ_{docked}) states were obtained by fitting with multi-exponential functions in OriginPro 8.5. Rate constants of docking and undocking were then calculated as $k_{\text{dock}} = 1/\tau_{\text{undocked}}$ or $1/\tau_{\text{prefolded}}$ and $k_{\text{undock}} = 1/\tau_{\text{docked}}$.

Since we observed two distinct population of traces, static docked as well as dynamic docked, the equilibrium constant for docking (K_{dock}) was estimated from both types of traces, which was further used to calculate an overall docking Gibbs free energy as follows:

$$K_{\text{dock}} = \frac{k_{\text{dock}}}{ak_{\text{undock}} + (1-a)k_{\text{photobleach}}} \quad (2)$$

$$\Delta G_{\text{dock}}^{0,\text{RNA64}} = -RT \ln K_{\text{dock}}(\text{RNA64}) \quad (3)$$

$$\Delta G_{\text{dock}}^{0,\text{EC64}} = -RT \ln K_{\text{dock}}(\text{EC64}) \quad (4)$$

where a and $(1-a)$ are the fractional populations of dynamic and static traces, respectively. Furthermore, the free energies of docking (ΔG_{dock}^0) observed for RNA64 and EC64 were used to compute the change in the equilibrium free energy $\Delta \Delta G_{\text{dock}}^0$ of folding:

$$\Delta \Delta G_{\text{dock}}^0 = \Delta G_{\text{dock}}^{0,\text{RNA64}} - \Delta G_{\text{dock}}^{0,\text{EC64}} \quad (5)$$

Note: Here we assumed a lower limit of K_{dock} , which is limited by the photobleaching rate of static docked traces. However, in case of dynamics traces, the cumulative dwell-time of docked state (τ_{docked}) was best fitted with three-exponential function having a slowest time constant of around 20 s (with 2–3% population), which is very close to the average time constant (19–24 s) for the static docked traces. Thus, we assume the observed K_{dock} would be very close to the actual value. Furthermore, we compute the free energy change in presence of only 2 mM Mg^{2+} (for both RNA64 and EC64), and not in presence of F^- due to the small number of dynamic traces and more number of static docked traces, which may contribute to high error.

Single-molecule transition-state analysis (TSA). For RNA64 and EC64, single-molecule TSA was carried out as previously described^{30,34}. The rate constants of docking k_{dock} observed for RNA64 and EC64 were used to compute the changes in the free energy barrier for docking ($\Delta \Delta G_{\text{dock}}^{++}$) induced by RNAP, as follows:

$$\Delta \Delta G_{\text{dock}}^{++} = -RT \ln (k_{\text{dock}}^{\text{RNA64}} / k_{\text{dock}}^{\text{EC64}}) \quad (7)$$

Although TSA has been extensively used to study the folding of single-domain proteins, it can also be applied to investigate bimolecular reactions where folding is coupled to ligand binding³⁵. Recently, a similar approach has been used to study the folding of RNA, using “mutation” of a ligand to alter its interactions with the RNA during its folding^{30,33–35}. In the current study, we measured $\Delta \Delta G_{\text{dock}}^{++}$ by treating the isolated riboswitch (RNA64) as the “mutant”, given that RNAP contacts are absent in this form.

Calculation of Gamma values. We measured the Gamma parameter (γ) for traces that show similar E_{FRET} , but have differences between the donor and acceptor time traces in terms of their quantum yield³⁹. γ is determined from the change in the intensities of donor and acceptor upon photobleaching of the acceptor as follows³⁸:

$$\gamma = (I_{A,\text{Pre}} - I_{A,\text{Post}}) / (I_{D,\text{Post}} - I_{D,\text{Pre}}) \quad (8)$$

where $I_{A \text{ or } D}$ is the fluorescence intensity of the donor or acceptor fluorophore, and *Pre* and *Post* subscripts indicate intensities before and after photobleaching of the acceptor, respectively.

Excited-state fluorescence lifetime of DY547 of single riboswitch. Fluorescence lifetimes were recorded using a fluorescence lifetime imaging (FLIM) system (Alpha 5, ISS Inc., Urbana-Champaign, IL). The sample was excited at 532 nm, selected by AOPT from a white light excitation source (Fianium WL-SC-400-8-PP) using a 582/75 nm emission filter. First, a laser scanning confocal image was recorded, then the excitation light was focused on a specific molecule and the fluorescence intensity recorded overtime to generate a time-dependent fluorescence decay. Individual decays were fitted using deconvolution of the instrument response function (IRF), yielding the fluorescence lifetimes of single riboswitches.

Single-round Transcription Assays

DNA templates. A 199-nucleotides DNA template including the fluoride riboswitch from *B. cereus* under the control of the T7A1 promoter was cloned into pUC19 plasmid between EcoRI and BamHI restriction sites. In addition, 25 nucleotides not found in the wild-type sequence were inserted after the promoter sequence in order to generate a 25-nucleotide stretch in which the RNA transcript lacks any uracil residues (EC-25) except for the +2 position dependent of the ApU dinucleotide used to initiate the transcription. Transcription templates for *in vitro* transcription were generated by PCR using the 'T7A1-PCR' forward DNA oligonucleotide and the 'crcB-reverse' reverse DNA oligonucleotide. (Supplementary Table 2). For *B. subtilis* transcription, the T7A1 promoter was replaced by the LambdaPr promoter using forward DNA oligonucleotide 'LambdaPr-CrcB-FWD' (Supplementary Table 2).

In vitro Transcription Assay. Halted complexes (EC-25) were prepared in transcription buffer (20 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM MgCl₂, 14 mM 2-mercaptoethanol, 0.1 mM EDTA) containing 25 μM ATP/CTP mix, 50 nM α³²P-GTP (3000 Ci/mmol), 10 μM ApU dinucleotide primer (Trilink), and 50 nM DNA template. *E. coli* RNAP holoenzyme (New England Biolabs) was added to 100 nM, and the mixture was incubated for 10 min at 37 °C. In the case of *B. subtilis* transcription core RNAP and SigA factor were incubated for 30 min at 30 °C with a 1:4 ratio (RNAP:SigA) to reconstitute the holoenzyme prior to the transcription reaction. The sample was passed through a G50 column to remove any free nucleotides. To complete the transcription reaction all four rNTPs were added concomitantly with heparin (450 μg/mL) to prevent the re-initiation of transcription. Time pausing experiments were performed using 10 μM rNTPs. The mixture was incubated at 37 °C, and reaction aliquots were quenched at the desired times into an equal volume of loading buffer (95% formamide, 1 mM EDTA, 0.1% SDS, 0.2% bromophenol blue, 0.2% xylene cyanol). Sequencing ladders were prepared by combining the halted complex with a chase solution containing 250 μM of each rNTP, in addition to one 3'-OMethyl rNTP (at 5 μM for 3'-OMe GTP and 2 μM for 3'-OMe ATP, UTP and CTP). Reaction aliquots were denatured before loading 5–8 μL of each onto a denaturing 8 M urea, 6% polyacrylamide sequencing gel. The gel was dried and exposed to a phosphor screen (typically overnight), which was then scanned on a Typhoon Phosphor Imager (GE Healthcare Life Sciences).

The half-life of transcriptional pausing was determined by calculating the fraction of each RNA pause species compared to the total amount of RNA for each time point, which was analyzed with pseudo-first-order kinetics to extract the half-life⁵⁰. For each determination we subtracted the background signal. Error bars in transcription quantification represent the standard deviation of the mean from at least two independent replicates.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study are available from the corresponding authors upon reasonable request. The source data underlying Figs. 2c, d, e, and 4d, e, Supplementary Fig. 1c, Supplementary Fig. 4, and Supplementary Fig. 6 are provided in a Source Data file and in Supplementary Table 1. Study data have been deposited in the University of Michigan Deep Blue Data repository (<https://doi.org/10.7302/pa6y-fb55>). Source data are provided with this paper.

Code availability

Custom scripts used to analyze the smFRET have been deposited in the University of Michigan Deep Blue Data repository (<https://doi.org/10.7302/pa6y-fb55>).

Received: 24 April 2021; Accepted: 2 December 2021;

Published online: 11 January 2022

References

- Sherwood, A. V. & Henkin, T. M. Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. *Annu. Rev. Microbiol.* **70**, 361–374 (2016).
- Mandal, M. & Breaker, R. R. Gene regulation by riboswitches. *Nat. Rev. Mol. Cell Biol.* **5**, 451–463 (2004).
- Serganov, A. & Nudler, E. A Decade of Riboswitches. *Cell* **152**, 17–24 (2013).
- Mironov, A. S. et al. Sensing Small Molecules by Nascent RNA: A Mechanism to Control Transcription in Bacteria. *Cell* **111**, 747–756 (2002).
- Jones, C. P. & Ferré-D'Amaré, A. R. Long-Range Interactions in Riboswitch Control of Gene Expression. *Annu. Rev. Biophys.* **46**, 455–481 (2017).
- Frieda, K. L. & Block, S. M. Direct Observation of Cotranscriptional Folding in an Adenine Riboswitch. *Sci. (80-.)* **338**, 397–400 (2012).
- Greenlee, E. B. et al. Challenges of ligand identification for the second wave of orphan riboswitch candidates. *RNA Biol.* **15**, 377–390 (2018).
- Sherlock, M. E. & Breaker, R. R. Former orphan riboswitches reveal unexplored areas of bacterial metabolism, signaling, and gene control processes. *RNA* **26**, 675–693 (2020).
- Tang, D.-J. et al. A SAM-I riboswitch with the ability to sense and respond to uncharged initiator tRNA. *Nat. Commun.* **11**, 2794 (2020).
- Deigan, K. E. & Ferré-D'Amaré, A. R. Riboswitches: Discovery of Drugs That Target Bacterial Gene-Regulatory RNAs. *Acc. Chem. Res.* **44**, 1329–1338 (2011).
- Baker, J. L. et al. Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. *Science*. **335**, 233–235 (2012).
- Li, S. et al. Eukaryotic resistance to fluoride toxicity mediated by a widespread family of fluoride export proteins. *Proc. Natl Acad. Sci. USA.* **110**, 19018–19023 (2013).
- Ren, A., Rajashankar, K. R. & Patel, D. J. Fluoride ion encapsulation by Mg²⁺ ions and phosphates in a fluoride riboswitch. *Nature* **486**, 85–89 (2012).
- Zhao, B., Guffy, S. L., Williams, B. & Zhang, Q. An excited state underlies gene regulation of a transcriptional riboswitch. *Nat. Chem. Biol.* **13**, 968–974 (2017).
- Wickiser, J. K., Winkler, W. C., Breaker, R. R. & Crothers, D. M. The Speed of RNA Transcription and Metabolite Binding Kinetics Operate an FMN Riboswitch. *Mol. Cell* **18**, 49–60 (2005).
- Ray, S., Chauvier, A. & Walter, N. G. Kinetics coming into focus: single-molecule microscopy of riboswitch dynamics. *RNA Biol.* **16**, 1077–1085 (2019).
- Uhm, H., Kang, W., Ha, K. S., Kang, C. & Hohng, S. Single-molecule FRET studies on the cotranscriptional folding of a thiamine pyrophosphate riboswitch. *Proc. Natl Acad. Sci. USA.* **115**, 331–336 (2018).
- Hua, B., Panja, S., Wang, Y., Woodson, S. A. & Ha, T. Mimicking Co-Transcriptional RNA Folding Using a Superhelicase. *J. Am. Chem. Soc.* **140**, 10067–10070 (2018).
- Hua, B. et al. Real-time monitoring of single ZTP riboswitches reveals a complex and kinetically controlled decision landscape. *Nat. Commun.* **11**, 4531 (2020).
- Watters, K. E., Strobel, E. J., Yu, A. M., Lis, J. T. & Lucks, J. B. Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat. Struct. Mol. Biol.* **23**, 1124–1131 (2016).
- Strobel, E. J., Cheng, L., Berman, K. E., Carlson, P. D. & Lucks, J. B. A ligand-gated strand displacement mechanism for ZTP riboswitch transcription control. *Nat. Chem. Biol.* **15**, 1067–1076 (2019).
- Watters, K. E., Yu, A. M., Strobel, E. J., Settle, A. H. & Lucks, J. B. Characterizing RNA structures *in vitro* and *in vivo* with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Methods* **103**, 34–48 (2016).
- Widom, J. R. et al. Ligand Modulates Cross-Coupling between Riboswitch Folding and Transcriptional Pausing. *Mol. Cell* **72**, 541–552 (2018). e6.
- Blanco, M. & Walter, N. G. Analysis of Complex Single-Molecule FRET Time Trajectories. in *Methods in Enzymology* **472**, 153–178 (Academic Press, 2010).
- Suddala, K. C. et al. Local-to-global signal transduction at the core of a Mn²⁺ sensing riboswitch. *Nat. Commun.* **10**, 4304 (2019).
- Weikl, T. R. & Paul, F. Conformational selection in protein binding and function. *Protein Sci.* **23**, 1508–1518 (2014).
- Haller, A., Soulière, M. F. & Micura, R. The Dynamic Nature of RNA as Key to Understanding Riboswitch Mechanisms. *Acc. Chem. Res.* **44**, 1339–1348 (2011).
- Daher, M., Widom, J. R., Tay, W. & Walter, N. G. Soft Interactions with Model Crowders and Non-canonical Interactions with Cellular Proteins Stabilize RNA Folding. *J. Mol. Biol.* **430**, 509–523 (2018).
- Dupuis, N. F., Holmstrom, E. D. & Nesbitt, D. J. Molecular-crowding effects on single-molecule RNA folding/unfolding thermodynamics and kinetics. *Proc. Natl Acad. Sci. USA.* **111**, 8464–8469 (2014).
- Suddala, K. C., Wang, J., Hou, Q. & Walter, N. G. Mg²⁺ Shifts Ligand-Mediated Folding of a Riboswitch from Induced-Fit to Conformational Selection. *J. Am. Chem. Soc.* **137**, 14075–14083 (2015).
- Daube, S. & von Hippel, P. Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes. *Sci. (80-.)* **258**, 1320–1324 (1992).
- Sidorenkov, I., Komissarova, N. & Kashlev, M. Crucial Role of the RNA:DNA Hybrid in the Processivity of Transcription. *Mol. Cell* **2**, 55–64 (1998).

33. Silverman, S. K. & Cech, T. R. An early transition state for folding of the P4-P6 RNA domain. *RNA* **7**, 161–166 (2001).
34. Bokinsky, G. et al. Single-molecule transition-state analysis of RNA folding. *Proc. Natl Acad. Sci.* **100**, 9302–9307 (2003).
35. Fersht, A. R. & Daggett, V. Protein Folding and Unfolding at Atomic Resolution. *Cell* **108**, 573–582 (2002).
36. Hwang, H. & Myong, S. Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions. *Chem. Soc. Rev.* **43**, 1221–1229 (2014).
37. Hwang, H., Kim, H. & Myong, S. Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proc. Natl Acad. Sci. USA.* **108**, 7414–7418 (2011).
38. Ray, S., Bandaria, J. N., Qureshi, M. H., Yildiz, A. & Balci, H. G-quadruplex formation in telomeres enhances POT1/TPP1 protection against RPA binding. *Proc. Natl Acad. Sci. USA.* **111**, 2990–2995 (2014).
39. Dahan, M. et al. Ratiometric measurement and identification of single diffusing molecules. *Chem. Phys.* **247**, 85–106 (1999).
40. Garst, A. D. & Batey, R. T. A switch in time: detailing the life of a riboswitch. *Biochim. Biophys. Acta* **1789**, 584 (2009).
41. Vogel, U. & Jensen, K. F. The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J. Bacteriol.* **176**, 2807–2813 (1994).
42. Suddala, K. C. et al. Single transcriptional and translational preQ1 riboswitches adopt similar pre-folded ensembles that follow distinct folding pathways into the same ligand-bound structure. *Nucleic Acids Res.* **41**, 10462–10475 (2013).
43. Gupta, A. & Gribskov, M. The Role of RNA Sequence and Structure in RNA-Protein Interactions. *J. Mol. Biol.* **409**, 574–587 (2011).
44. Yadav, R., Sengupta, B. & Sen, P. Effect of sucrose on chemically and thermally induced unfolding of domain-I of human serum albumin: Solvation dynamics and fluorescence anisotropy study. *Biophys. Chem.* **211**, 59–69 (2016).
45. Kay, B. K., Thai, S. & Volgina, V. V. High-throughput Biotinylation of Proteins. *Methods Mol. Biol.* **498**, 185–196 (2009).
46. Kolb, K. E., Hein, P. P. & Landick, R. Antisense Oligonucleotide-stimulated Transcriptional Pausing Reveals RNA Exit Channel Specificity of RNA Polymerase and Mechanistic Contributions of NusA and RfaH. *J. Biol. Chem.* **289**, 1151–1163 (2014).
47. Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An Oxygen Scavenging System for Improvement of Dye Stability in Single-Molecule Fluorescence Experiments. *Biophys. J.* **94**, 1826–1835 (2008).
48. Ishii, Y., Yoshida, T., Funatsu, T., Wazawa, T. & Yanagida, T. Fluorescence resonance energy transfer between single fluorophores attached to a coiled-coil protein in aqueous solution. *Chem. Phys.* **247**, 163–173 (1999).
49. Qin, F. & Li, L. Model-Based Fitting of Single-Channel Dwell-Time Distributions. *Biophys. J.* **87**, 1657–1671 (2004).
50. Landick, R., Wang, D. & Chan, C. L. Quantitative analysis of transcriptional pausing by *Escherichia coli* RNA polymerase: his leader pause site as paradigm. *Methods Enzymol.* **274**, 334–353 (1996).

Acknowledgements

We thank Dr. Julius B Lucks for generous gifts of some plasmids containing the *Bacillus cereus* crcB fluoride riboswitch, and Dr. Katsuhiko Murakami for the generous gift of *Bacillus subtilis* RNAP and SigA factor. We thank Bennett Hendricks for help with some experiments. We also acknowledge NSF MRI-ID grant DBI-0959823 (to N.G.W.) for seeding the Single Molecule Analysis in Real-Time (SMART) Center, whose Fluorescence Lifetime Imaging (FLIM) system was used with Dr. Damon Hoff's help for fluorescence lifetime measurements. This work was supported by National Institutes of Health (NIH) grants R01 GM118524 and R35 GM131922 (N.G.W.) and GM120457 (J.R.W.).

Author contributions

R.Y., J.R.W. and N.G.W. designed the experiments; R.Y. performed all the experiments and analyzed the data; A.C. performed transcription assays and analyzed transcription data; R.Y. wrote the original draft; R.Y., J.R.W., A.C. and N.G.W. wrote, reviewed and edited the manuscript. All authors read and commented on the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-27827-y>.

Correspondence and requests for materials should be addressed to Nils G. Walter.

Peer review information *Nature Communications* thanks Evgeny Nudler and the other, anonymous, reviewer for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

Supplementary Information:

An anionic ligand snap-locks a long-range interaction in a magnesium-folded riboswitch

Rajeev Yadav^{1,2}, Julia R Widom^{1,3}, Adrien Chauvier¹ & Nils G Walter^{1,*}

¹Single Molecule Analysis Group, Department of Chemistry and Center for RNA Biomedicine, University of Michigan, Ann Arbor, MI 48109, USA

²Present address: Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA

³Present address: Department of Chemistry and Biochemistry, University of Oregon, Eugene, OR 97403, USA

*e-mail: nwalter@umich.edu

Supplementary Note 1: Experiments with an alternative donor labeling position

To validate our findings, we further performed smFRET experiments at RNA lengths of 48 and 58 with a different donor labeling position (Supplementary Fig. 9a). In this design, the donor (DY547) is placed at A39, rather than U33 as in the previous construct; the acceptor position remains the same at the 5'-end. When referring to this new design, we place “n” after the number designating a particular RNA species (EC58n and RNA58n, for example).

For the isolated RNA58n, we observed only one FRET state and no changes upon addition of Mg^{2+} and F^- (Supplementary Fig. 9b,d). The FRET histograms were best fitted with a single Gaussian peak centered around $E_{FRET} \approx 0.9$ in the absence and presence of Mg^{2+} and F^- (Supplementary Fig. 9d). In contrast, the elongation complex EC58n exhibits an additional mid-FRET state in the absence of Mg^{2+} and F^- , which becomes more prominent in the presence of Mg^{2+} and F^- (Supplementary Fig. 9c,e). Since F^- favors the docked pseudoknot conformation of the riboswitch, the additional peak in EC58n presumably results from pseudoknot formation in the presence of RNAP. To ensure that the mid-FRET state represents the docked conformation in this new design, we performed control experiments in which the pseudoknot was disrupted either by mutation (U45A/C46U)¹ or with the L1 block sequence (Supplementary Fig. 10). Both of these alterations depress the population of the mid-FRET state and abolish its response to Mg^{2+} and F^- , confirming that the mid-FRET state represents the docked conformation.

Having noticed that the dynamics are very slow at this transcript length for constructs with the original labeling scheme, we performed smFRET measurements with an exposure time of 400 ms. For EC58n, the FRET histograms were best fitted with two Gaussian peaks centered around $E_{FRET} \approx 0.55$ and 0.9 and no effect was observed upon addition of F^- in the absence of Mg^{2+} (Supplementary Fig. 11c). In the presence of 5 mM Mg^{2+} , the population of the mid-FRET state

was 20%, which increased to 45% in the presence of 0.5 mM F^- . These results are in agreement with those shown in the main text for the corresponding RNA designs with the original labeling scheme (Fig. 4), indicating that F^- binds only in the presence of Mg^{2+} and induces folding into the docked conformation.

We observed both dynamic and static traces as shown by diagonal and off-diagonal features in TODPs, respectively (Supplementary Fig. 11c,d). In the absence of Mg^{2+} , a static population is observed in the high FRET state, which on addition of F^- remains almost unchanged. However, in the presence of Mg^{2+} F^- addition leads to a decrease in the static high-FRET population and a concomitant increase in the dynamic population, as well as a small static population in the mid-FRET state at high F^- concentration (Supplementary Fig. 11d). We observed very slow dynamics for this RNA design (Supplementary Fig. 11e), where the rate constants are observed as $k_{\text{dock}} \approx 0.06 \text{ s}^{-1}$ and $k_{\text{undock}} \approx 0.062 \text{ s}^{-1}$ in the absence of Mg^{2+} and almost no change was observed on addition of F^- . In the presence of 5 mM Mg^{2+} , the rates are observed as $k_{\text{dock}} \approx 0.03 \text{ s}^{-1}$ $k_{\text{undock}} \approx 0.032 \text{ s}^{-1}$, which on addition of 0.1 mM F^- changes to $k_{\text{dock}} \approx 0.04 \text{ s}^{-1}$ $k_{\text{undock}} \approx 0.025 \text{ s}^{-1}$. All these results support that the docked state is stabilized in the presence of F^- .

Because the intrinsic dynamics of the aptamer are fast (discussed in the main text for RNA64), the comparably slow dynamics of EC58n point towards involvement of other factors. Additionally, the mid-FRET (docked) state was observed in EC58n but not RNA58n (Supplementary Fig. 9). These additional factors could include interactions of the riboswitch with either the whole transcription machinery or with only the DNA template or RNAP. To differentiate between these possibilities, we performed smFRET experiment on RNA58n in the presence of only the DNA template, where the riboswitch was immobilized through a biotinylated non-template DNA (ntDNA) strand. Under these conditions, we observed only the high-FRET state

(Supplementary Fig. 12), similar to what was seen for isolated RNA58n (Supplementary Fig. 9d). This clearly indicates that the docked state observed in EC58n appears due to the presence of RNAP; potential interactions of the riboswitch with RNAP are discussed in the main text and the following section.

We observed mainly the high-FRET state for isolated RNA48n as well as EC48n and did not observe any detectable change upon addition of Mg^{2+} and F^- (Supplementary Fig. 13). Based on these results and those discussed in the main text for RNA48 and EC48, we conclude that this state is prevalent before the riboswitch has been fully transcribed, and we term it the precursor state.

Supplementary Note 2: Interaction of the riboswitch with RNAP at transcript length 58

We first tested the hypothesis that the riboswitch interacts with RNAP by performing protein induced fluorescence enhancement (PIFE)^{2,3} measurements on EC58n, labeled with only a donor fluorophore at A39 (Supplementary Fig. 14a). The fluorescence intensity varied over time in most traces, in both the absence and presence of F^- (Supplementary Fig. 14b), indicating that the local environment of the donor dye is changing (a restricted environment leads to fluorescence enhancement). This could only be possible if there is a structural change in the riboswitch that changes the extent to which the region in the vicinity of the fluorophore (possibly the L3 loop) is interacting with RNAP. As discussed in the main text, this effect was also seen for EC48.

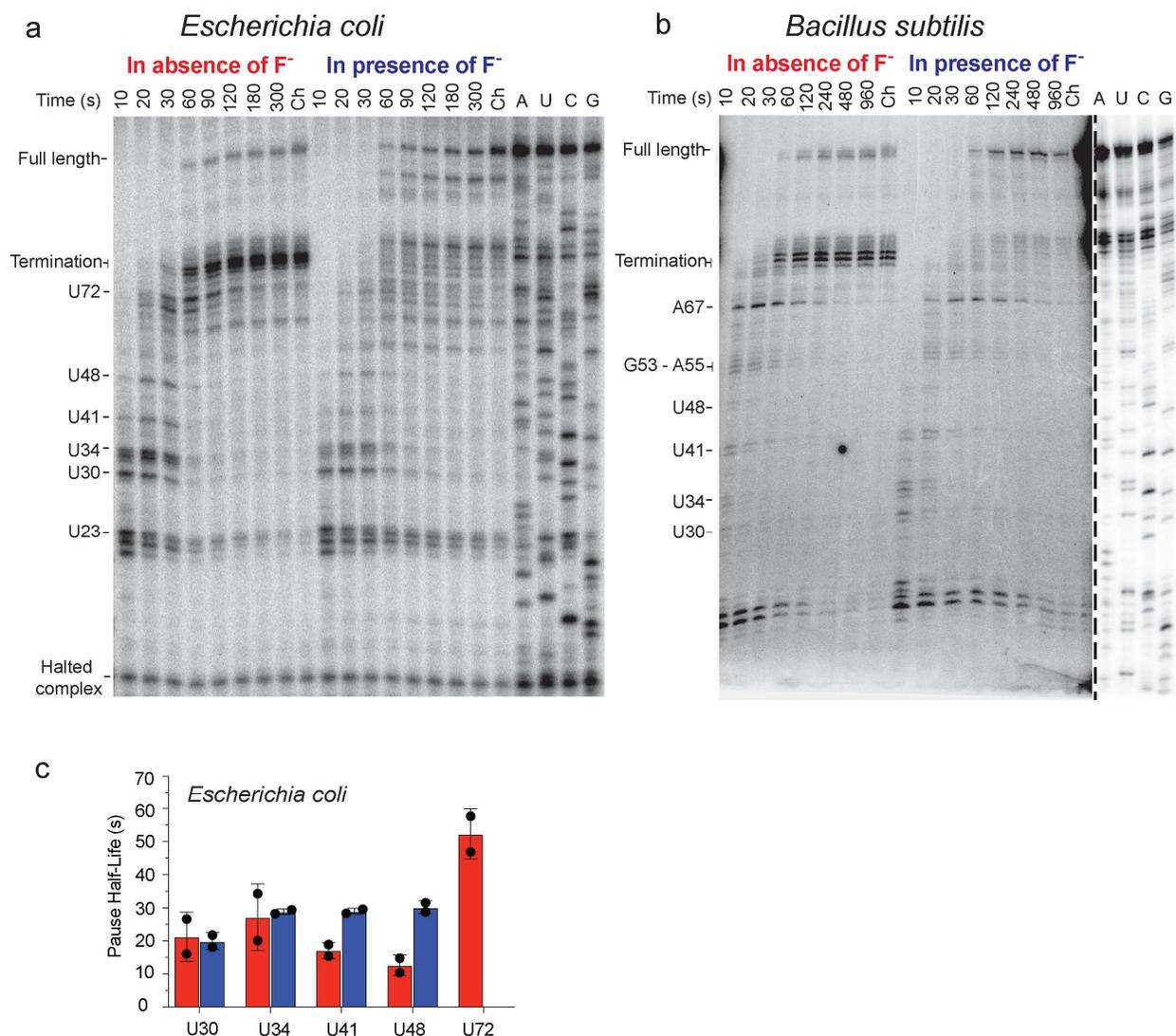
Furthermore, we used confocal fluorescence microscopy to record the fluorescence lifetime of the donor at the single-molecule level (Supplementary Fig. 14c,d). For this experiment, we immobilized the riboswitch (EC58n with only Dy547) through biotinylated *E. coli* RNAP on a PEG-passivated, streptavidin coated cover glass. A laser scanning confocal image was recorded

using a pulsed laser, which was then focused onto a specific molecule. Time-dependent fluorescence decays were then collected and fitted by deconvolution of the instrument response function (IRF), yielding the fluorescence lifetimes of single riboswitches. The fluorescence lifetime distributions in the presence of 2 mM Mg^{2+} and 2 mM Mg^{2+} plus 1 mM F^- are shown in Supplementary Fig. 14e. In the presence of only Mg^{2+} , the distribution was best fitted with a single Gaussian peak centered around 1.8 ns, and in the presence of Mg^{2+} and F^- was best fitted with two peaks centered around 1.3 ns and 1.8 ns. We conclude that the longer lifetime observed in the absence of F^- results from confinement of the fluorophore as shown schematically in Supplementary Fig. 14a. When the riboswitch converts into the docked conformation upon addition of F^- , the fluorophore becomes free to move and exhibits a shorter lifetime as a result.

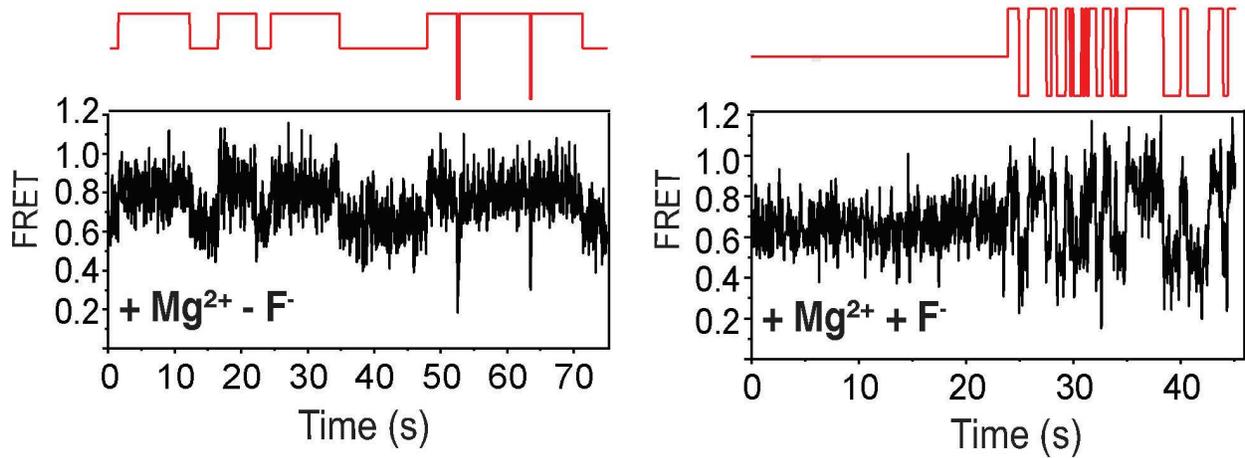
Supplementary Note 3: *In vitro* transcription assays

Overall, our single-molecule results indicate that stabilizing interactions within an EC facilitate F^- - and Mg^{2+} -dependent folding of the nascent fluoride riboswitch. During RNA synthesis, pausing by RNAP plays a critical role in guiding the folding of RNAs within a short time window.^{4,5} We used single-round transcription assays to probe how pausing by RNAP impacts the folding pathway of the riboswitch and ligand-mediated regulation of transcription. In our *in vitro* transcription assay, we observed transcriptional pauses at various locations, some of them (U30, U34 and U41) located in the aptamer domain. In agreement with a prior SHAPE-seq study,¹ these pauses are independent of the concentration of F^- (Supplementary Fig. 1). They extend the time available for the folding of the P1 hairpin. Two pauses (U48 and U72) are observed in the expression platform. The U72 pause is located downstream of the terminator hairpin and is not detected in the presence of ligand. It could therefore act as a decision point where

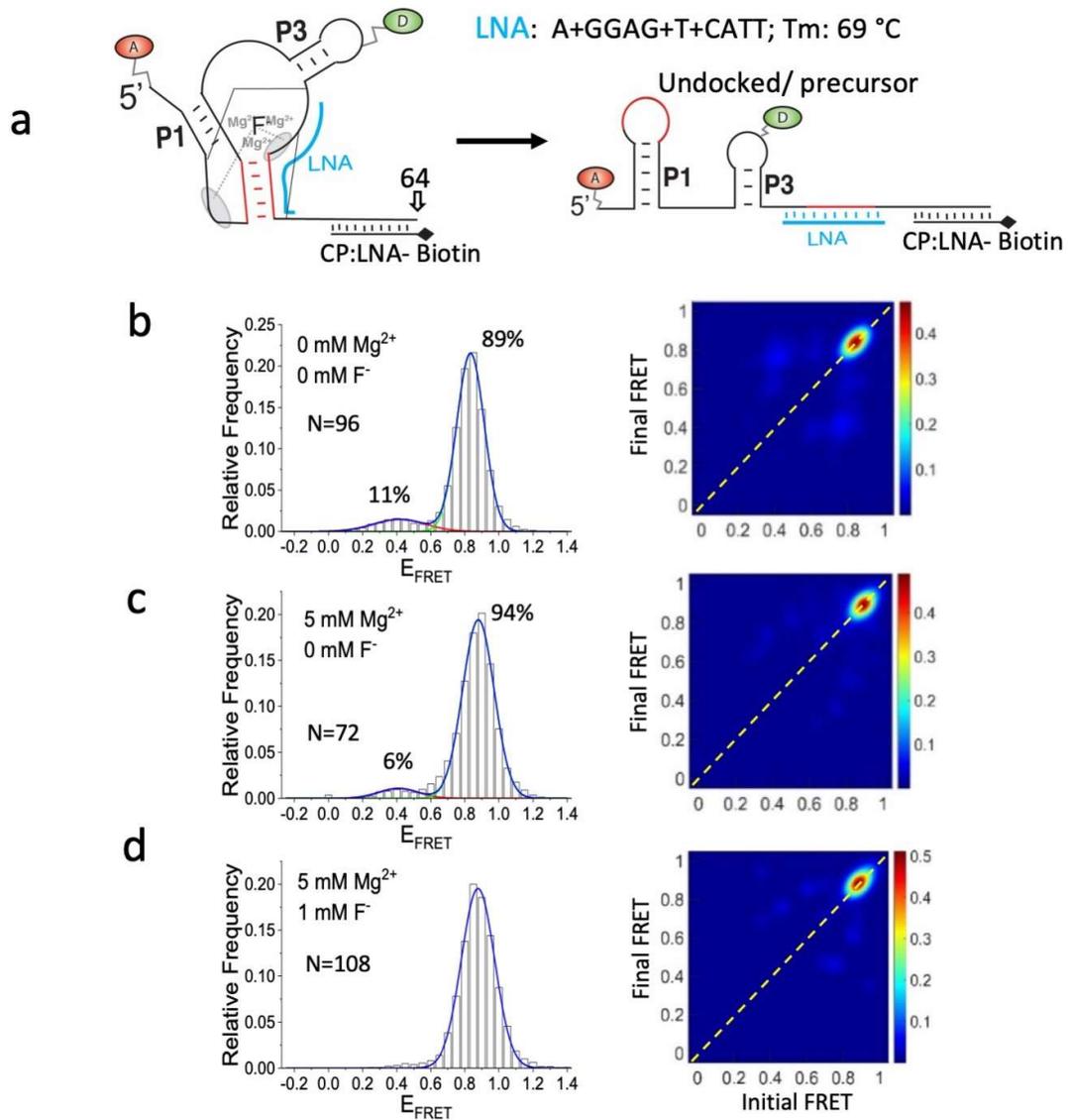
pausing by RNAP allows time for F⁻ to bind; upon further addition of 5 nucleotides by transcription finally the terminator hairpin nucleates¹.



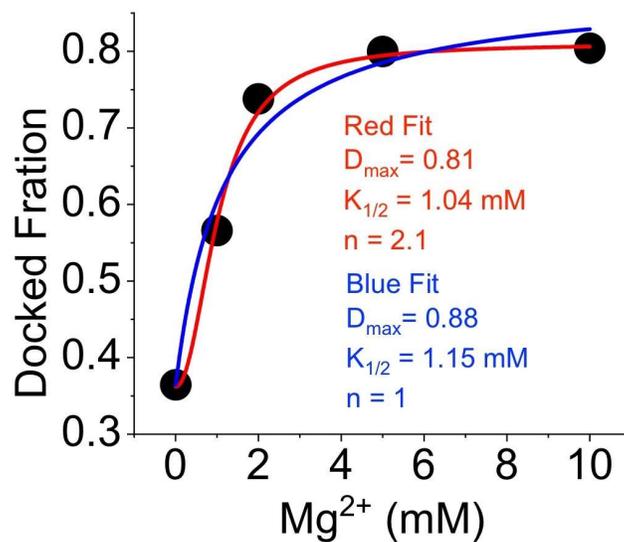
Supplementary Figure 1. *In vitro* transcriptional pausing during synthesis of the fluoride riboswitch. (a, b) Representative denaturing urea-acrylamide gel with RNA polymerase from *Escherichia coli* (*E. coli*) and *Bacillus subtilis* (*B. subtilis*), respectively. The pausing pattern is different with *B. subtilis* compare to *E. coli* transcription. The pauses at U30, U34, U41, U48 are common for both the polymerase, while the pause at U72 is specific to *E. coli* and the pauses at G53, U54, A55 and A67 are specific for *B. subtilis*. Both *In vitro* transcription assays were repeated independently for at least two times and got similar results in both cases. (c) Pause half-life with *E. coli* transcription in the absence (red) and presence (blue) of F⁻ as indicated. Error bars are presented as \pm standard deviation for n = 2 independent experiments.



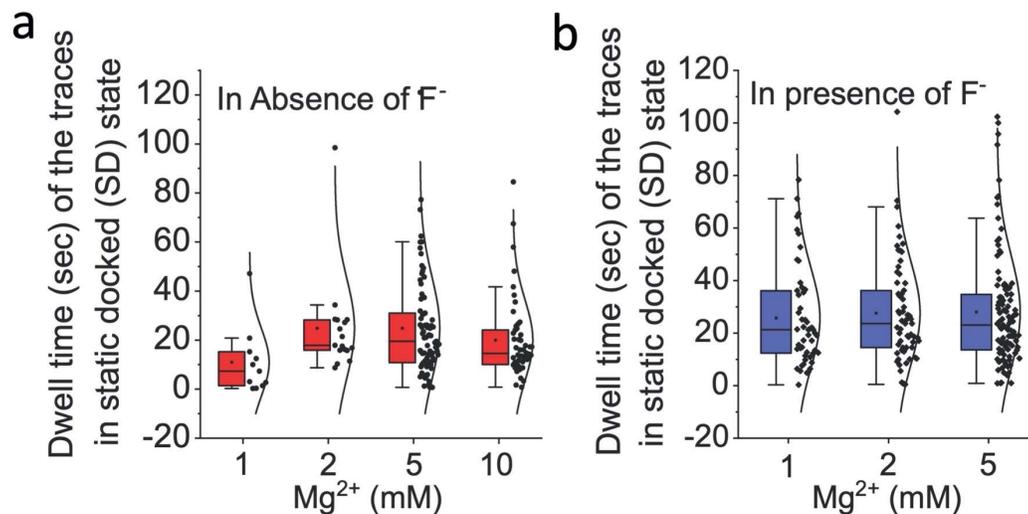
Supplementary Figure 2. Examples of time traces for RNA64 exhibiting all three FRET states. Examples of smFRET traces that show transitions between all three FRET states for RNA64 are observed in both the absence and presence of Mg²⁺ and F⁻ (observed only 5-10% of total traces in both the absence and presence of F⁻). The left trace is in presence of Mg²⁺ only, while the right trace is in the presence of both Mg²⁺ and F⁻. The HMM fit (red line) for each trace shows the three FRET states.



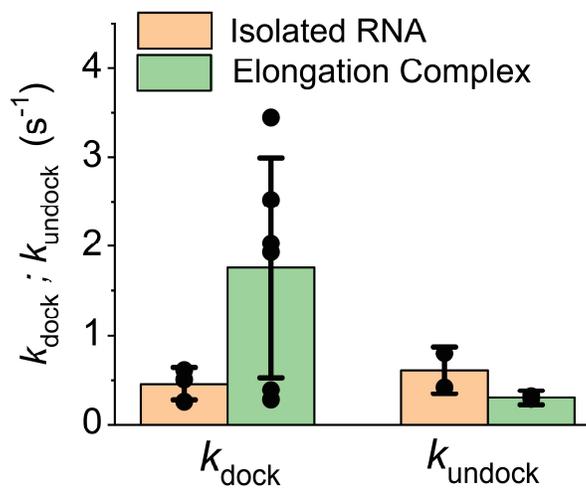
Supplementary Figure 3. smFRET data on RNA64 with LNA sequence complementary to A39-U48. This LNA sequence, which has three nucleotides (+G, +T, +C) modified with a 2'-O, 4'-C methylene bridge on their ribose, blocks pseudoknot formation along with A40-U48 long-range interactions. (a) Structure of the riboswitch with the LNA sequence shown in cyan. (b-d) FRET histograms and their corresponding TODPs (b) in the absence of Mg²⁺ and F⁻, (c) in the presence of 5 mM Mg²⁺ only and (c) in the presence of 5 mM Mg²⁺ and 1 mM F⁻.



Supplementary Figure 4. The population of the docked state changes with Mg^{2+} concentration. This titration curve is well fitted by a Hill equation with a cooperativity coefficient of 2.1 (red); the fit is poor when using a non-cooperative model (blue).

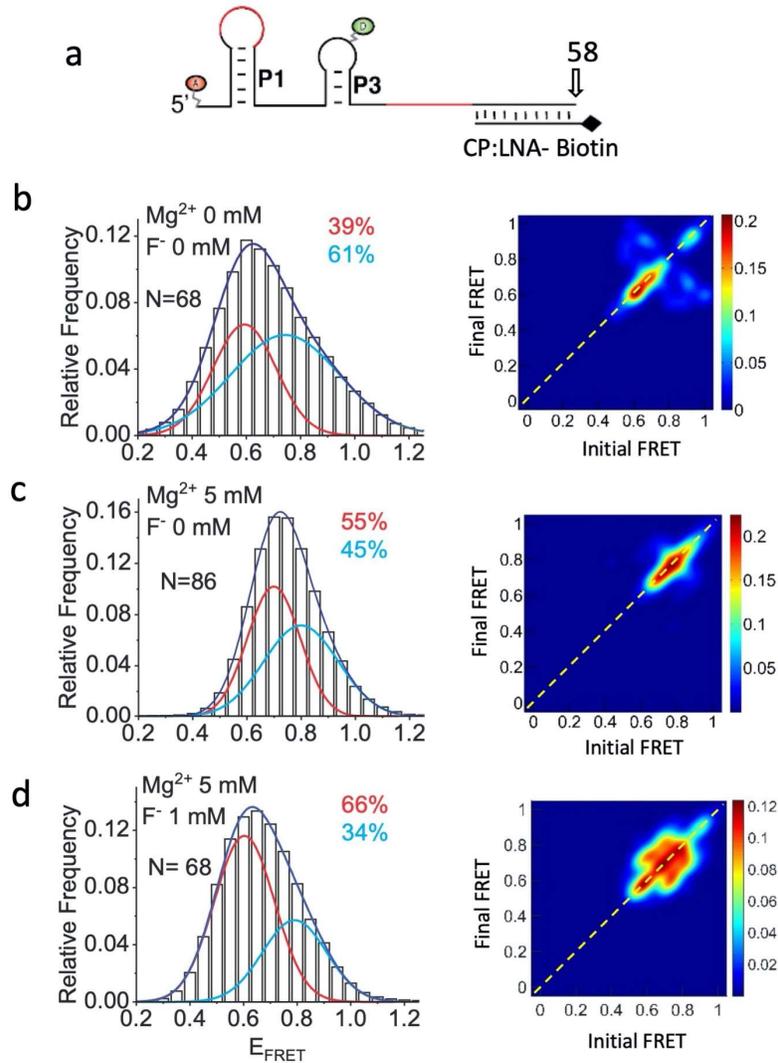


Supplementary Figure 5. Photobleaching times of static traces in the mid-FRET state of RNA64 (SD traces). Data are shown at different concentrations of Mg^{2+} (a) in the absence of F^- and (b) in the presence of 1 mM F^- . The mean dwell times with the interquartile range (IQR) are 11 s (IQR: 1.7 s - 12 s, $n = 11$), 25 s (IQR: 16 s - 27 s, $n = 17$), 25 s (IQR: 11 s - 29 s, $n = 79$) and 20 s (IQR: 10 s - 24 s, $n = 50$) in the presence of 1, 2, 5 and 10 mM Mg^{2+} (without F^-), and 26 s (IQR: 12 s - 36 s, $n = 51$), 28 s (IQR: 15 s - 35 s, $n = 65$), 28 s (IQR: 14 s - 35 s, $n = 107$) in the presence of 1, 2 and 5 mM Mg^{2+} (with 1 mM F^-). These data indicate that the molecules remain in the SD conformation longer than the mean photobleaching time of around 20 s.

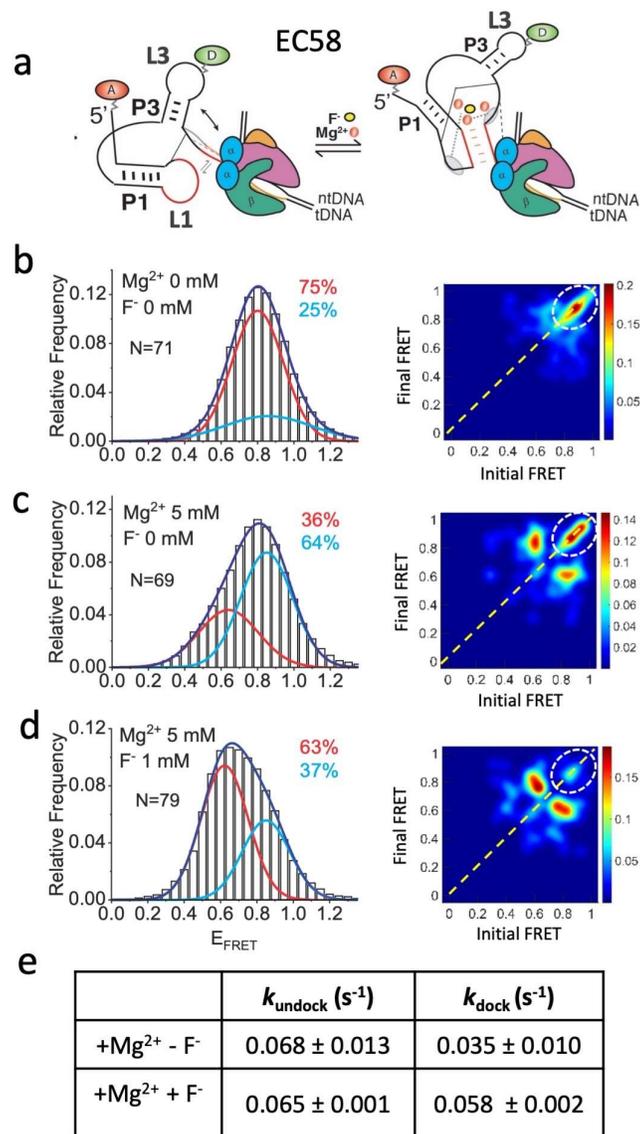


Supplementary Figure 6. Proximity of RNA polymerase favors the docked conformation of RNA64.

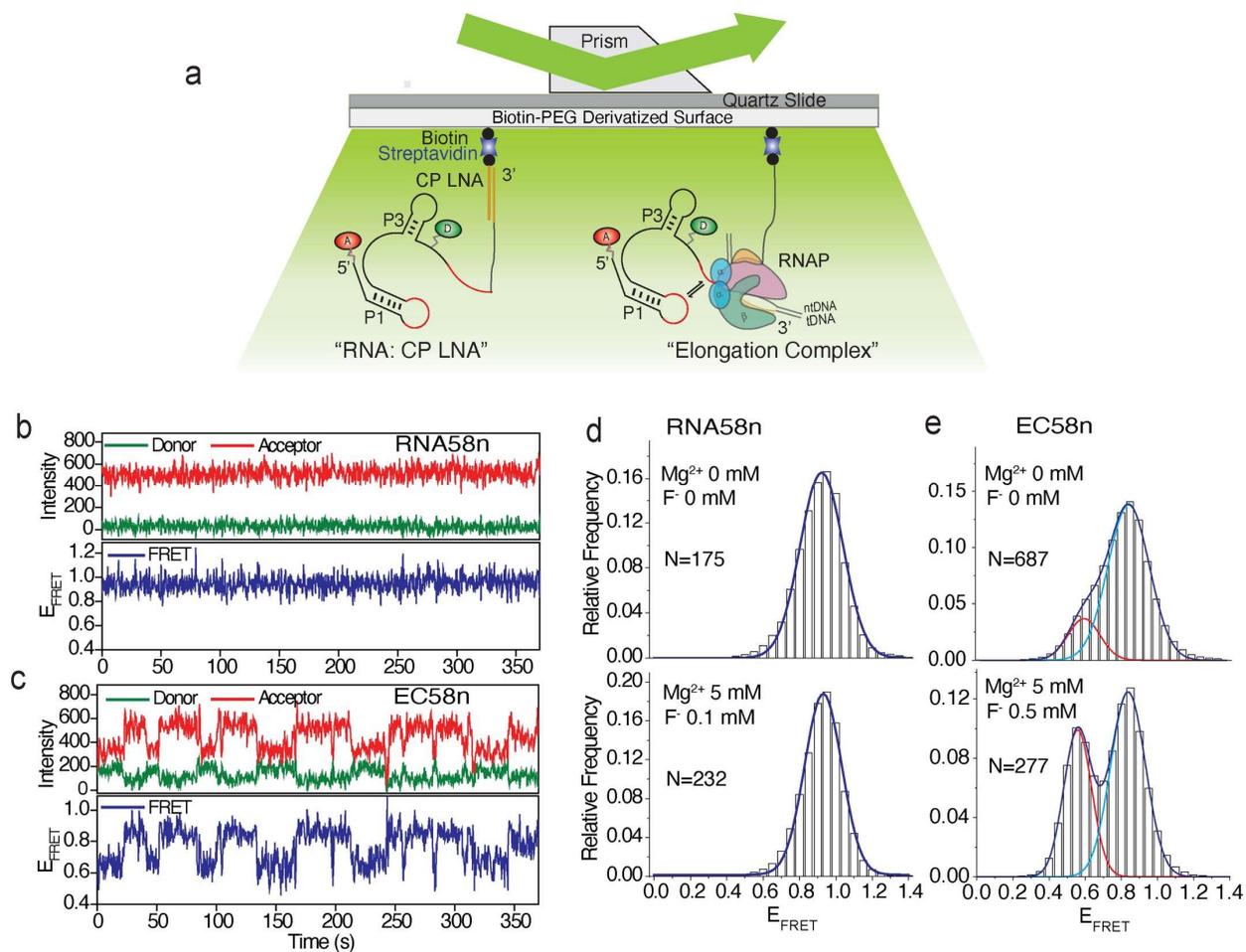
Kinetic parameters for high- to mid-FRET (prefolded to docked state) transitions. The errors represent the standard deviation of two separate datasets. Due to the small number of dynamic traces, bootstrapping was used to calculate error bars for the EC data and is presented as \pm standard deviation for $n \geq 2$ independent data sets (with the values k_{dock} : $0.5 \pm 0.2 \text{ s}^{-1}$ and k_{undock} : $0.6 \pm 0.3 \text{ s}^{-1}$ for the isolated RNA and k_{dock} : $1.7 \pm 1.2 \text{ s}^{-1}$ and k_{undock} : $0.3 \pm 0.1 \text{ s}^{-1}$ for the Elongation Complex).



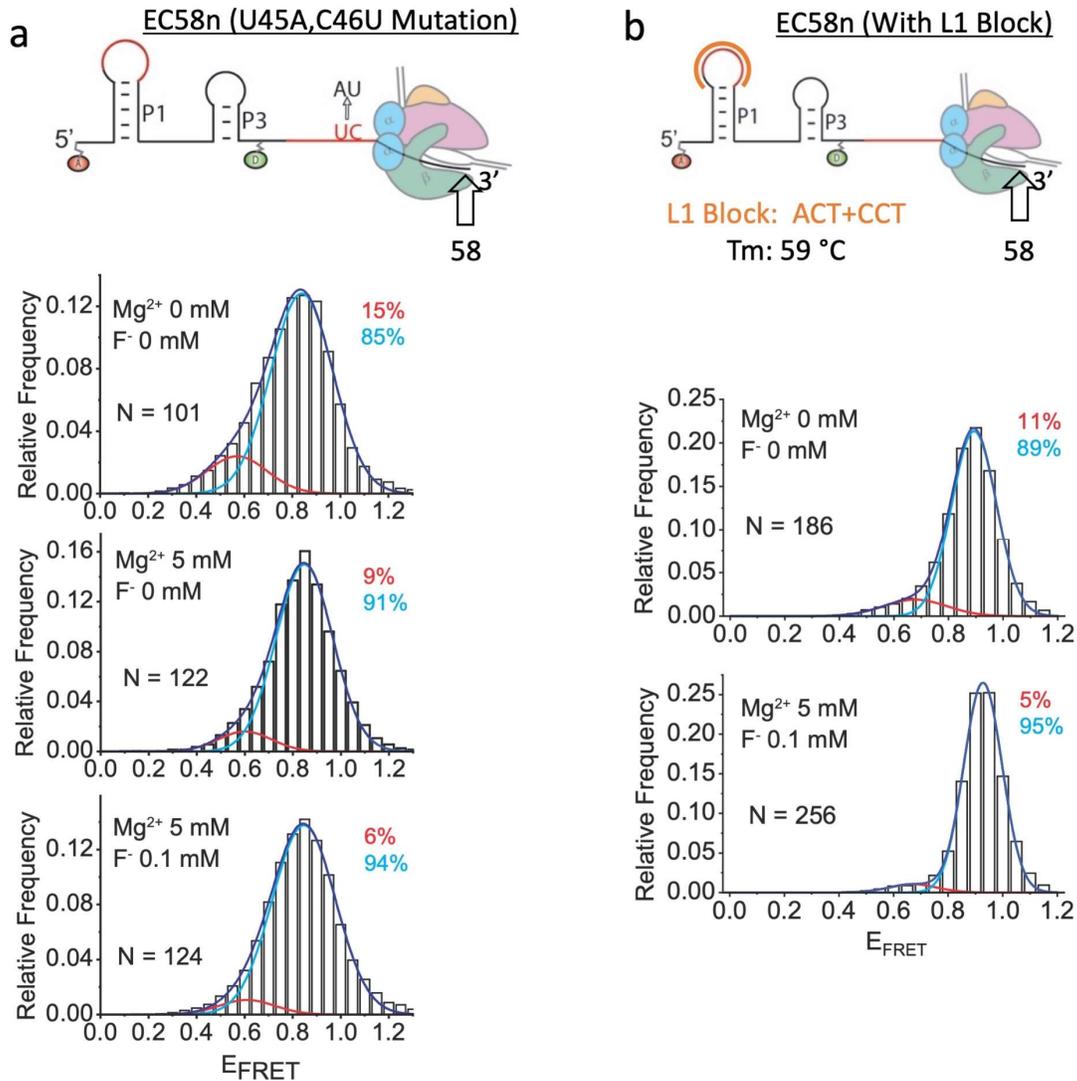
Supplementary Figure 7. smFRET histograms for RNA58. (a) A schematic representation of RNA58. (b-d) smFRET histograms and their corresponding TODPs in the absence and presence of Mg²⁺ and F⁻ as indicated on each plot. The percent population of both fitted peaks are shown in respective colors in each histogram panel, and the number of molecules that were analyzed is indicated by "N".



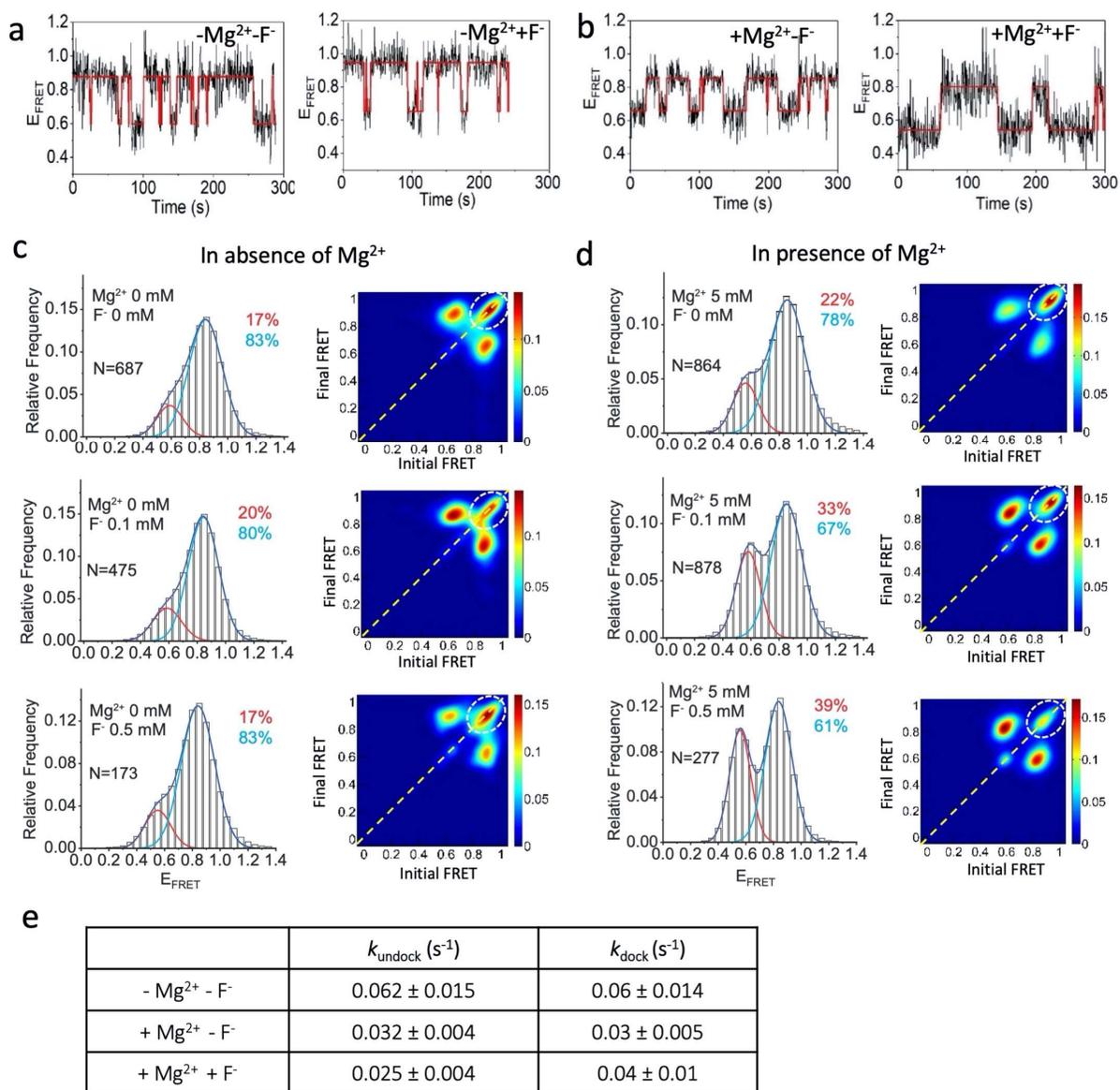
Supplementary Figure 8. Dynamics for EC58. (a) Schematic of precursor (left) and docked (right) RNA structures for EC58. (b-d) FRET histograms and TODPs for EC58, recorded at an exposure time of 400 ms. In the absence of Mg²⁺ and F⁻, most of the population is in the high-FRET (precursor) state; on addition of Mg²⁺ and F⁻ transitions occur between the high- and mid-FRET states. The population of the mid-FRET state increases on addition Mg²⁺ and F⁻ due to replacement of high-FRET static traces (highlighted by on-diagonal white circles) with dynamic traces as shown in the TODPs. The percent population of both fitted peaks are shown in respective colors, and the number of molecules that were analyzed is indicated by "N" in each FRET histogram panel. (e) Kinetic parameters for transitions between the high- and mid-FRET states. The reported errors are the standard error of the fit.



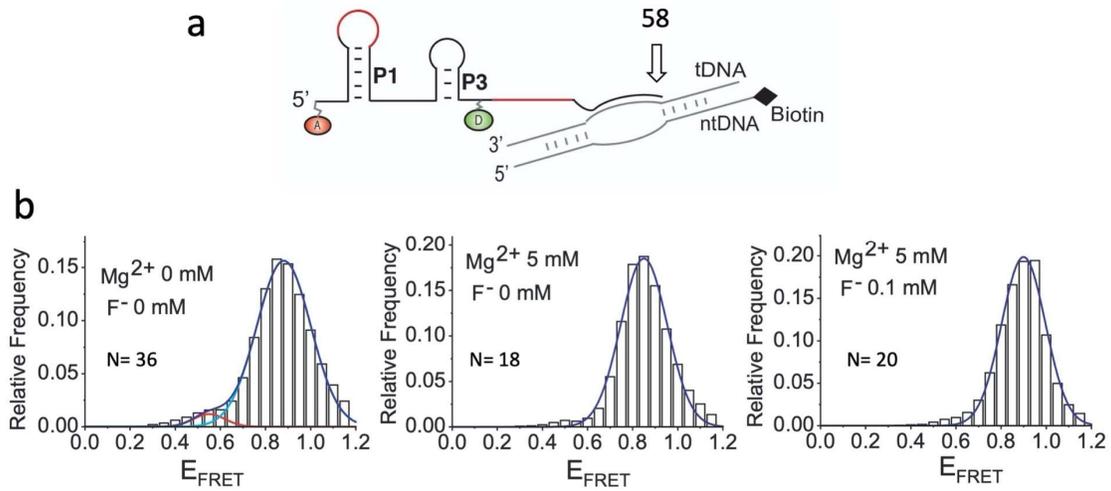
Supplementary Figure 9. RNA58n and EC58n with donor fluorophore at A39. (a) smFRET experimental setup for isolated RNA and elongation complex (note the new donor labeling position). The RNA is immobilized through biotinylated CP LNA or biotinylated *E. coli* RNAP, as indicated. (b, c) Representative FRET time traces for RNA58n and EC58n in presence of 5mM Mg^{2+} , respectively. (d, e) FRET distribution observed for RNA58n (d) and EC58n (e), in the absence and presence of Mg^{2+} and F^- . Only the high-FRET state is observed for RNA58n, while for EC58n a mid-FRET state is also observed. The number of molecules that were analyzed is indicated by "N" in the FRET histogram panel.



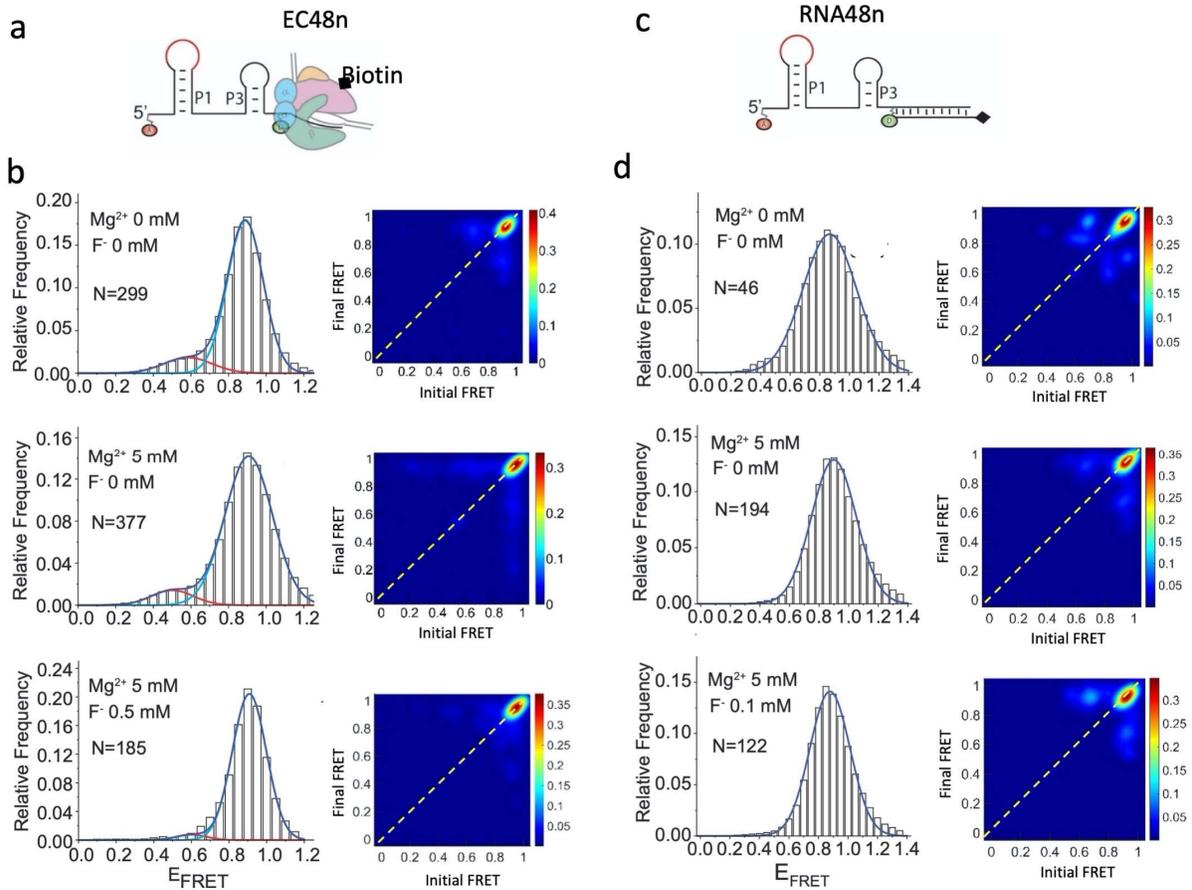
Supplementary Figure 10. The mid-FRET state in EC58n represents the docked conformation. smFRET measurements for EC58n (a) with a U45A,C46U double mutation and (b) with L1 block sequence, both of which disrupt pseudoknot formation. In both cases the population of the mid-FRET state is reduced relative to the wild-type and without the L1 block sequence. The number of molecules that were analyzed is indicated by "N" in the FRET histogram panel.



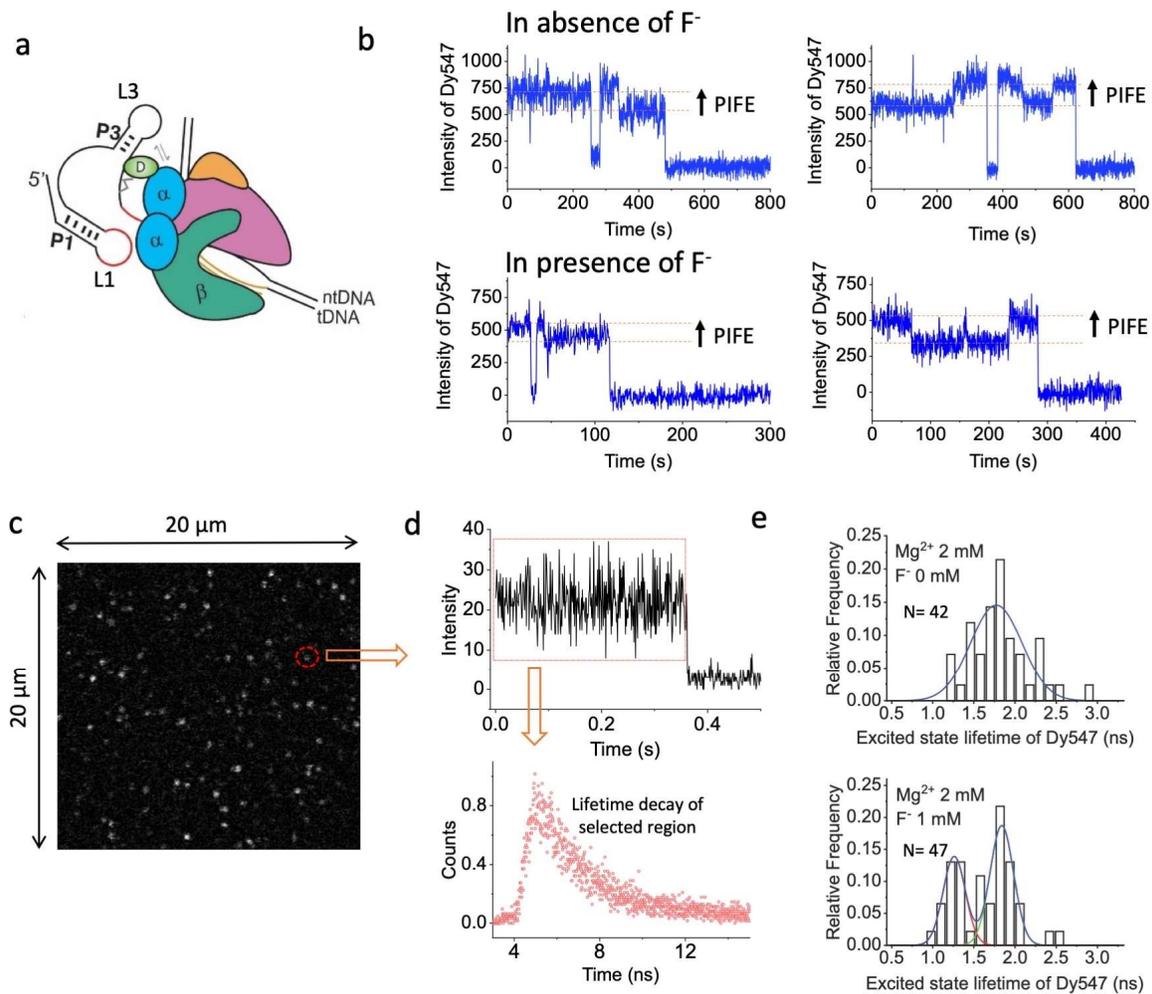
Supplementary Figure 11. Dynamics for EC58n. (a, b) Representative FRET efficiency traces for EC58n in the absence (a) and presence (b) of Mg^{2+} at an exposure time of 400 ms. (c, d) smFRET histograms and their corresponding TODPs at varying concentrations of F^{-} in the absence (c) and presence (d) of Mg^{2+} . In both cases FRET histograms were best fitted with two Gaussian peaks. The percent population of both fitted peaks are shown in respective colors in each histogram panel. In the presence of Mg^{2+} , the population of the mid-FRET state increases on addition of F^{-} due to replacement of high-FRET static traces (highlighted by on-diagonal white circles) with dynamic traces as shown in the TODPs in panel d. The number of molecules that were analyzed is indicated by "N" in the FRET histogram panel. (e) Kinetic parameters for interconversion between the mid- and high-FRET states. The reported errors are the standard deviation of at least three independent datasets.



Supplementary Figure 12. RNA58n with DNA template. This control experiment suggests that the mid-FRET peak (around $E_{\text{FRET}} \approx 0.55$) observed in EC58n requires the presence of RNAP, as it is not observed when only the DNA template is present. (a) Schematic of RNA58n with DNA template experiment (b) smFRET histograms observed in the absence and presence of Mg^{2+} and F^- .



Supplementary Figure 13. smFRET data for EC48n and RNA48n. (a) Schematic of EC48n. (b) smFRET histograms and their corresponding TODPs for EC48n in the absence and presence of Mg²⁺ and F⁻. (c) Schematic of RNA48n. (d) smFRET histograms and their corresponding TODPs for RNA48n in the absence and presence of Mg²⁺ and F⁻.



Supplementary Figure 14. Change in micro-environment upon binding of F^- to EC58n (a) A schematic representation of EC58n labeled with a single fluorophore (DY547) at A39. (b) Representative DY547 intensity traces exhibiting PIFE in the absence (top two) and presence (bottom two) of F^- . (c-e) Excited state fluorescence lifetime measurements for single EC58n complexes. (c) Representative laser scanning confocal image of immobilized EC58n complexes. (d) Procedure for extracting the fluorescence lifetime: a time window without any bleaching or blinking events is selected from the photon count time traces (top). This gets converted into a decay profile (bottom), which was best fitted upon deconvolution with the IRF to extract the excited state lifetime of the dye. (e) Lifetime distribution in the presence of 2 mM Mg^{2+} and the absence and presence of F^- as indicated. Two lifetime populations are observed in the presence of F^- .

Supplementary Table 1. Observed kinetic parameters. Errors represent the standard deviation of two independent datasets.

Isolated RNA (RNA64)						
	Low- to mid-FRET		High- to mid-FRET		% SD traces	
	k_{dock} (s ⁻¹)	k_{undock} (s ⁻¹)	k_{dock} (s ⁻¹)	k_{undock} (s ⁻¹)	0mM F ⁻	1mM F ⁻
0 mM Mg ²⁺	1.8 ± 0.2	3.6 ± 0.6	2.6 ± 0.2	3.0 ± 0.2	0	0
1 mM Mg ²⁺	1.8 ± 0.2	0.5 ± 0.0	1.2 ± 0.3	0.6 ± 0.1	5	32
2 mM Mg ²⁺	1.4 ± 0.1	0.5 ± 0.2	0.5 ± 0.2	0.6 ± 0.3	10	52
5 mM Mg ²⁺	1.1 ± 0.0	0.5 ± 0.4	0.9 ± 0.3	0.4 ± 0.1	32	63
10 mM Mg ²⁺	0.7 ± 0.1	0.4 ± 0.3	1.3 ± 0.5	0.5 ± 0.3	40	
2 mM Mg ²⁺ + 1mM F ⁻	2.1 ± 0.1	1.3 ± 0.3	0.6 ± 0.1	0.7 ± 0.5		
RNA in elongation complex (EC64)						
2 mM Mg ²⁺	3.6 ± 0.2	0.5 ± 0.2	1.7 ± 1.2	0.3 ± 0.1	28	69
	$\Delta\Delta G_{dock}^0$ (kcal/mol)		$\Delta\Delta G_{dock}^{++}$ (kcal/mol)			
RNA64/EC64 (2mM Mg ²⁺) *	0.61		0.58			

Equations used to compute energies (see methods);

$$* \Delta\Delta G_{dock}^0 = \Delta G_{dock}^{0, RNA64} - \Delta G_{dock}^{0, EC64} \text{ and } \Delta\Delta G_{dock}^{++} = -RT \ln(k_{dock}^{RNA64} / k_{dock}^{EC64}).$$

Supplementary Table 2. Oligonucleotides used in this study. The underlined sequences are complementary to allow for surface or transcription bubble capture.

Dye-labeled RNA oligonucleotides	
RNA oligonucleotide-1	5'-(Cy5-9S-U)AG GCG AUG GAG UUC GCC AUA AAC GCU GC(5-LC-N-U) UAG CUA AU-3'
RNA oligonucleotide-2 "n"	5'-(N3-U)AG GCG AUG GAG UUC GCC AUA AAC GCU GCU UAG CU(A[2'Dy547]) AU-3'
Non-labeled RNA oligonucleotides	
EC RNA48p	5'- pCAC <u>ACG A</u> -3'
EC RNA58	5'- pGAC UCC UAA <u>CCA CAC GA</u> -3'
EC RNA64	5'- pGAC UCC UAC CAG UAA <u>CCA CAC GA</u> -3'
EC RNA64 mutant 'U48A'	5'- pGAC UCC AAC CAG UAA <u>CCA CAC GA</u> -3'
EC RNA58n mutant 'M19'	5'-pGAC AUC UAA CCA CAC GA-3'
Splint DNA/ LNA	
Splint LNA48p	5'- <u>TC+G +TG+T</u> GAT TAG CTA AGC AGC GTT TAT-3'
Splint DNA58	5'-TCG TGT GGT TAG GAG TCA TTA GCT AAG CAG CGT TTA T-3'
Splint DNA64	5'-GTG GTT ACT GGT AGG AGT CAT TAG CTA AGC AGC GTT TAT-3'
Splint DNA64 mutant 'U48A'	5'-GTG GTT ACT GGT TGG AGT CAT TAG CTA AGC AGC GTT TAT-3'
Splint DNA58n mutant 'M19'	5'-TCG TGT GGT TAG ATG TCA TTA GCT AAG CAG CGT TTA T-3'
Other oligonucleotides	
tDNA for RNA48p	5'-CTC TGA ATC TCT TCC <u>TCG TGT</u> GAT TCC AAC TTG CTG GGA-3'
tDNA for RNA58, RNA64	5'-CTC TGA ATC TCT TCC <u>TCG TGT</u> GGT TCC AAC TTG CTG GGA-3'
biotin-ntDNA	5'-TCC CAG CAA GTT GCT GAG CTC TTC GGA AGA GAT TCA GAG CCC CCC CCC CCC C/3Bio/ -3'
ntDNA	5'-TCC CAG CAA GTT GCT GAG CTC TTC GGA AGA GAT TCA GAG-3'
CP LNA for RNA48p	/5Biosg/ T+CGT+G+TGA+TT
CP LNA for RNA58, RNA64	/5Biosg/ T+CG+TG+TGG+TT
10nt LNA complementary to A39-U48	5'-A+GG AG+T +CAT T-3'
L1 block LNA	5'-ACT +CCT-3'
Oligonucleotides used for <i>in vitro</i> transcription assays	
T7A1-PCR	5'-TCC AGA TCC CGA AAA TTT ATC AAA AAG AGT ATT G-3'
crcB-reverse	5'-TCT CAC CTC TTT AAA TAG CTT GCT CAA AAA AAT AG-3'
LambdaPr-CrcB-FWD	5'-GTG CGT GTT GAC TAT TTT ACC TCT GGC GGT GAT AAT GGT TGC ATC CAG AGG GAC ACC CAG AAG A-3'

Supplementary References

1. Watters, K. E., Strobel, E. J., Yu, A. M., Lis, J. T. & Lucks, J. B. Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat. Struct. Mol. Biol.* **23**, 1124–1131 (2016).
2. Hwang, H. & Myong, S. Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions. *Chem. Soc. Rev.* **43**, 1221–9 (2014).
3. Hwang, H., Kim, H. & Myong, S. Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 7414–8 (2011).
4. Pan, T. & Sosnick, T. RNA Folding During Transcription. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 161–175 (2006).
5. Perdriquet, G. A., Artsimovitch, I., Furman, R., Sosnick, T. R. & Pan, T. Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proc. Natl. Acad. Sci.* **109**, 3323–3328 (2012).