Molecular Cell
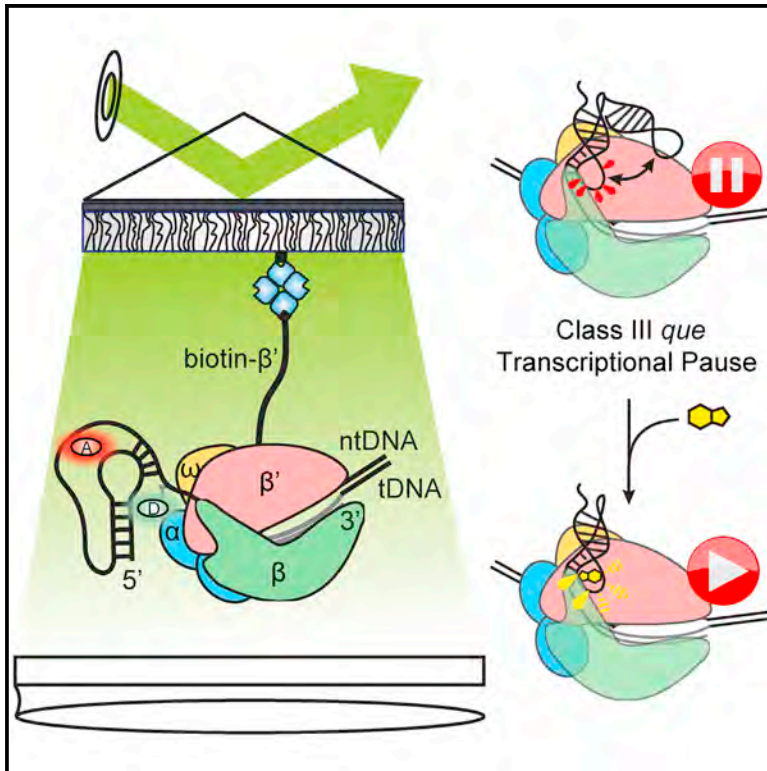
Ligand Modulates Cross-Coupling between Riboswitch Folding and Transcriptional Pausing

Graphical Abstract

Highlights

- smFRET, biochemistry, and simulations probe co-transcriptional riboswitch folding

- Nascent RNA folding is disfavored by the DNA template and aided by RNAP

- Transcriptional pausing is stabilized by RNA pseudoknot and destabilized by ligand

Authors

Julia R. Widom, Yuri A. Nedialkov, Victoria Rai, Ryan L. Hayes, Charles L. Brooks III, Irina Artsimovitch, Nils G. Walter

Correspondence

nwalter@umich.edu

In Brief

Proper folding of RNA during transcription is a critical prerequisite to the many functions of RNA. Widom et al. show that the co-transcriptional folding of a riboswitch is governed by interactions with RNA polymerase that concurrently affect the speed of transcription.

CellPress

Article

Ligand Modulates Cross-Coupling between Riboswitch Folding and Transcriptional Pausing

Julia R. Widom,[1,2,6] Yuri A. Nedialkov,[3,4,7] Victoria Rai,[1,2,5,8] Ryan L. Hayes,[5] Charles L. Brooks III,[5] Irina Artsimovitch,[3,4] and Nils G. Walter[1,2,9,*]

[1]Single Molecule Analysis Group, Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA
[2]Center for RNA Biomedicine, University of Michigan, Ann Arbor, MI 48109, USA
[3]Department of Microbiology, The Ohio State University, Columbus, OH 43210, USA
[4]Center for RNA Biology, The Ohio State University, Columbus, OH 43210, USA
[5]Biophysics Program and Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA
[6]Present address: Department of Chemistry and Biochemistry, University of Oregon, Eugene, OR 97403, USA
[7]Present address: Division of Pharmaceutics and Pharmaceutical Chemistry, College of Pharmacy, The Ohio State University, Columbus, OH 43210, USA
[8]Present address: Department of Gastroenterology, University of Chicago, Chicago, IL 60637, USA
[9]Lead Contact
*Correspondence: nwalter@umich.edu
https://doi.org/10.1016/j.molcel.2018.08.046

SUMMARY

Numerous classes of riboswitches have been found to regulate bacterial gene expression in response to physiological cues, offering new paths to antibacterial drugs. As common studies of isolated riboswitches lack the functional context of the transcription machinery, we here combine single-molecule, biochemical, and simulation approaches to investigate the coupling between co-transcriptional folding of the pseudoknot-structured preQ$_1$ riboswitch and RNA polymerase (RNAP) pausing. We show that pausing at a site immediately downstream of the riboswitch requires a ligand-free pseudoknot in the nascent RNA, a precisely spaced sequence resembling the pause consensus, and electrostatic and steric interactions with the RNAP exit channel. While interactions with RNAP stabilize the native fold of the riboswitch, binding of the ligand signals RNAP release from the pause. Our results demonstrate that the nascent riboswitch and its ligand actively modulate the function of RNAP and vice versa, a paradigm likely to apply to other cellular RNA transcripts.

INTRODUCTION

To render bacteria responsive to rapidly changing environmental conditions, gene expression is regulated by a variety of means, including mechanisms in which riboswitches modulate messenger RNA (mRNA) synthesis, translation, or stability in response to physiological cues. Riboswitches, structured non-coding RNA elements typically found in 5′ untranslated re-

gions, are thought to regulate the expression of up to 4% of genes in certain bacteria (Sherwood and Henkin, 2016). A typical riboswitch contains a ligand-binding aptamer domain upstream of an expression platform whose conformation changes in response to ligand binding by the aptamer domain, thereby regulating expression of the downstream gene. The at least 40 classes of riboswitches for which ligands have been identified thus far (Greenlee et al., 2018) sense diverse ligands including metabolites (Roth et al., 2007), nucleobases (Frieda and Block, 2012), other RNAs (Zhang and Ferré-D'Amaré, 2013), and more. Due to their critical importance in bacterial gene regulation, riboswitches present attractive targets for the development of new antibiotic therapies (Deigan and Ferré-D'Amaré, 2011).

Mechanistic studies have been performed on many isolated riboswitches (Duesterberg et al., 2015; Holmstrom et al., 2014; Liberman et al., 2015; Suddala et al., 2013); however, in the cell, riboswitch folding occurs in the context of the transcription machinery. The speed of transcription by RNA polymerase (RNAP) imposes kinetic constraints on the rate of riboswitch folding, while the nascent RNA itself has the potential to modulate transcription in a variety of different ways. Most studies of co-transcriptional RNA folding utilize readouts that result from transcription itself, including SHAPE probing to detect backbone flexibility (Watters et al., 2016) and single-molecule force and fluorescence techniques (Frieda and Block, 2012; Uhm et al., 2018). When naked RNA is compared to RNA folded during active transcription, the effects of interactions with the transcription machinery are convoluted with the kinetic effects of non-equilibrium folding. As a result, potential direct effects of the presence of the DNA template and RNAP on folding remain largely unexplored.

The preQ$_1$ riboswitch from the Gram-positive bacterium *Bacillus subtilis* (*Bsu*) is thought to act through a transcriptional mechanism in which the binding of its ligand, the tRNA nucleotide precursor 7-methylamino-7-deazaguanine
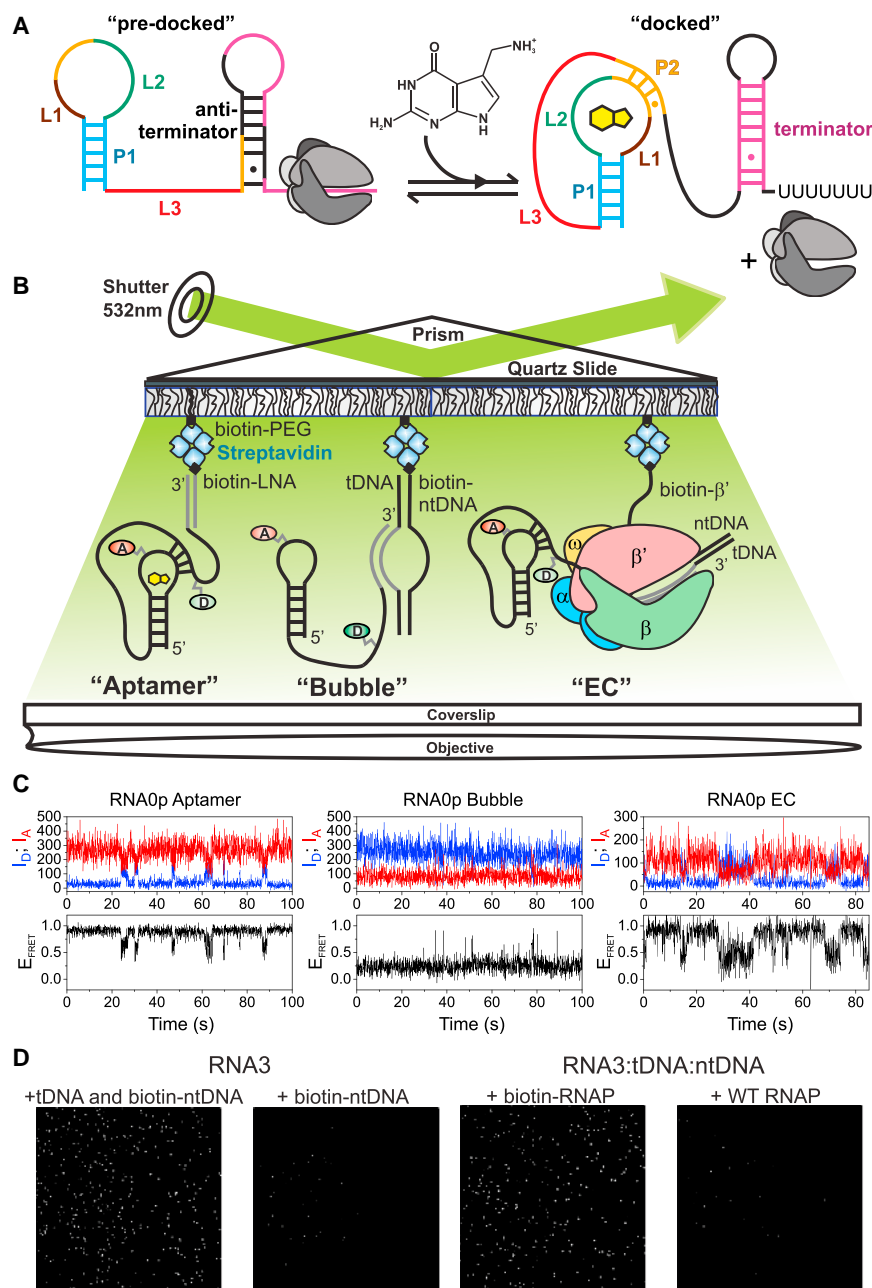
Figure 1. smFRET Investigation of the preQ$_1$ Riboswitch in Active Transcription Complexes

(A) Regulation of gene expression by the preQ$_1$ riboswitch. Ligand-induced pseudoknot docking favors the formation of a terminator hairpin, leading to decreased expression of the downstream genes.

(B) smFRET experimental setup. "Aptamer" immobilization utilizes a biotinylated LNA capture probe. "Bubble" immobilization utilizes a partially complementary DNA bubble with a biotinylated non-template DNA. Elongation complex ("EC") immobilization utilizes biotinylated *Eco* RNAP. Locations of donor ("D," green) and acceptor ("A," red) fluorophores are indicated.

(C) Representative smFRET traces for RNA0p under aptamer (left), bubble (middle), and EC (right) immobilization in the presence of 1 mM MgCl$_2$ and 100 nM preQ$_1$.

(D) iCCD images show that bubble immobilization is dependent on the tDNA in addition to the biotinylated ntDNA and the fluorophore-labeled RNA (left). Efficient EC immobilization requires biotinylated RNAP (right).

observed immediately after the aptamer domain is synthesized (U46, Figure S1), finding that the presence of RNAP stabilizes the folded conformation of the riboswitch. Additionally, we found that the U46 pause itself is stabilized by interactions between a nascent RNA pseudoknot and RNAP in conjunction with a precisely spaced sequence that resembles the pause consensus identified *in vivo* (Figure S1D) (Larson et al., 2014; Vvedenskaya et al., 2014) and is destabilized (released) upon ligand binding. This pause, hereafter called the "*que* pause," exhibits features distinct from the two previously defined mechanisms for transcriptional pausing in bacteria (Artsimovitch and Landick, 2000), leading us to term it "class III." Our observations demonstrate an unprecedented cross-coupling between riboswitch folding and RNAP pausing that is likely to govern the co-transcriptional folding of numerous RNAs.

RESULTS

Direct Observation of Riboswitch Folding in Transcription Complexes

Due to the sensitivity of RNA folding to macromolecular crowding (Daher et al., 2018; Dupuis et al., 2014), ionic conditions (Suddala et al., 2015), and other factors, we hypothesized that the behavior of the preQ$_1$ riboswitch would be significantly altered upon incorporation into a transcription elongation complex

(preQ$_1$), stabilizes a pseudoknot structure that favors the formation of a terminator hairpin, decreasing expression of enzymes involved in queuosine biosynthesis (Figure 1A) (Roth et al., 2007). With a minimal aptamer containing only 36 nucleotides, it is one of the smallest known riboswitches. By investigating this riboswitch in the context of an active transcription complex, we show that a network of interactions between the nascent RNA, DNA template, and RNAP enable the transcription machinery to cross-couple regulation of nascent RNA folding and transcriptional pausing. Specifically, we used single-molecule FRET (smFRET) to probe the conformation of the nascent riboswitch at and beyond a transcriptional pause

(EC) containing a DNA template and RNAP. We therefore used smFRET to investigate the effects of the transcription machinery on the folding of the nascent riboswitch. We immobilized the riboswitch on polyethylene glycol (PEG)-passivated, streptavidin-coated quartz slides for imaging via prism-based total internal reflection fluorescence (p-TIRF) microscopy. The donor fluorophore was placed at the $2'$ position of a guanosine residue within the L3 "tail" of the riboswitch, and the acceptor was placed on a 4-aminoallyl-uracil residue within the loop L2 (Figure 1). In an isolated RNA aptamer, the "docked" pseudoknot conformation in which the helix P2 is formed yields a high FRET efficiency (E_{FRET}), while a mid-FRET state is observed for the "pre-docked" conformation in which P2 is not fully intact (Santner et al., 2012; Suddala et al., 2013, 2015) (Figure 1C).

We used a systematic series of immobilization strategies to study the riboswitch at varying degrees of biological complexity (Figure 1B). Immobilization through a biotinylated LNA capture probe (CP) allowed investigation of the isolated aptamer (henceforth called "aptamer immobilization"). A partially complementary DNA bubble with a biotinylated non-template (nt) strand enabled investigation of the nucleic acid framework of an EC ("bubble immobilization"). Finally, immobilization through biotinylated *Escherichia coli* (*Eco*) RNAP allowed observation of the riboswitch in a complete EC ("EC immobilization") (Daube and von Hippel, 1992; Sidorenkov et al., 1998). While *Eco* RNAP has been found to exhibit more pronounced pausing than *Bsu* RNAP in some cases (Grundy and Henkin, 2004), many pauses are recognized by both RNAPs, including several that have been identified in riboswitches (Helmling et al., 2018; Lemay et al., 2011; Steinert et al., 2017). In our *in vitro* transcription assays, we found that while *Eco* and *Bsu* RNAPs did not yield identical results, they both recognized the *que* pause (Figure S1). As is frequently done, we therefore chose to use an experimental system based on the better-studied *Eco* RNAP to identify not only specific features of preQ$_1$ riboswitch folding in the context of a known transcription machinery but also factors that may contribute to the folding of nascent RNA in general.

To investigate co-transcriptional folding wherein riboswitch properties and interactions with the transcription machinery may evolve as additional RNA is synthesized, we utilized RNAs containing varying amounts of the riboswitch expression platform. The shortest RNA species studied by smFRET is analogous to the species present at the *que* pause ("RNA0 pause," henceforth called "RNA0p"); we also investigated RNAs with an additional three (RNA3) or ten (RNA10) nucleotides of expression platform (Figure S2A). We first confirmed that our assembly procedures yielded the desired complexes and immobilization mechanisms. Efficient bubble immobilization required that the template DNA (tDNA) be present along with the fluorophore-labeled RNA and biotinylated ntDNA, indicating that the molecules we observe contain all three components. EC immobilization was efficient only with biotinylated RNAP, indicating the absence of nonspecific sticking to the slide (Figure 1D).

To determine the correct number of states with which to model our data, we fit the traces from several experimental conditions using 1-, 2- and 3-state kinetic models and computed the mean Bayesian information criterion (BIC) and modified Bayesian information criterion (BIC') (Lerner et al., 2018) across all of the traces (Table S3). Consistent with the interpretations made in earlier work on this riboswitch (Suddala et al., 2013, 2015), we found that the BIC indicated that all of the datasets tested are best modeled with two states, and in no case did either criterion suggest more than two states. For a number of our datasets, however, the conformational fluctuations identified in the kinetic model exhibited bi-exponential kinetics. This indicates that, although a 2-state model may be the most parsimonious, those states likely include subpopulations with slight differences in structure that affect their propensity to interconvert between the global docked and pre-docked conformations, again similar to our prior observations in the isolated riboswitch (Suddala et al., 2013, 2015). In the current work, we focus on the effects of the transcription machinery on the equilibrium partitioning of population between the global docked and pre-docked states.

Addition of the DNA Template Electrostatically Disfavors Riboswitch Folding

Consistent with previous studies of the isolated riboswitch aptamer (Suddala et al., 2013, 2015), we observed that under aptamer immobilization in the absence of Mg^{2+} or preQ$_1$, RNA0p existed almost exclusively in the pre-docked mid-FRET state with $E_{FRET} \sim 0.5$, while RNA3 also exhibited a minor fully docked high-FRET population (15%, $E_{FRET} = 0.88$) (Figure 2A). For both RNAs, addition of Mg^{2+} shifted the mid-FRET state to $E_{FRET} \sim 0.7$, indicative of a more compact pre-docked conformation, while increasing the population of the high-FRET state (Figure 3A). Addition of preQ$_1$ to either RNA (with or without Mg^{2+}) shifted additional population to the high-FRET state while only slightly changing the E_{FRET} of either state.

In contrast, most RNA10 molecules occupied a very high-FRET state ($E_{FRET} \sim 0.97$) regardless of Mg^{2+} or preQ$_1$ concentration (Figures 2A and 3A), consistent with an alternate structure. Specifically, additional nucleotides in RNA10 are predicted to form Watson-Crick base pairs with nucleotides that make up P2 and L2 in the native conformation (Figure S2B). To test this notion, we used mutagenesis to generate a sequence whose most stable predicted fold was once again the native structure. This mutant responded to ligand, confirming that when wild-type (WT) RNA10 is heat-annealed with CP, the native aptamer fold is disfavored.

To isolate the effects of the DNA template from those of RNAP, we next probed the riboswitch as part of a bubble complex that mimics the nucleic acid scaffold of an EC. We found that addition of the DNA bubble led to a shift of population from the high-FRET to the mid-FRET state and a shift of the mid-FRET state to lower E_{FRET} (~ 0.35 for RNA0p and RNA3, Figure 2B). Furthermore, the Mg^{2+}-dependent shift of the mid-FRET state to higher E_{FRET} was also greatly diminished (Figure 3B). RNA0p is rendered nearly insensitive to preQ$_1$ and Mg^{2+}, but sensitivity is restored as additional bases are inserted to form RNA3 and RNA10. The effects of the bubble are relieved by increasing the K$^+$ concentration (Figures S2C and S2D), indicating electrostatic repulsion by the DNA as the underlying cause for diminished pseudoknot docking. Bubble-immobilized RNA10 properly responds to Mg^{2+} and preQ$_1$, in contrast to its misfolding under aptamer immobilization. These results show that the DNA template may
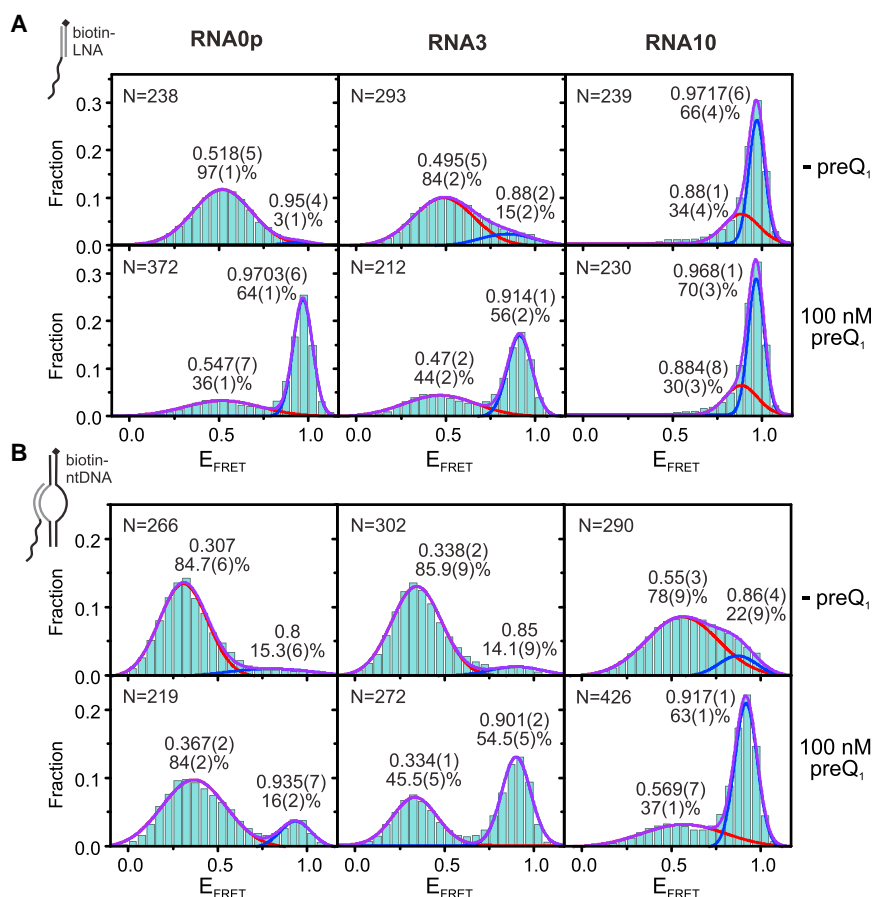
Figure 2. Riboswitch Folding under Aptamer and Bubble Immobilization in the Absence of Mg^{2+}

(A) smFRET histograms of RNA0p (left), RNA3 (middle), and RNA10 (right) under aptamer immobilization in the absence of MgCl$_2$ and the absence (top row) or presence (bottom row) of 100 nM preQ$_1$.

(B) Corresponding smFRET histograms under bubble immobilization.

Each histogram was fit with two Gaussians, and the resulting E_{FRET} and population of each state are noted adjacent to the corresponding peak. The standard deviation of the last digit of each parameter (obtained via bootstrapping) is reported in parentheses. Values for which no standard deviation is reported were constrained during fitting. The number of traces included in each histogram ("N") is noted in the upper left corner. See also Figure S2.

Compared to bubble-immobilized complexes at the same low K$^+$ concentration, the addition of RNAP shifted the mid-FRET state back to a higher E_{FRET} value ($\sim$0.53 for RNA0p) and shifted a significant fraction of population back to the high-FRET state (Figures 3C and 4). Similar to bubble complexes, the closer the aptamer to the EC, the less stabilized the high-FRET state, which is evidence of steric hindrance. Surprisingly, however, the shift toward higher E_{FRET} occurs even for ECs containing RNA0p, in which the 3′ segment of the P2 helix is expected to be within the RNAP exit channel (Kang et al., 2017).

To better rationalize these observations, we explored the steric accessibility of P2 folding using MD simulations with an all-heavy-atom structure-based model built on the cryo-EM structure of an *Eco* RNAP EC (Kang et al., 2017), together with crystal (Klein et al., 2009) and NMR (Kang et al., 2009) structures of the preQ$_1$ riboswitch (Figure 5). We found that while RNA0p is predicted to form fewer P2 base pairs than RNAs extended by 1–3 nucleotides (RNA1, RNA2, and RNA3), an average of 1–2 base pairs are intact (Table S4). In order for these base pairs to form, the 5′ segment of P2 must "bend back" toward RNAP. This bent-back pose places the 5′ leg of P2 near a cluster of basic residues in the β' subunit and the flap-tip (FT) of the β subunit (Figures 5 and S4), conserved regions in the vicinity of the RNA exit channel (Guo et al., 2018; Kang et al., 2018).

We probed for these predicted interactions through cross-linking experiments with 4-thio-UMP incorporated in or near P2. Indeed, when RNAP is stalled at the *que* pause site, both U10 and U14 efficiently cross-link to β', confirming that, as predicted, the 5′ segment of P2 bends back to interact with RNAP (Figures 5 and S4). We found cross-linking to be dependent on the stability and accessibility of P2, being more efficient in the presence of preQ$_1$ and less efficient in a G12-to-U mutant, in
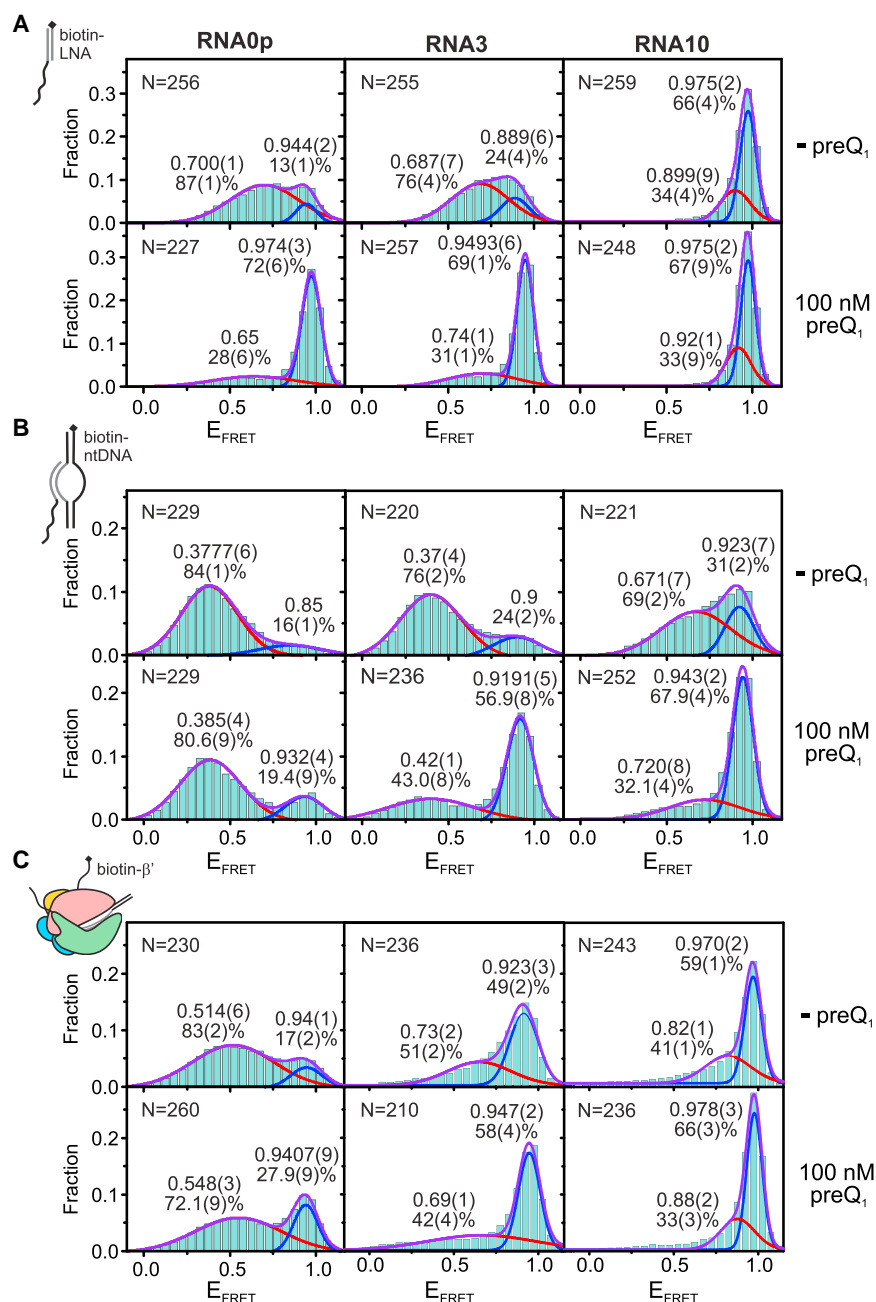
play a role in shaping the folding free-energy landscape of nascent RNAs when in direct proximity.

Addition of RNAP to the Bubble Scaffold Restores Riboswitch Folding

To probe the direct effects of RNAP on riboswitch folding, we assembled active ECs on a bubble scaffold identical to that described above. Immobilization of ECs through biotinylated RNAP proved to be critical to proper quantification of these effects, as experiments with the biotinylated ntDNA revealed that $\sim$50% of bubble scaffolds were not bound by RNAP (Figure 4A). Specifically, the FRET histogram of RNA3-containing ECs immobilized through biotinylated ntDNA can be fit as a linear combination of the bubble histogram and the RNAP-immobilized EC histogram, with the EC histogram contributing 51%. Consistent with this result, addition of GTP to ECs under identical conditions resulted in an elongation efficiency of 46% (Figure 4B). On-slide addition of the substrate nucleotide (GTP) to RNA0p-containing ECs caused the histogram of RNA0p to evolve toward that of RNA3, indicating that the ECs remain active on the slide (Figure 4C). Notably, obtaining ECs active under imaging conditions required optimization of the oxygen scavenging system (OSS) used to minimize photobleaching and blinking (Aitken et al., 2008) (Figure S3) to allow for stable binding of RNAP to the bubble scaffold and avoid nuclease degradation of the RNA.

PreQ$_1$ Releases a Pseudoknot-Stabilized Transcriptional Pause

Our results so far are consistent with a model in which the balance between electrostatic repulsion by the DNA template and stabilizing interactions with RNAP plays a significant role in preQ$_1$- and Mg^{2+}-dependent nascent RNA folding. Because RNAP pausing plays a critical role in guiding the co-transcriptional folding of numerous RNAs (Pan et al., 1999; Perdrizet et al., 2012), we next used single-round transcription assays to further probe the relationship between riboswitch folding and pausing. In addition to the *que* pause at U46, we observed a terminator that the riboswitch is expected to regulate (U70) and a second U-tract terminator further downstream (U108) (Figures 6 and S1). Nearly 100% of complexes paused at U46, with a pause half-life of 42 s in the absence of preQ$_1$ and a vast majority population (97%) with a significantly shortened half-

which P2 is weakened by disrupting its only G-C base pair. Interactions between the A-rich tail L3 and the minor groove of P1 may also contribute to stabilization of the cross-linking competent conformation (Feng et al., 2011; Klein et al., 2009). Together, our simulations and cross-linking results indicate that, at the *que* pause, residues near P2 interact with the β' subunit, potentially contributing to the stabilization of the docked conformation we observed through smFRET. Overall, our smFRET studies reveal that DNA and RNAP exert significant, opposing effects on nascent RNA folding, which are likely to play an important role in the regulation of transcription by other riboswitches.

life of 15 s in the presence of preQ$_1$. In addition, we found that the raw dataset for *Bsu* provided by Larson et al. (Larson et al., 2014) indicates that a transcriptional pause does, in fact, occur at this location *in vivo* (Figure S1E). Saturating concentrations of preQ$_1$ increased termination efficiency modestly from 27% to 33%. Both effects can be fit globally with an apparent preQ$_1$ binding affinity ($K_{1/2}$) of ~400 nM (Figure S5A), much weaker than the 5–10 nM for pseudoknot docking in the isolated aptamer, a reduction in apparent affinity also observed in other transcriptional riboswitches (Wickiser et al., 2005). The effect of preQ$_1$ on termination remains modest upon the addition of transcription factors, cell extract, crowding agents, additional riboswitch RNA,
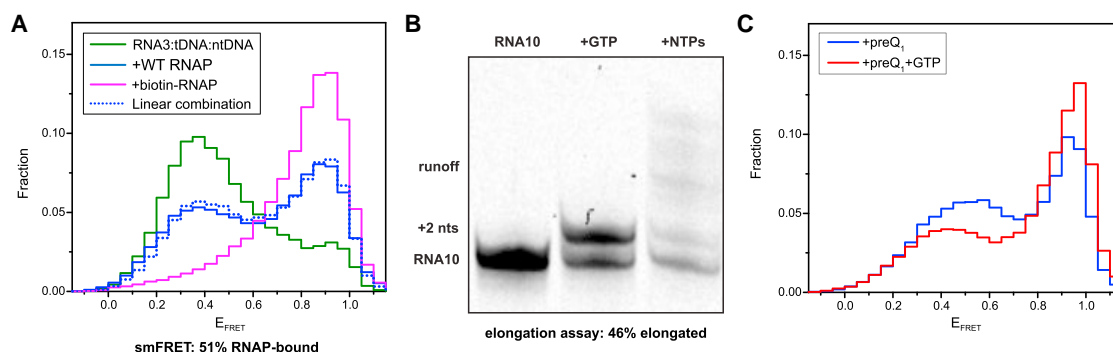
Figure 4. Assembly and Activity of Riboswitch-Containing ECs
(A) When RNA3-containing ECs are immobilized on a slide through biotinylated ntDNA, the FRET histogram (blue) corresponds to a linear combination (dashed blue) of the RNA3:tDNA:biotin-ntDNA histogram (green) and the RNA3:tDNA:ntDNA+biotin-RNAP histogram (purple), each contributing about 50%.
(B) A representative gel from a bubble-initiated elongation assay (the gel that was quantified is shown in Figure S3A).
(C) Incubation with GTP on the microscope slide causes the histogram of RNA0p ECs to evolve toward that of RNA3, indicating that nucleotides are being added.

and varied NTP concentrations (Figures S5B–S5D), suggesting that *in vivo* non-canonical factors may be involved in regulation of the downstream genes (Baird et al., 2010; Rinaldi et al., 2016).

The effect of preQ$_1$ and the similarity of the sequence around U46 to the pause consensus suggested that the *que* pause has both sequence and structural determinants, so to further uncover its origins, several RNA mutations were analyzed (Figures 6C–6E and S6). Altering the C18 base that forms a Watson-Crick pair with preQ$_1$ (Roth et al., 2007) abolished the impact of preQ$_1$ on pausing and termination, confirming that the riboswitch is responsible for the observed effects. A mutation that strengthens P2, stabilizing the docked state (C9U, which converts a non-canonical C-A base pair to U-A), doubled the pause half-life (to 81 s), while the P2-weakening G12U mutation, which we also utilized in the cross-linking experiments presented above, shortened the pause (to 26 s) and eliminated the effect of preQ$_1$. These aptamer mutations have smaller effects in the presence of preQ$_1$, suggesting that the details of sequence- and structure-dependent interactions with RNAP are more critical in the less-folded *apo* state. As expected, mutation of key bases in the pause consensus sequence (G37U and G36U/G37U) greatly reduced pause half-life in the absence of preQ$_1$ to 12 and 10 s, respectively. Addition of one U residue between the aptamer and the pause site (1U) also weakened pausing in the absence of preQ$_1$ (23 s), with little effect in the presence of preQ$_1$, and adding a second U (2U) had no significant further effect (22 s). The position of the pause is dictated by the consensus sequence rather than the riboswitch, as these insertions shift the pause site downstream by one and two nucleotides, respectively. These results indicate that the consensus sequence and strongly distance- and conformation-dependent interactions between the riboswitch and RNAP cooperate to stabilize the paused state, with ligand binding stabilizing the pseudoknot in a distinct way that facilitates pause release.

Interactions with the β Flap-Tip of RNAP Stabilize Pausing

We next tested the effects of RNAP-riboswitch contacts revealed by MD simulations. We first substituted basic β' residues (R77A, K79A, and R81A, Figure 5B) that lie near the stem of the

his pause hairpin in *E. coli* ECs (Guo et al., 2018; Kang et al., 2018) and have been implicated in the function of *put* antiterminator RNA structures (Sen et al., 2002). Arg77 is conserved between *Eco* and *Bsu* RNAPs, and Lys79 and Arg81 are also positively charged in *Bsu* RNAP. We found that individual point mutations led to modest decreases in pause half-life in the absence of preQ$_1$, whereas a triple mutant had the most significant effect (∼2-fold). In contrast, effects on pausing were negligible in the presence of preQ$_1$ (Figure S6B). This suggests that the pause is partially stabilized by electrostatic interactions between the exit channel and the pause-promoting, ligand-free docked conformation of the riboswitch. All variants showed decreased termination at U70, with the R77A and triple mutants showing the most extreme effects, though the termination-enhancing activity of preQ$_1$ was retained (Figures S5E and S5I).

We next tested the effect of the β FT, which is required for the pause-stabilizing effects of RNA duplexes (Hein et al., 2014). In contrast to the point mutants, which had effects only in the absence of preQ$_1$, deleting the FT by removing residues K890–K914 of β (ΔFT) decreased the pause half-life by about 30% in both the presence and absence of preQ$_1$ (Figure S6C). The effect of preQ$_1$-dependent conformational changes on pausing was therefore maintained for ΔFT RNAP, despite an overall decrease in pause half-life. In contrast, we observed no effect of the G12U mutation on pausing by ΔFT RNAP, indicating that pausing by ΔFT is not stabilized by the pseudoknot. Together, our experiments on RNAP variants suggest that both electrostatic interactions with exit channel residues and steric interactions with the FT contribute to pseudoknot-dependent pause stabilization in the absence of preQ$_1$. By contrast, in the preQ$_1$-bound docked state, only steric interactions with the FT contribute to pause *de*stabilization. These observations support the notion that the docked states in the presence and absence of preQ$_1$ are distinct and involve different interactions with RNAP.

DISCUSSION

We used smFRET, biochemical transcription assays, and MD simulations to study the co-transcriptional folding of the *Bsu*
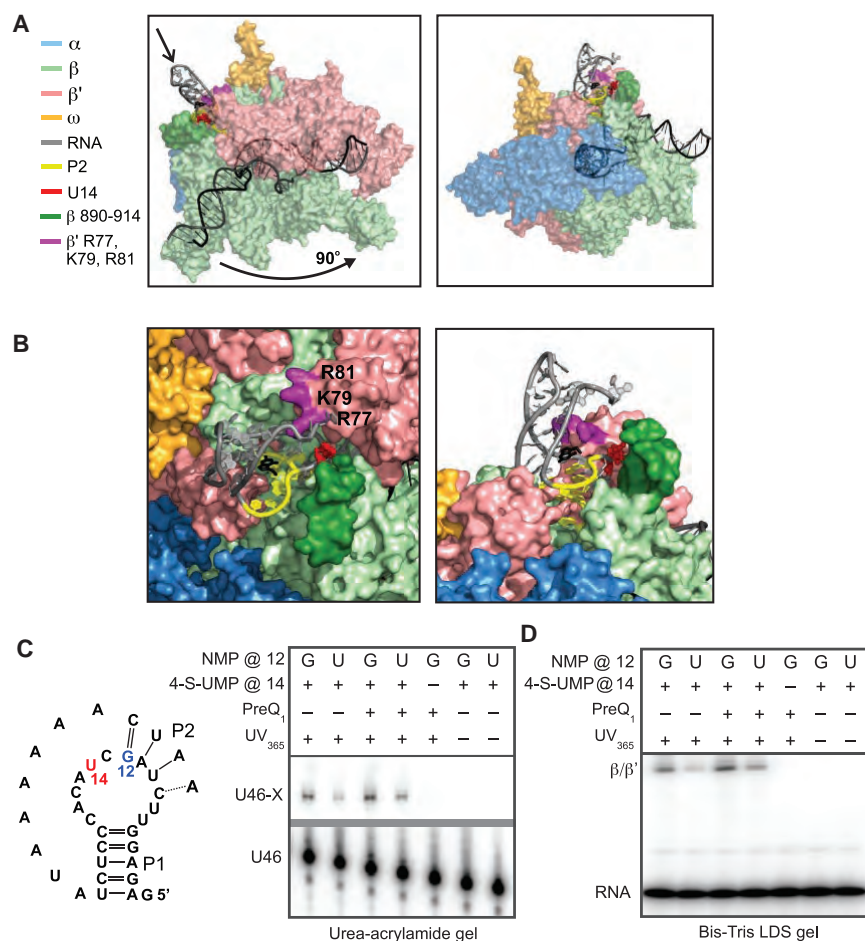
preQ$_1$ riboswitch. Isolated aptamers exhibited the previously observed pre-docked (mid-FRET) and docked (high-FRET) conformations, with the docked conformation being stabilized upon addition of preQ$_1$ and Mg^{2+} and the pre-docked conformation becoming more compact upon addition of Mg^{2+}. Similar Mg^{2+}-dependent changes in the absence of ligand have previously been observed by NMR in the isolated *Bsu* and *Fusobacterium nucleatum* (*Fnu*) preQ$_1$ riboswitches. In the *Bsu* riboswitch, addition of Mg^{2+} leads to changes in the ligand-binding pocket that make it resemble but not perfectly mimic the ligand-bound state (Suddala et al., 2013). In the *Fnu* riboswitch, addition of Mg^{2+} leads to signals suggestive of P2 base pairing and interactions between P1 and the 3' tail (Santner et al., 2012). These structural insights may explain some of the variation in E_{FRET} that we observe upon addition of Mg^{2+} and ligand. For example, we found that addition of preQ$_1$ shifts the E_{FRET} of the docked conformation to slightly higher values, a change that accompanies the concomitant rearrangement of the ligand-binding pocket. The significant changes in the E_{FRET} of the pre-docked conformation upon addition of Mg^{2+} are consistent with formation of the P1-3' tail interactions observed in both the *Bsu* (Suddala et al., 2013) and *Fnu* (Santner et al., 2012) preQ$_1$ riboswitches.

We also found that the DNA template disfavors nascent RNA folding in a distance-dependent manner. The E_{FRET} of the pre-docked conformation was significantly lowered and its dependence on Mg^{2+} weakened by addition of the DNA. We found that addition of KCl increased the E_{FRET} of the pre-docked conformation in a continuous fashion (Figures S3C and S3D), suggesting that there is no one distinct pre-docked conformation or that its dynamics remain unresolved. The addition of RNAP reverses this effect, restoring preQ$_1$-dependent docking and compaction of the pre-docked state despite steric hindrance of folding in the exit channel (Figure 3). Our observations provide concrete evidence for a previously hypothesized chaperone activity of RNAP, which has proven difficult to investigate by other means (Pan and Sosnick, 2006; Zhang and Landick, 2016).

Few studies exist in which naked RNA was directly compared to RNA in bubble complexes or ECs, with the only changes being the addition of those components of the transcription machinery. A study utilizing the fluorescent adenine analog 2-aminopurine showed that addition of the DNA template had little effect on base stacking in an RNA hairpin, while addition of bacteriophage T7 RNAP to the bubble decreased stacking significantly (Datta and von Hippel, 2008). A similar pattern was observed with single-stranded nascent RNA (Datta et al., 2006). Our investigation reveals very different behavior for a pseudoknot, with the DNA template having a significant effect and RNAP favoring RNA
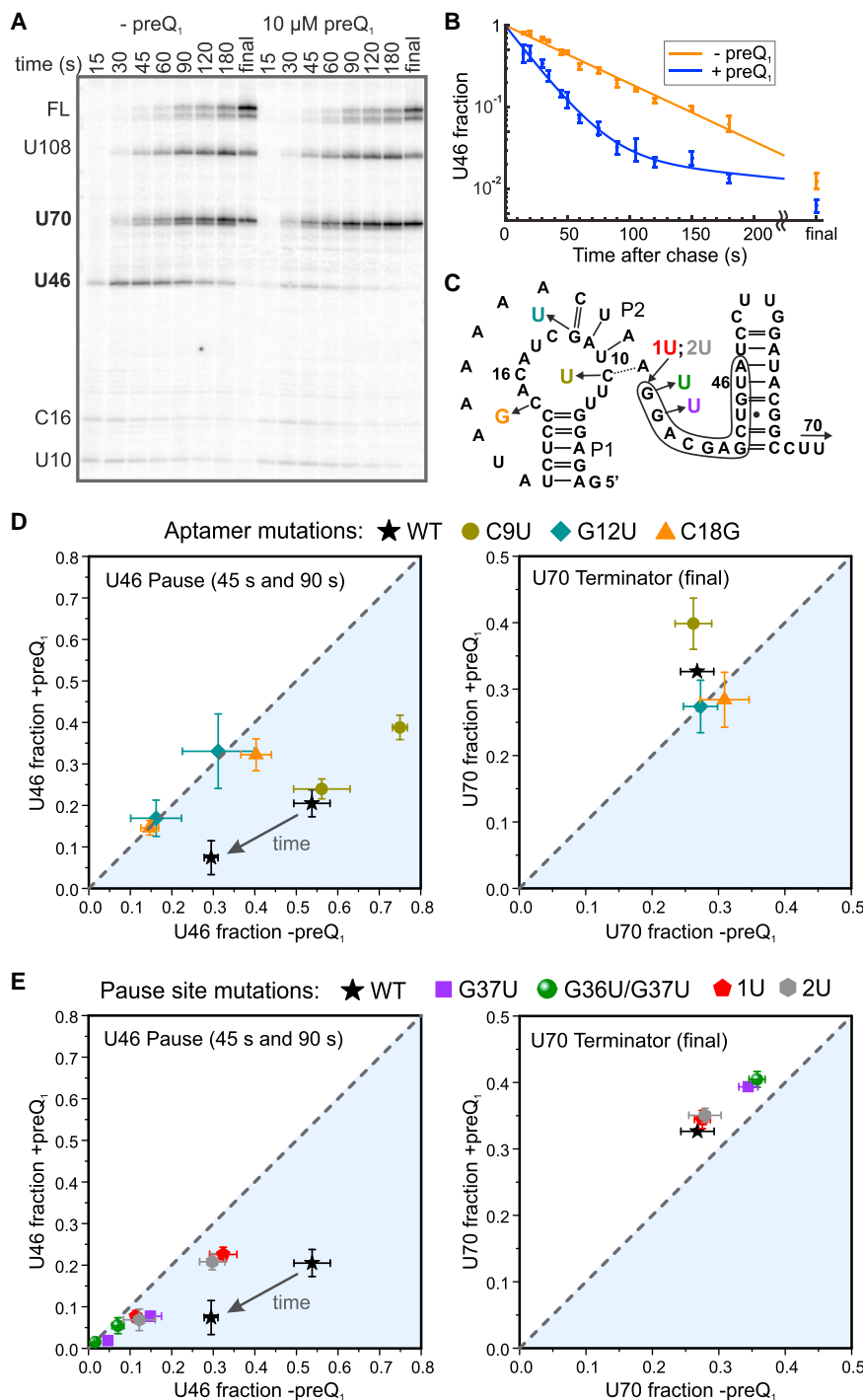
Figure 6. Effects of PreQ$_1$ on Transcriptional Pausing and Termination

(A) Representative denaturing urea-acrylamide gel close-up showing two weak pauses within the aptamer domain (U10 and C16), the *que* pause immediately following the aptamer (U46), the terminator regulated by the riboswitch (U70), a U-tract arrest site (U108), and full-length RNA (FL). Here and elsewhere, all lanes and data points marked "final" were recorded after a 3–5 min incubation with 200 μM NTPs.

(B) Fraction of complexes at the U46 site as a function of reaction time. A single-exponential fit is shown for the time course in the absence of preQ$_1$ (orange), and a double-exponential fit is shown for the time course in the presence of 10 μM preQ$_1$ (blue).

(C) Sequence of the aptamer and expression platform showing the final nucleotide added at the bands indicated in (A), as well as the color-coded locations of mutations reported in (D) and (E). U70 and U108 are beyond the region shown. The consensus-like pause sequence is outlined.

(D) Effects on pausing at U46 (left) and termination at U70 (right) of mutations within the aptamer domain. Data points above the dashed line indicate that addition of preQ$_1$ increased the population of a particular species, while points below the dashed line indicate that preQ$_1$ decreased the population of that species. PreQ$_1$ had no effect on data points that fall on the dashed line.

(E) Effects on pausing (left) and termination (right) of mutations at the pause site.

All error bars represent the standard deviation of three independent replicates. See also Figures S5 and S6 and Tables S5 and S6.

Conversely, using transcription assays, we showed that riboswitch folding modulates RNAP pausing (Figure 6). Current evidence points to the existence of an "elemental" paused state, in which sequences similar to the consensus trigger conformational changes that interfere with nucleotide addition. Elemental pauses have so far been found to be stabilized by two primary mechanisms (Artsimovitch and Landick, 2000; Zhang and Landick, 2016) (Figure 7A). In "hairpin-stabilized" or class I pausing, a hairpin in the nascent RNA causes conformational rearrangements in the exit channel that stabilize the pause by allosterically promoting a "swiveled" conformation of RNAP (Kang et al., 2018). Class I pauses are further stabilized by the transcription factor NusA through interactions with the nascent RNA and the β flap domain (Hein et al., 2014; Toulokhonov et al., 2001) and are resolved by the trailing ribosome (Landick et al., 1985). In a second mechanism (class II), the elemental pause enables backtracking by RNAP, leading to disengagement of the 3′ end of the RNA from the active site. Class II pauses

folding. Addition of *Eco* RNAP to a bubble scaffold decreased the rate of annealing of an antisense oligonucleotide (designed to mimic folding of the *his* pause hairpin) to the nascent RNA, but to a lesser extent than might be expected, indirectly implying a chaperoning function of RNAP (Hein et al., 2014). However, Hein et al. investigated a bimolecular annealing process, and their results therefore cannot be directly compared to our measurements of intramolecular pseudoknot folding.
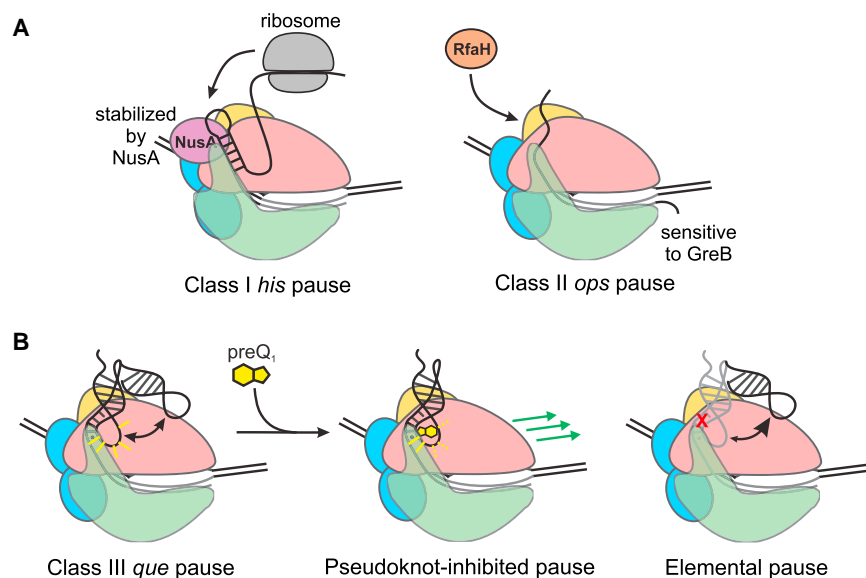
Figure 7. Mechanisms for Pause Stabilization and Control of Gene Expression

(A) At a class I pause, the trailing ribosome promotes RNAP escape and couples transcription to translation to regulate attenuation in amino acid biosynthesis operons. At a class II pause, RfaH binds to the EC and antiterminates transcription of long operons.

(B) At the class III *que* pause, the preQ$_1$ ligand chases RNAP from the pause and rearranges the riboswitch structure to promote termination, converting a pseudoknot-stabilized pause (left) into a pseudoknot-inhibited pause (middle). Disrupting the pseudoknot yields an elemental pause (right). See also Figure S7.

are sensitive to GreB, which enhances RNA cleavage by RNAP, and NusG, which antagonizes backtracking (Herbert et al., 2010). The class II *ops* pause provides the opportunity for the antitermination factor RfaH to bind, which temporarily stabilizes the pause before promoting elongation and suppressing pausing elsewhere (Artsimovitch and Landick, 2002).

The *que* pause we studied similarly requires an initial elemental pause, evidenced by the decrease in pause lifetime upon mutation of the consensus sequence (G37U and G36U/G37U mutants; Figures 6E and S7). The *que* pause is destabilized by moving the riboswitch further from RNAP (1U and 2U insertions), destabilizing P2 (G12U mutant), or deleting the β FT. We propose that these variants represent elemental or only weakly pseudoknot-stabilized pauses, lasting 22–25 s under our experimental conditions (Figure 7). In the WT and C9U variants, in contrast, the elemental pause is further stabilized by interactions between the pseudoknot and RNAP, resulting in a 2- to 4-fold increase in the pause half-life over that of the elemental pause.

These results suggest an analogy to the hairpin-stabilized class I pause, and the *que* pause described here indeed shares intriguing characteristics with other pauses that are dependent on nascent RNA structure. Notably, the first base pair of P2 is the same distance from the 3′ end of the nascent RNA as is the first base pair of the *his* pause hairpin (Toulokhonov et al., 2001). This is also identical to the distance at which initial folding of a pseudoknot was observed in the fluoride riboswitch (Watters et al., 2016). Differences in RNAP translocation register yield a 1- to 2-nucleotide variation in the position of the paired base relative to RNAP; nevertheless, these similarities indicate a common length-scale for nucleation of RNA folding. Insertion of two nucleotides between the hairpin and the *his* pause site reduces the pause lifetime and eliminates the ability of an oligonucleotide that disrupts the hairpin to inhibit pausing (Toulokhonov et al., 2001), suggesting that the hairpin no longer affects pausing at this distance. Our 1U and 2U variants exhibit the same effect, with potential modulation of nascent RNA structure through

preQ$_1$ binding having little effect on pausing. Similar to our results, it has previously been found that deletion of the β FT has no effect on the duration of an elemental pause, while it destabilizes a *his* pause-complex mimic in which the hairpin is replaced with a duplex (Hein et al., 2014). The *que* pause therefore shares certain characteristics with the class I *his* pause despite being stabilized by a pseudoknot rather than a hairpin.

However, there are several significant differences between the properties of the *que* pause and canonical class I and class II pauses. Notably, *Bsu* and *Eco* RNAPs both respond to the *que* pause signal, whereas *Bsu* RNAP does not recognize the *his* pause at all (Artsimovitch et al., 2000). The FT has a much larger effect on class I pauses than on the *que* pause, and the presence of a pseudoknot rather than a hairpin also renders the *que* pause insensitive to NusA (Figures S5G and S6D), which requires either a simple duplex of at least 10 base pairs (Kolb et al., 2014) or a loop of greater than 4 nucleotides (Toulokhonov et al., 2001) for maximum stimulation of pausing. The folded riboswitch instead presents a 4-base-pair helix at the same position as the *his* pause hairpin and, instead of a loop, presents a triplex-like structure in which L3 forms Hoogsteen-face interactions with the helix P1 (Klein et al., 2009). PreQ$_1$ binding has the effect of chasing RNAP away from the *que* pause, a function performed by the ribosome in the case of class I pauses (Landick et al., 1985). Unlike class II pauses, the RNA at the *que* pause does not undergo cleavage in response to GreB (Figure S5H), indicating that RNAP is not in a backtracked position. Our results suggest that the *que* pause represents a new mechanism in which an elemental pause is stabilized by a pseudoknot in the nascent RNA next to the RNAP exit channel, which we term "pseudoknot-stabilized" or "class III" pausing (Figure 7B).

Our cross-linking results would seem to indicate that addition of preQ$_1$ strengthens interactions with RNAP (Figure 5), and on the basis of our transcription assays in the absence of preQ$_1$, this would be expected to promote pausing. However, in our cross-linking experiments, pause release is impossible due to the absence of NTPs. In this situation, preQ$_1$ has the sole effect of stabilizing a bent-back, cross-linking-competent conformation in which P2 is intact. When the presence of

NTPs makes pause release possible, this conformation, which clearly differs from the pause-promoting bent-back conformation, appears to instead yield a ''pseudoknot-inhibited'' analog of the hairpin-inhibited pause escape mechanism (Zhang and Landick, 2016). The G12U mutant does not transition into a pseudoknot-inhibited pause because it binds preQ$_1$ too weakly, leading to minimal effects of preQ$_1$ on pausing and termination. The 1U and 2U variants do not transition into a pseudoknot-inhibited state for one or both of two reasons: (1) the increased distance between the pseudoknot and RNAP means that RNAP does not ''feel'' the conformational change upon preQ$_1$ binding, and/or (2) preQ$_1$ binding does not cause an extensive structural rearrangement, as seen in our MD and smFRET results.

In summary, we have demonstrated an intimate cross-coupling between bacterial riboswitch folding and transcriptional pausing that is likely to govern the co-transcriptional folding of numerous RNAs. Using a combination of approaches that are readily applicable to a wide variety of RNA species, we found that the presence of stalled RNAP provides a favorable free energy landscape for folding of the nascent preQ$_1$ riboswitch. In turn, the ligand-free but pre-folded riboswitch pseudoknot stabilizes the paused state of RNAP via interactions with the β FT and with positively charged residues on the β' subunit, while preQ$_1$ binding stabilizes a distinct docked conformation that counteracts pausing. As pseudoknots are a common RNA structural motif, we expect that numerous examples of class III pausing will be found among the thousands of transcriptional pauses that have been identified (Larson et al., 2014; Vvedenskaya et al., 2014). Our results reveal an additional, so far underappreciated layer to bacterial gene regulation that can be exploited for antibiotic drug intervention (Blount and Breaker, 2006; Deigan and Ferré-D'Amaré, 2011; Howe et al., 2015).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Bacterial strains, Plasmids and Growth Conditions
- METHOD DETAILS
  - Single-molecule FRET
  - Transcription Assays
  - Simulations
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and six tables and can be found with this article online at https://doi.org/10.1016/j.molcel.2018.08.046.

AUTHOR CONTRIBUTIONS

Conceptualization, J.R.W. and N.G.W.; Methodology, J.R.W., Y.A.N., R.L.H.; Investigation, J.R.W., Y.A.N., V.R., R.L.H., I.A.; Writing - Original Draft, J.R.W.; Writing - Review & Editing, J.R.W., Y.A.N., V.R., R.L.H., C.L.B., I.A., N.G.W.; Supervision, N.G.W., I.A., C.L.B.; Funding Acquisition, N.G.W., I.A., J.R.W.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Aitken, C.E., Marshall, R.A., and Puglisi, J.D. (2008). An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. Biophys. J. *94*, 1826–1835.

Artsimovitch, I., and Landick, R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. Proc. Natl. Acad. Sci. USA *97*, 7090–7095.

Artsimovitch, I., and Landick, R. (2002). The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. Cell *109*, 193–203.

Artsimovitch, I., Svetlov, V., Anthony, L., Burgess, R.R., and Landick, R. (2000). RNA polymerases from Bacillus subtilis and Escherichia coli differ in recognition of regulatory signals in vitro. J. Bacteriol. *182*, 6027–6035.

Baird, N.J., Kulshina, N., and Ferré-D'Amaré, A.R. (2010). Riboswitch function: flipping the switch or tuning the dimmer? RNA Biol. *7*, 328–332.

Blanco, M., and Walter, N.G. (2010). Analysis of complex single-molecule FRET time trajectories. Methods Enzymol. *472*, 153–178.

Blount, K.F., and Breaker, R.R. (2006). Riboswitches as antibacterial drug targets. Nat. Biotechnol. *24*, 1558–1564.

Daher, M., Widom, J.R., Tay, W., and Walter, N.G. (2018). Soft Interactions with Model Crowders and Non-canonical Interactions with Cellular Proteins Stabilize RNA Folding. J. Mol. Biol. *430*, 509–523.

Datta, K., and von Hippel, P.H. (2008). Direct spectroscopic study of reconstituted transcription complexes reveals that intrinsic termination is driven primarily by thermodynamic destabilization of the nucleic acid framework. J. Biol. Chem. *283*, 3537–3549.

Datta, K., Johnson, N.P., and von Hippel, P.H. (2006). Mapping the conformation of the nucleic acid framework of the T7 RNA polymerase elongation complex in solution using low-energy CD and fluorescence spectroscopy. J. Mol. Biol. *360*, 800–813.

Daube, S.S., and von Hippel, P.H. (1992). Functional transcription elongation complexes from synthetic RNA-DNA bubble duplexes. Science *258*, 1320–1324.

Deigan, K.E., and Ferré-D'Amaré, A.R. (2011). Riboswitches: discovery of drugs that target bacterial gene-regulatory RNAs. Acc. Chem. Res. *44*, 1329–1338.

Duesterberg, V.K., Fischer-Hwang, I.T., Perez, C.F., Hogan, D.W., and Block, S.M. (2015). Observation of long-range tertiary interactions during ligand binding by the TPP riboswitch aptamer. eLife *4*, e12362.

Dupuis, N.F., Holmstrom, E.D., and Nesbitt, D.J. (2014). Molecular-crowding effects on single-molecule RNA folding/unfolding thermodynamics and kinetics. Proc. Natl. Acad. Sci. USA *111*, 8464–8469.

Elgaher, W.A.M., Fruth, M., Groh, M., Haupenthal, J., and Hartmann, R.W. (2014). Expanding the scaffold for bacterial RNA polymerase inhibitors: design, synthesis and structure–activity relationships of ureido-heterocyclic-carboxylic acids. RSC Advances *4*, 2177–2194.

Feng, J., Walter, N.G., and Brooks, C.L., 3rd (2011). Cooperative and directional folding of the preQ$_1$ riboswitch aptamer domain. J. Am. Chem. Soc. *133*, 4196–4199.

Frieda, K.L., and Block, S.M. (2012). Direct observation of cotranscriptional folding in an adenine riboswitch. Science *338*, 397–400.

Greenlee, E.B., Stav, S., Atilho, R.M., Brewer, K.I., Harris, K.A., Malkowski, S.N., Mirihana Arachchilage, G., Perkins, K.R., Sherlock, M.E., and Breaker, R.R. (2018). Challenges of ligand identification for the second wave of orphan riboswitch candidates. RNA Biol. *15*, 377–390.

Grundy, F.J., and Henkin, T.M. (2004). Kinetic analysis of tRNA-directed transcription antitermination of the Bacillus subtilis glyQS gene in vitro. J. Bacteriol. *186*, 5392–5399.

Guo, X., Myasnikov, A.G., Chen, J., Crucifix, C., Papai, G., Takacs, M., Schultz, P., and Weixlbaumer, A. (2018). Structural Basis for NusA Stabilized Transcriptional Pausing. Mol. Cell *69*, 816–827.e4.

Hayes, R.L., Noel, J.K., Whitford, P.C., Mohanty, U., Sanbonmatsu, K.Y., and Onuchic, J.N. (2014). Reduced model captures Mg(2+)-RNA interaction free energy of riboswitches. Biophys. J. *106*, 1508–1519.

Hein, P.P., Kolb, K.E., Windgassen, T., Bellecourt, M.J., Darst, S.A., Mooney, R.A., and Landick, R. (2014). RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. Nat. Struct. Mol. Biol. *21*, 794–802.

Helmling, C., Klötzner, D.-P., Sochor, F., Mooney, R.A., Wacker, A., Landick, R., Fürtig, B., Heckel, A., and Schwalbe, H. (2018). Life times of metastable states guide regulatory signaling in transcriptional riboswitches. Nat. Commun. *9*, 944.

Herbert, K.M., Zhou, J., Mooney, R.A., Porta, A.L., Landick, R., and Block, S.M. (2010). E. coli NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. J. Mol. Biol. *399*, 17–30.

Holmstrom, E.D., Polaski, J.T., Batey, R.T., and Nesbitt, D.J. (2014). Single-molecule conformational dynamics of a biologically functional hydroxocobalamin riboswitch. J. Am. Chem. Soc. *136*, 16832–16843.

Hoopes, B.C., and McClure, W.R. (1981). Studies on the selectivity of DNA precipitation by spermine. Nucleic Acids Res. *9*, 5493–5504.

Howe, J.A., Wang, H., Fischmann, T.O., Balibar, C.J., Xiao, L., Galgoci, A.M., Malinverni, J.C., Mayhood, T., Villafania, A., Nahvi, A., et al. (2015). Selective small-molecule inhibition of an RNA structural element. Nature *526*, 672–677.

Kang, M., Peterson, R., and Feigon, J. (2009). Structural Insights into riboswitch control of the biosynthesis of queuosine, a modified nucleotide found in the anticodon of tRNA. Mol. Cell *33*, 784–790.

Kang, J.Y., Olinares, P.D.B., Chen, J., Campbell, E.A., Mustaev, A., Chait, B.T., Gottesman, M.E., and Darst, S.A. (2017). Structural basis of transcription arrest by coliphage HK022 Nun in an *Escherichia coli* RNA polymerase elongation complex. eLife *6*, e25478.

Kang, J.Y., Mishanina, T.V., Bellecourt, M.J., Mooney, R.A., Darst, S.A., and Landick, R. (2018). RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. Mol. Cell *69*, 802–815.e1.

Kay, B.K., Thai, S., and Volgina, V.V. (2009). High-throughput biotinylation of proteins. Methods Mol. Biol. *498*, 185–196.

Klein, D.J., Edwards, T.E., and Ferré-D'Amaré, A.R. (2009). Cocrystal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. Nat. Struct. Mol. Biol. *16*, 343–344.

Kolb, K.E., Hein, P.P., and Landick, R. (2014). 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.

Kumar, S., Bouzida, D., Swendson, R., Kollman, P.A., and Rosenberg, J.M. (1992). The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. J. Comput. Chem. *13*, 1011–1021.

Landick, R., Carey, J., and Yanofsky, C. (1985). Translation activates the paused transcription complex and restores transcription of the *trp* operon leader region. Proc. Natl. Acad. Sci. USA *82*, 4663–4667.

Larson, M.H., Mooney, R.A., Peters, J.M., Windgassen, T., Nayak, D., Gross, C.A., Block, S.M., Greenleaf, W.J., Landick, R., and Weissman, J.S. (2014). A pause sequence enriched at translation start sites drives transcription dynamics in vivo. Science *344*, 1042–1047.

Lemay, J.-F., Desnoyers, G., Blouin, S., Heppell, B., Bastet, L., St-Pierre, P., Massé, E., and Lafontaine, D.A. (2011). Comparative study between transcriptionally- and translationally-acting adenine riboswitches reveals key differences in riboswitch regulatory mechanisms. PLoS Genet. *7*, e1001278.

Lerner, E., Ingargiola, A., and Weiss, S. (2018). Characterizing highly dynamic conformational states: The transcription bubble in RNAP-promoter open complex as an example. J. Chem. Phys. *148*, 123315.

Liberman, J.A., Suddala, K.C., Aytenfisu, A., Chan, D., Belashov, I.A., Salim, M., Mathews, D.H., Spitale, R.C., Walter, N.G., and Wedekind, J.E. (2015). Structural analysis of a class III preQ$_1$ riboswitch reveals an aptamer distant from a ribosome-binding site regulated by fast dynamics. Proc. Natl. Acad. Sci. USA *112*, E3485–E3494.

Lutz, B., Faber, M., Verma, A., Klumpp, S., and Schug, A. (2014). Differences between cotranscriptional and free riboswitch folding. Nucleic Acids Res. *42*, 2687–2696.

Noel, J.K., Whitford, P.C., Sanbonmatsu, K.Y., and Onuchic, J.N. (2010). SMOG@ctbp: simplified deployment of structure-based models in GROMACS. Nucleic Acids Res. *38*, W657-61.

Noel, J.K., Whitford, P.C., and Onuchic, J.N. (2012). The shadow map: a general contact definition for capturing the dynamics of biomolecular folding and function. J. Phys. Chem. B *116*, 8692–8702.

Noel, J.K., Levi, M., Raghunathan, M., Lammert, H., Hayes, R.L., Onuchic, J.N., and Whitford, P.C. (2016). SMOG 2: A Versatile Software Package for Generating Structure-Based Models. PLoS Comput. Biol. *12*, e1004794.

Onuchic, J.N., and Wolynes, P.G. (2004). Theory of protein folding. Curr. Opin. Struct. Biol. *14*, 70–75.

Pan, T., and Sosnick, T. (2006). RNA folding during transcription. Annu. Rev. Biophys. Biomol. Struct. *35*, 161–175.

Pan, T., Artsimovitch, I., Fang, X.-W., Landick, R., and Sosnick, T.R. (1999). Folding of a large ribozyme during transcription and the effect of the elongation factor NusA. Proc. Natl. Acad. Sci. USA *96*, 9545–9550.

Perdrizet, G.A., 2nd, Artsimovitch, I., Furman, R., Sosnick, T.R., and Pan, T. (2012). Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. Proc. Natl. Acad. Sci. USA *109*, 3323–3328.

Rinaldi, A.J., Lund, P.E., Blanco, M.R., and Walter, N.G. (2016). The Shine-Dalgarno sequence of riboswitch-regulated single mRNAs shows ligand-dependent accessibility bursts. Nat. Commun. *7*, 8976.

Roth, A., Winkler, W.C., Regulski, E.E., Lee, B.W.K., Lim, J., Jona, I., Barrick, J.E., Ritwik, A., Kim, J.N., Welz, R., et al. (2007). A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. Nat. Struct. Mol. Biol. *14*, 308–317.

Santner, T., Rieder, U., Kreutz, C., and Micura, R. (2012). Pseudoknot preorganization of the preQ1 class I riboswitch. J. Am. Chem. Soc. *134*, 11928–11931.

Sen, R., King, R.A., Mzhavia, N., Madsen, P.L., and Weisberg, R.A. (2002). Sequence-specific interaction of nascent antiterminator RNA with the zinc-finger motif of *Escherichia coli* RNA polymerase. Mol. Microbiol. *46*, 215–222.

Senavirathne, G., Liu, J., Lopez, M.A., Jr., Hanne, J., Martin-Lopez, J., Lee, J.-B., Yoder, K.E., and Fishel, R. (2015). Widespread nuclease contamination in commonly used oxygen-scavenging systems. Nat. Methods *12*, 901–902.

Shea, J.-E., Nochomovitz, Y.D., Guo, Z., and Brooks, C.L., III (1998). Exploring the space of protein folding Hamiltonians: The balance of forces in a minimalist β-barrel model. J. Chem. Phys. *109*, 2895–2903.

Sherwood, A.V., and Henkin, T.M. (2016). Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. Annu. Rev. Microbiol. *70*, 361–374.

Sidorenkov, I., Komissarova, N., and Kashlev, M. (1998). Crucial role of the RNA:DNA hybrid in the processivity of transcription. Mol. Cell *2*, 55–64.

Steinert, H., Sochor, F., Wacker, A., Buck, J., Helmling, C., Hiller, F., Keyhani, S., Noeske, J., Grimm, S., Rudolph, M.M., et al. (2017). Pausing guides RNA folding to populate transiently stable RNA structures for riboswitch-based transcription regulation. eLife *6*, e21297.

Suddala, K.C., Rinaldi, A.J., Feng, J., Mustoe, A.M., Eichhorn, C.D., Liberman, J.A., Wedekind, J.E., Al-Hashimi, H.M., Brooks, C.L., 3rd, and Walter, N.G. (2013). 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.

Suddala, K.C., Wang, J., Hou, Q., and Walter, N.G. (2015). Mg(2+) shifts ligand-mediated folding of a riboswitch from induced-fit to conformational selection. J. Am. Chem. Soc. *137*, 14075–14083.

Svetlov, V., and Artsimovitch, I. (2015). Purification of bacterial RNA polymerase: tools and protocols. Methods Mol. Biol. *1276*, 13–29.

Toulokhonov, I., Artsimovitch, I., and Landick, R. (2001). Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. Science *292*, 730–733.

Uhm, H., Kang, W., Ha, K.S., Kang, C., and Hohng, S. (2018). Single-molecule FRET studies on the cotranscriptional folding of a thiamine pyrophosphate riboswitch. Proc. Natl. Acad. Sci. USA *115*, 331–336.

Vvedenskaya, I.O., Vahedian-Movahed, H., Bird, J.G., Knoblauch, J.G., Goldman, S.R., Zhang, Y., Ebright, R.H., and Nickels, B.E. (2014). Interactions between RNA polymerase and the "core recognition element" counteract pausing. Science *344*, 1285–1289.

Watters, K.E., Strobel, E.J., Yu, A.M., Lis, J.T., and Lucks, J.B. (2016). Cotranscriptional folding of a riboswitch at nucleotide resolution. Nat. Struct. Mol. Biol. *23*, 1124–1131.

Whitford, P.C., Schug, A., Saunders, J., Hennelly, S.P., Onuchic, J.N., and Sanbonmatsu, K.Y. (2009). Nonlocal helix formation is key to understanding S-adenosylmethionine-1 riboswitch function. Biophys. J. *96*, L7–L9.

Whitford, P.C., Geggier, P., Altman, R.B., Blanchard, S.C., Onuchic, J.N., and Sanbonmatsu, K.Y. (2010). Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. RNA *16*, 1196–1204.

Wickiser, J.K., Winkler, W.C., Breaker, R.R., and Crothers, D.M. (2005). The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. Mol. Cell *18*, 49–60.

Zhang, J., and Ferré-D'Amaré, A.R. (2013). Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. Nature *500*, 363–366.

Zhang, J., and Landick, R. (2016). A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure. Trends Biochem. Sci. *41*, 293–310.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
E. coli: JM109 competent cells	Promega	Cat#L2001
E. coli: BLR (DE3) competent cells	EMD Millipore	Cat#69053
E. coli: XJb (DE3) competent cells	Zymo Research	Cat#T3051
E. coli: XL10 gold ultracompetent cells	Stratagene	Cat#200314
Chemicals, Peptides, and Recombinant Proteins		
Cy5 mono-reactive dye pack	GE Healthcare Life Sciences	Cat#PA25001
T4 RNA ligase 2	New England Biolabs	Cat#M0239S
E. coli RNA polymerase, core enzyme	New England Biolabs	Cat#M0550S
E. coli RNA polymerase, holoenzyme	New England Biolabs	Cat#M0551S
E. coli RNA polymerase, core enzyme with biotinylation tag	This study	N/A
E. coli RNA polymerase, holoenzyme with R77A mutation on β' subunit	This study	N/A
E. coli RNA polymerase, holoenzyme with K79A mutation on β' subunit	This study	N/A
E. coli RNA polymerase, holoenzyme with R81A mutation on β' subunit	This study	N/A
E. coli RNA polymerase, holoenzyme with R77A/K79A/R81A triple mutation on β' subunit	This study	N/A
E. coli RNA polymerase, holoenzyme with β flap tip deleted	Irina Artsimovitch	N/A
ApC RNA dinucleotide	TriLink Biotechnologies	Cat#O-31002
High purity rNTP set	GE Healthcare Life Sciences	Cat#27-2025-01
α-^{32}P GTP, 3000 Ci/mmol	Perkin Elmer	Cat#BLU006H250UC
Trolox	Fisher Scientific	Cat#218940050
Protocatechuic acid	Soltec Ventures	Cat#CL107
Protocatechuate 3,4-dioxygenase	Sigma Aldrich	Cat#P8279
Glucose	Sigma Aldrich	Cat#G8270
Glucose oxidase from *Aspergillus Niger*	Sigma Aldrich	Cat#G7141
Catalase from bovine liver	Sigma Aldrich	Cat#C40
Catalase from *Corynebacterium glutamicum*	Sigma Aldrich	Cat#02071
Deposited Data		
NET-seq data from *B. subtilis*	Gene Expression Omnibus	GEO: GSE56720
Oligonucleotides		
Sequences provided in Tables S1 and S2	Dharmacon/Integrated DNA Technologies	N/A
Recombinant DNA		
pVR: *E. coli* RecA promoter followed by 5′UTR of QueC gene from *B. subtilis*	This study	N/A
pVR2: pVR but with a 12-nucleotide stretch lacking A residues inserted after the promoter	This study	N/A
pVR2-m1: C9U mutant	This study	N/A
pVR2-m2: C12U mutant	This study	N/A
pVR2-m3: C18G mutant	This study	N/A
pVR2-m5: 1U mutant	This study	N/A
pVR2-m6: 2U mutant	This study	N/A
pVR2-m7: G37U mutant	This study	N/A
pVR2-m8: G36U/G37U mutant	This study	N/A
pIA999: overexpression plasmid for *E. coli* RNAP with biotinylation tag	Irina Artsimovitch	N/A
pIA1291: overexpression plasmid for *E. coli* RNAP with β' R77A mutation	This study	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pIA1292: overexpression plasmid for *E. coli* RNAP with β' K79A mutation	This study	N/A
pIA1293: overexpression plasmid for *E. coli* RNAP with β' R81A mutation	This study	N/A
pIA1294: overexpression plasmid for *E. coli* RNAP with β' R77A/K79A/R81A triple mutation	This study	N/A
Software and Algorithms		
ImageQuant	Molecular Dynamics	N/A
Custom smFRET analysis code in MATLAB and IDL	This study	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Nils Walter (nwalter@umich.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial strains, Plasmids and Growth Conditions

E. coli JM109 competent cells were used for maintenance of all transcription template plasmids. *E. coli* XL10 gold ultracompetent cells were used for maintenance of biotinylated RNAP expression plasmids. *E. coli* BLR (DE3) competent cells were used for over-expression of biotinylation-tagged RNAP, and XJb (DE3) cells were used for overexpression of RNAP variants. All strains were grown at 37°C in Luria-Bertani media or on Luria-Bertani media-agar plates supplemented with 100 μg/mL carbenicillin. Plasmids used in this study are listed in the Key Resources Table.

METHOD DETAILS

Single-molecule FRET
Design and preparation of RNA constructs
It is widely known that transcription initiation can be bypassed by incubating RNAP with a pre-assembled scaffold consisting of a partially complementary DNA bubble and an RNA primer (Daube and von Hippel, 1992; Hein et al., 2014; Kolb et al., 2014). Our DNA bubble was designed based on a sequence from Kolb et al., 2014, with minor changes so that it was predicted to not interfere with formation of the native riboswitch structure (Table S1). "EC RNAs" 0, 3 and 10 were then designed to contain 10 nucleotides that were complementary to the template strand of the bubble, as well as the desired number of bases from the expression platform. The sequence of the capture probe (CP) was then designed to sequester in a duplex the same nucleotides that form the RNA:DNA hybrid in the bubble complex, and LNA bases were included to achieve a T_m of ~80°C. The RNA:LNA or RNA:DNA hybrid sequence was not based on the expression platform, and was identical in the three RNA constructs with the exception of two bases that had to be changed to accommodate the portion of the aptamer oligonucleotide that became part of the hybrid in the shortest construct, RNA0p. The riboswitch aptamer sequence included a G-to-C mutation at the second position to suppress formation of alternate structures containing helix P1a (Kang et al., 2009). Aptamer labeling locations were chosen to match those used in previous work on the isolated aptamer (Suddala et al., 2013, 2015). Ligation enabled smFRET constructs of multiple lengths to be prepared from one aptamer construct.

All oligonucleotides used in smFRET experiments were purchased HPLC-purified from Dharmacon (in the case of fluorophore-labeled RNA) or Integrated DNA Technologies (in the case of DNA and unlabeled RNA). The riboswitch aptamer construct contained a 2′ Dy547 label at G37 and a 5-aminoallyl-uridine (5-N-U) at position 14. To label the 5-N-U, 4 nmol RNA was incubated for 4 hr at room temperature in a 50 μL reaction containing 30 μL of DMSO, one Cy5 mono-reactive dye pack (GE Healthcare Life Sciences) and 0.1 M sodium bicarbonate (pH 9). The RNA was isolated by ethanol precipitation. The labeled aptamer (200 pmol) and 5′ phosphorylated EC RNA (400 pmol) were annealed with a DNA or LNA splint (400 pmol-1 nmol) in a 100 μL mixture containing 150 mM NaCl, 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA. This annealing mixture was diluted to 500 μL 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 2 hr. The reaction mixture was extracted with an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol, then concentrated down to ~20 μL in a vacuum centrifuge at 45°C and purified by denaturing, 8 M urea, 8% polyacrylamide gel electrophoresis (PAGE). The product band was extracted and the RNA isolated by ethanol precipitation.

Protein preparation
E. coli core RNAP bearing an "AviTag" biotinylation tag (GLNDIFEAQKIEWH) on the C terminus of the β' subunit was expressed from plasmid pIA999 in BLR(DE3) cells (Kay et al., 2009). The cells were grown in LB media supplemented with 20 μM biotin and 100 μg/mL

carbenicillin, induced with 1 mM IPTG at an OD600 of 0.7 and harvested 3 hr post-induction. The protein was purified essentially as described (Svetlov and Artsimovitch, 2015). Briefly, the *his*-tagged protein was purified on a nickel-NTA agarose column, followed by a heparin column, and finally by a Mono Q ion-exchange column. The purified protein was dialyzed into storage buffer (10 mM Tris-HCl, pH 7.5, 50% glycerol, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT) and stored at −80°C. Biotinylation of the β' subunit by endogenous BirA was verified by incubating biotin-tagged and wild-type proteins with streptavidin, followed by native PAGE analysis. Further *in vitro* biotinylation was not performed. RNAP mutants were prepared by the same protocol, except that a Resource Q ion-exchange column was used instead of Mono Q.

smFRET experiments

The RNA of interest was combined at a final concentration of 0.5 μM with either capture probe (CP) or template DNA (tDNA) and non-template DNA (ntDNA) in buffer A (50 mM Tris-HCl, pH 7.5, 100 mM KCl). When CP was used, its concentration was 5 μM; DNA concentrations were 10 μM when immobilizing through 3'-biotinylated ntDNA and 1 μM when immobilizing through biotinylated RNAP. The mixture was annealed by incubating at 90°C for 2 min, then 37°C for 10 min, then RT for 10 min. For experiments in the absence of RNAP, the annealing mixture was diluted in buffer A to a concentration of 15-50 pM RNA and flowed onto a slide that 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 performed in the presence of RNAP, the annealed nucleic acid scaffold was diluted to 50 nM RNA in buffer B (50 mM Tris-HCl, pH 7.5, 100 mM KCl, 1 mM MgCl$_2$), and E. coli RNAP was added to 250 nM. This mixture was incubated for 15 min at 37°C, then diluted in buffer B to 50 pM RNA for immobilization. Imaging was performed at 62 ms time resolution in buffer A or B (where indicated, supplemented with preQ$_1$ or additional KCl), containing 44 mM glucose, 165 U/mL glucose oxidase from Aspergillus niger, 2170 U/mL catalase from Corynebacterium glutamicum and 5 mM trolox as an oxygen scavenger (Aitken et al., 2008). This oxygen scavenger was chosen based on measurements of RNAP elongation activity in buffers containing various OSS components (Figure S4). The slide was incubated in each imaging buffer for 10 min prior to data collection. A laser power of 10-15 mW at 532 nm was used to excite the sample in a prism-based total internal reflection geometry, and emission from DY547 and Cy5 were recorded simultaneously using an intensified CCD camera (I-Pentamax, Princeton Instruments). The sample was additionally excited by a 640 nm laser at the beginning and end of each dataset to verify the presence of Cy5. For on-slide elongation assays, ECs were prepared, immobilized, and imaged as described above. Imaging buffer containing 1 mM GTP was then added, and imaging was repeated after a 10-min incubation.

smFRET data analysis

Locations of molecules and fluorophore intensity traces for each molecule were extracted from raw video files using IDL (Research Systems) and analyzed using Matlab (The Math Works). Traces were manually selected for further analysis using the following criteria: single-step photobleaching, a Cy5 fluorescence intensity of >200 (arbitrary units) when excited at 640 nm, a total (DY547+Cy5) intensity of >200 when excited at 532 nm, and a fluorescence duration (prior to photobleaching) of >3 s and, if transitions were observed in the trace, anticorrelation between donor and acceptor signals. The FRET efficiency (E_{FRET}) was calculated as $I_A/(I_A+I_D)$, where I_A and I_D represent the background corrected fluorescence intensities of the acceptor (Cy5) and donor (DY547) fluorophores, respectively. FRET histograms were created by sorting the E_{FRET} observed for the first 100 frames of a large number of traces (>200 for main text data, >100 for supplementary data) into 22 bins with centers ranging from E_{FRET} = -0.05 to E_{FRET} = 1.05. Plotting and fitting of FRET efficiency histograms was performed in OriginLab 8.1. Uncertainties in the histogram fit parameters were computed by bootstrapping, splitting the entire set of data points into three batches and fitting them independently.

Hidden Markov Modeling (HMM) analysis was performed using the segmental k-means algorithm in the QuB software suite as previously described (Blanco and Walter, 2010). To select the number of states to model our data with, HMM was performed on a selection of data sets using 1-, 2-, and 3-state models. We used fits to the corresponding histograms to select the initial values for optimization of the average E_{FRET} and relative populations of the states. For each trace, we computed the Bayesian information criterion (BIC) and modified Bayesian information criterion (BIC'), as defined by Lerner et al. (Lerner et al., 2018), according to Equations 1-3:

$$BIC(q) = -2L + Kln(n) \qquad \text{Equation 1}$$

$$\Delta BIC(q) = BIC(q) - min_{p=1}^{3}\{BIC(p)\} \qquad \text{Equation 2}$$

$$BIC'(q) = \Delta BIC(q)/(n-K) \qquad \text{Equation 3}$$

where q is the number of states in the model being tested, K is the number of parameters in that model, L is the log-likelihood of the HMM fit, n is the number of data points in a given trace, and the number of states p varies over all of the models being considered. We then computed the mean BIC or BIC' over all of the traces in that data set. We found that for our data, BIC overvalued long traces while BIC' overvalued short traces. Unsurprisingly, then, BIC' suggested models with less than or equal to the number of states suggested by BIC. In none of the cases we tested did either metric suggest more than two states, and BIC suggested 2 states in all cases tested. We therefore used a 2-state model for the fitting and interpretation of all of our smFRET data.

Transcription Assays
Bubble-initiated transcription assays
Annealed nucleic acid scaffolds were prepared using RNA10 as described above for smFRET. Additional buffer components such as oxygen scavengers were added either immediately after annealing the nucleic acid scaffold, or after incubating the scaffold with RNAP (noted as "after RNAP" in Figure S4). Scaffolds were diluted to 50 nM in buffer B, *E. coli* RNAP core enzyme (New England Biolabs) was added to 200 nM, and the mixture was incubated at 37°C for 15 min. The reaction mixture was then treated with 200 μM GTP and incubated for 10 min, after which an aliquot was removed and quenched into loading buffer (1x = 4 M urea, 25 mM EDTA, 45 mM Tris-borate, pH 8.3). The remaining reaction was treated with 200 μM ATP, CTP and UTP and incubated for an additional 10 min, after which another aliquot was removed and quenched. The reaction aliquots were analyzed by denaturing, 8 M urea, 15% PAGE and visualized by detecting the Cy5 fluorescence using a Typhoon 9410 Variable Mode Imager (GE Healthcare Life Sciences). Gel images for these and all following experiments were analyzed in ImageQuant (Molecular Dynamics). For RNA degradation assays (Figure S4), a 10 nM solution of RNA10 was prepared in buffer B, including the specified OSS components and RNase inhibitors at the same concentrations used for smFRET. The mixture was incubated at 37°C for 30 min and analyzed by denaturing, 8 M urea, 15% PAGE.

We found through these assays that a number of common OSS components interfered with transcription in various ways. The OSS substrate protocatechuic acid (PCA) was found to interfere with transcription, as was the combined glucose/glucose oxidase/catalase (GOD/CAT) OSS, indicating inhibition by its product, D-gluconic acid. PCA and D-gluconic acid are both carboxylic acids, as are many known inhibitors of bacterial RNAP (Elgaher et al., 2014), and are therefore negatively charged under our imaging conditions. Inhibition was relieved by adding the OSS after incubating the bubble scaffold with RNAP, suggesting that PCA and D-gluconic acid interfere with binding rather than elongation. It is therefore likely that at the millimolar concentrations used for oxygen scavenging, they bind to and neutralize the electropositive cleft of RNAP where the transcription bubble resides. The maximum elongation efficiency was obtained by adding the GOD/CAT OSS after incubation with RNAP; therefore, this protocol was used for all experiments reported herein. Certain OSS enzymes, particularly catalase from bovine liver, were found to have RNase contamination and were therefore avoided, echoing reports of DNase contamination in OSS enzymes (Senavirathne et al., 2015).

Promoter-initiated transcription assays
A 151-nucleotide portion of the 5′ UTR of the queCDEF operon from *B. subtilis*, including the preQ$_1$ riboswitch, was cloned into the pUC18 plasmid between the KpnI and EcoRI restriction sites. The *E. coli* RecA promoter was inserted immediately upstream of the riboswitch between the BamHI and KpnI sites. In addition, 12 nucleotides not found in the wild-type sequence were inserted after the promoter in order to generate a 19-nucleotide stretch in which the RNA transcript lacks any adenosine residues. Mutant plasmids were generated by site-directed mutagenesis using the primers in Table S2. Transcription templates for all experiments except those utilizing ΔFT RNAP were prepared by PCR using the "Forward PCR primer" and "Reverse PCR primer" indicated in Table S2. Transcription templates for ΔFT RNAP and associated WT controls contained the Gal1 promoter instead of RecA, and were generated in two PCR steps using first "Forward Gal1 primer 1" and "Reverse PCR primer," then "Forward Gal1 primer 2" and "Reverse PCR primer."

For transcription assays, halted complexes (HCs) were prepared in buffer C (20 mM Tris-acetate, pH 8.0, 20 mM sodium acetate, 2 mM magnesium acetate, 5% glycerol, 14 mM 2-mercaptoethanol, 0.1 mM EDTA) containing 1 μM labeling NTP, 5 μM of two other NTPs (omitting ATP), 56 nM α^{32}P-GTP or CTP (3000 Ci/mmol), 100 μM ApC dinucleotide primer, and 40 nM DNA template. *E. coli* RNAP holoenzyme (New England Biolabs) was added to 80 nM, and the mixture was incubated for 15 min at 37°C. 0.9 reaction volumes of HC was treated with 0.1 reaction volumes of pre-warmed chase solution in buffer C, which contained 200 μg/mL rifampicin and 10x the desired final concentrations of all four NTPs and preQ$_1$. The mixture was incubated at 37°C, and reaction aliquots were quenched at the desired times into an equal volume of loading buffer (8 M urea, 0.8 × TBE, 0.2% bromophenol blue, 0.2% xylene cyanol, 1 mM EDTA). For most experiments, additional NTPs were then added to 200 μM, and a final aliquot was removed 3-5 min later. For experiments with RNAP variants (point mutants in the β' subunit and ΔFT RNAP), the concentration of RNAP was 50 nM instead of 80 nM. Sequencing ladders were prepared by combining 9 μL of HC with 1 μL of chase solution containing 250 μM of each NTP, in addition to one 3′-OMe NTP (at 25 μM for 3′-OMe GTP and 15 μM for 3′-OMe ATP, UTP and CTP). Reactions were incubated for 15 min at 37°C before being quenched with 10 μL of loading buffer. Reaction aliquots were denatured before loading 1.5-2.5 μL of each onto a denaturing 8 M urea, 10% polyacrylamide sequencing gel. The gel was dried and exposed to a phosphor screen (typically overnight), which was then scanned on a Typhoon 9410 Variable Mode Imager.

For reactions including transcription factors, cell extract or RNA hairpin and the associated control reactions, the template was purified by spermine precipitation (Hoopes and McClure, 1981). These reactions were performed by preparing HCs at twice the concentration described above, adding an equal volume of 2x the desired final concentration of factor, extract or hairpin, and 30 s later adding the chase solution. The final concentration of all transcription factors was 100 nM. In the case of *Eco* NusA, it was verified that concentrations as high as 500 nM had no effect beyond that seen at 100 nM.

Transcription data analysis
Volumes of the pause band following the aptamer domain (U46), the terminator in the expression platform (U70), the terminator following the expression platform (U108), and the two bands resulting from runoff were determined using ImageQuant (Molecular Dynamics). For each lane, these volumes were background corrected by determining the volume above background of all points that had a higher intensity than the average intensity of a background box positioned in the same lane. Two background boxes

were used in each lane: one to correct band U46 (positioned just below that band), and one to correct bands U70, U108 and the two runoff bands (positioned between U70 and U108). The relative intensity of each band in a given lane was determined by dividing the background-corrected volume of that band by the sum of the volumes of all five bands mentioned above. Two pause bands were observed within the aptamer domain, but these were omitted from the analysis due to their low intensity and lack of variation in response to $preQ_1$ and mutations. Error bars in transcription quantifications represent the standard deviation of the intensity fraction of each band across three independent experiments. To determine the $K_{1/2}$ for regulation of pausing and termination by $preQ_1$, the relative intensities of U46 and U70 at 45 s and 90 s after chasing, and 5 min after adding a further chase to 200 μM of each NTP were plotted as a function of $preQ_1$ concentration. These six datasets were globally fit with Equation 4:

$$F([preQ_1]) = y_0 + \frac{A * [preQ_1]}{K_{1/2} + [preQ_1]}$$

Equation 4

where F is the intensity fraction in a given band, y_0 is the intensity fraction at zero $preQ_1$, and A is the amplitude of the change that occurs upon addition of $preQ_1$. A global fit was performed in which y_0 and A were allowed to vary between datasets, but all were required to have the same value of $K_{1/2}$.

To determine the pause efficiency and lifetime from detailed time courses, the natural logarithm of the fraction paused was calculated at each time point t (seconds) and the following equation was fit to the resulting data:

$$F(t) = Ln\left[a_1 \cdot e^{m_1 t} + a_2 \cdot e^{m_2 t}\right]$$

Equation 5

Fitting to the logarithm of the data rather than the raw data allowed the time points with low paused fraction to still contribute to the fit. a_1 and a_2 correspond to the fraction of elongation complexes that pause with half-lives $-Ln[2]/m_1$ and $-Ln[2]/m_2$, respectively. The constraint of $a_1 + a_2 \leq 1$ was used in all fits.

Cross-linking Assays

100 nM DNA template and 20 nM RNAP holoenzyme were incubated in buffer D (20 mM Tris-HCl, 120 mM KCl, 5 mM MgCl$_2$, 5% glycerol (w/v), 1 mM 2-mercaptoethanol, pH 7.9) at 37°C for 15 min. ApU (100 μM) and starting NTPs (30 μM CTP, GTP, and ATP) were added to the promoter bound RNAP, and the reaction was incubated for 30 min at 37°C. The EC at position 6 was immobilized on 10 μL Ni^{2+}-NTA agarose beads at 37°C for 15 min. Each following walk (indicated in Figure S5) was completed with 5 intermediate washes with buffer D to eliminate unbound NTPs. Each wash involved addition of 1 mL of buffer D followed by a 5 s spin on a tabletop centrifuge to pellet the Ni^{2+} beads with attached ECs. The supernatant was then removed by pipette. The cross-linkable analog 4-thio-UTP or 6-thio-GTP (10 μM) was introduced into the RNA at the desired position. Upon reaching A36, 10 μCi [α-^{32}P]-GTP, 3000 Ci/mmol was added for 2 min. 5 μM ATP, CTP, and GTP were then added to reach C43, then 5 μM UTP and GTP to extend to U46. ECs C43 and U46 were incubated with 40 μM $preQ_1$ at 37°C for 5 min and exposed to 365 nm UV light (UVP, 8W Model UVLMS-38) for 10 min on ice. The reactions were either quenched with stop buffer (9M Urea, 20 mM EDTA, 1XTBE, 0.5% Brilliant Blue R and 0.5% Xylene Cyanol FF) and the sample resolved on denaturing 7M urea 15% polyacrylamide-gel, or in LDS loading dye and separated on a 4%–12% Bis-Tris Novex gel (Invitrogen). The gels were dried and exposed to phosphor screens.

Simulations

Molecular dynamics simulations were performed to explore the folding landscape of the $preQ_1$ riboswitch bound to the RNA polymerase elongation complex. Structure-based or Gō-type simulations are particularly useful for exploring slow processes like folding that are computationally expensive to probe with more precise explicit solvent simulations. Structure-based models were originally used in protein folding (Shea et al., 1998) and are based on energy landscape theory and the principle of minimal frustration (Onuchic and Wolynes, 2004). In order to fold on biological timescales, native interactions must be more stabilizing than nonnative interactions, so structure-based models only include interactions between native contacts. We chose to use an all heavy atom structure-based model that has been extensively calibrated, and has been used for RNA folding (Lutz et al., 2014; Whitford et al., 2009) and functional studies of the ribosome, a mixed protein and RNA system (Whitford et al., 2010).

To generate a structure-based model, a native structure is needed to specify the interactions that ought to be stabilized. Atoms in close proximity within a known structure are chosen to be contact pairs that interact within the model; the shadow algorithm was used to calculate the contact map (Noel et al., 2012). Since no structure of the $preQ_1$ riboswitch docked to RNAP existed, a model was created from structures of various components. Run input files for various structural elements were prepared using SMOG 2 (Noel et al., 2010, 2016), and then stitched together as described. The $preQ_1$ riboswitch was modeled using contacts present in its crystal structure (PDB: 3FU2) (Klein et al., 2009). Its NMR structure (PDB: 2L1V) (Kang et al., 2009) was used to fill in 5′ and L2 gaps from the crystal structure: any contacts and interactions with missing residues were used to supplement the interactions defined by the crystal structure. An *E. coli* RNAP model was adopted from a cryo-EM structure of an EC (PDB: 6ALH) (Kang et al., 2017). The transcription bubble was modeled by creating single- and double-stranded sections of RNA and DNA using the nucleic acid builder (http://structure.usc.edu/make-na/server.html), and then stitching the resulting topologies together appropriately. Interactions of residues in either the riboswitch or the polymerase overrode those interactions in the DNA bubble model, so the bubble was only used to supplement interactions for missing residues. Since several topologies were being combined, and contact and dihedral interactions are scaled slightly differently in each, the dihedral and contact energies were both set to typical values

of 0.7 energy units. We refer to the supplemental information of Hayes et al., 2014 for a discussion of units in the model. Models with 0 to 3 nucleotides inserted between the riboswitch and RNAP (RNA0p, RNA1, RNA2 and RNA3) were made by modifying the sequences generated with the nucleic acid builder (sequences are reported in Table S1).

The resulting model essentially stitches together models of the riboswitch and the RNAP. Since the riboswitch and RNAP fold independently, the model is valid, but only accounts for steric interactions between the riboswitch and RNAP, and neglects any specific interactions that could be obtained from a structure of the complex. Consequently, some questions are beyond the scope of the model, while others can be robustly answered. The average number of base pairs formed in the model should be fairly robust, as the model identifies structures of the riboswitch that lead to a minimal disruption of RNAP structure. Likewise, the model should be able to elucidate the sterically allowed binding poses for the folded riboswitch, whereas identifying the dominant pose is beyond the scope of the model, as the populations of competing poses can be shifted by specific interactions between the riboswitch and RNAP.

Equilibrium simulations of riboswitch folding were run using umbrella sampling on the Q (number of formed contacts) reaction coordinate with replica exchange between umbrellas to allow trapped high-energy folded poses to relax. 11 replicas were used with an exchange attempts every 5000 steps, an umbrella k constant of 0.005 energy units per contact squared, and umbrella equilibria Q_0 evenly spaced among replicas from 0 to 282 contacts. The weighted histogram analysis method (WHAM) (Kumar et al., 1992) was then used to reweight each frame of each replica to an unbiased simulation. Simulations started from an unfolded $preQ_1$ riboswitch structure, and allowed the riboswitch to fold at a temperature of T = 85 units. (The folding temperature of the riboswitch is approximately 90 units in the presence of the polymerase.) Simulations were run for 50 million steps with the first 12.5 million steps discarded for equilibration. Simulations with the $preQ_1$ ligand included 5 ligands in a cubic box of 20 nm side length. The binding affinity of $preQ_1$ was not calibrated.

Simulations were run with and without $preQ_1$ and varying numbers of inserted nucleotides to determine whether steric constraints allowed P2 to fold (Table S3). A base pair was considered formed if the central hydrogen bond between purine N1 and pyrimidine N3 was within a factor of 1.5 of its native distance. (In the A·C mismatch, the only hydrogen bond was used instead.) The number of formed base pairs was counted for each frame and averaged using the weight given by WHAM over all replicas in the simulation. The average number of base pairs was largely unchanged in the absence or presence of $preQ_1$ for RNA1, RNA2 and RNA3 (1.8-2.4 versus 2.0-2.4, respectively), and depends weakly on number of inserts, but changes substantially in the zero insert case (RNA0p), where it is lower by about half a base pair with $preQ_1$ and a whole base pair without $preQ_1$. To check the robustness of the results, the contacts in the RNA linker (all residues between the 3′ end of P2 and the 5′ end of the RNA-DNA duplex) were removed. This allows the RNA to pull slightly out of the exit tunnel, and makes the RNA more flexible so it can twist more. While about half an additional base pair could form with the more flexible model, the trends remained the same: one to three inserts formed roughly the same number of base pairs as each other, and no inserts formed half a base pair less than them. Finally, a model of the flap tip deletion mutant was made in which residues K890-K914 of the beta subunit were removed. Simulations of the flap tip deletion mutant achieved results intermediate between the other two models, but showed less sensitivity to $preQ_1$ in the RNA0p system.

To connect with cross-linking studies and to elucidate the folded pose, the frequency with which the cross-linking residue U14 was within 8 Å of each protein residue was calculated. Two poses of the riboswitch were identified that involved interactions primarily between the riboswitch and β′ (collectively called "mode 1"), along with a third pose where U14 is wedged between the clamp and flap tip ("mode 2"). Within mode 1, simulations with the default parameters (rigid linker) yielded a pose where U14 was positioned near K76-K79, while with the floppy linker, an additional pose was observed in which it is positioned near T393. These results are in agreement with cross-linking results that show that U14 interacts with the β′ subunit.

DATA AND SOFTWARE AVAILABILITY

Custom scripts used for this work are available upon request.

Supplemental Information

Ligand Modulates Cross-Coupling

between Riboswitch Folding

and Transcriptional Pausing

Julia R. Widom, Yuri A. Nedialkov, Victoria Rai, Ryan L. Hayes, Charles L. Brooks III, Irina Artsimovitch, and Nils G. Walter

Supplemental Figures and Tables for Widom et al., "Ligand Modulates Cross-Coupling between Riboswitch Folding and Transcriptional Pausing"
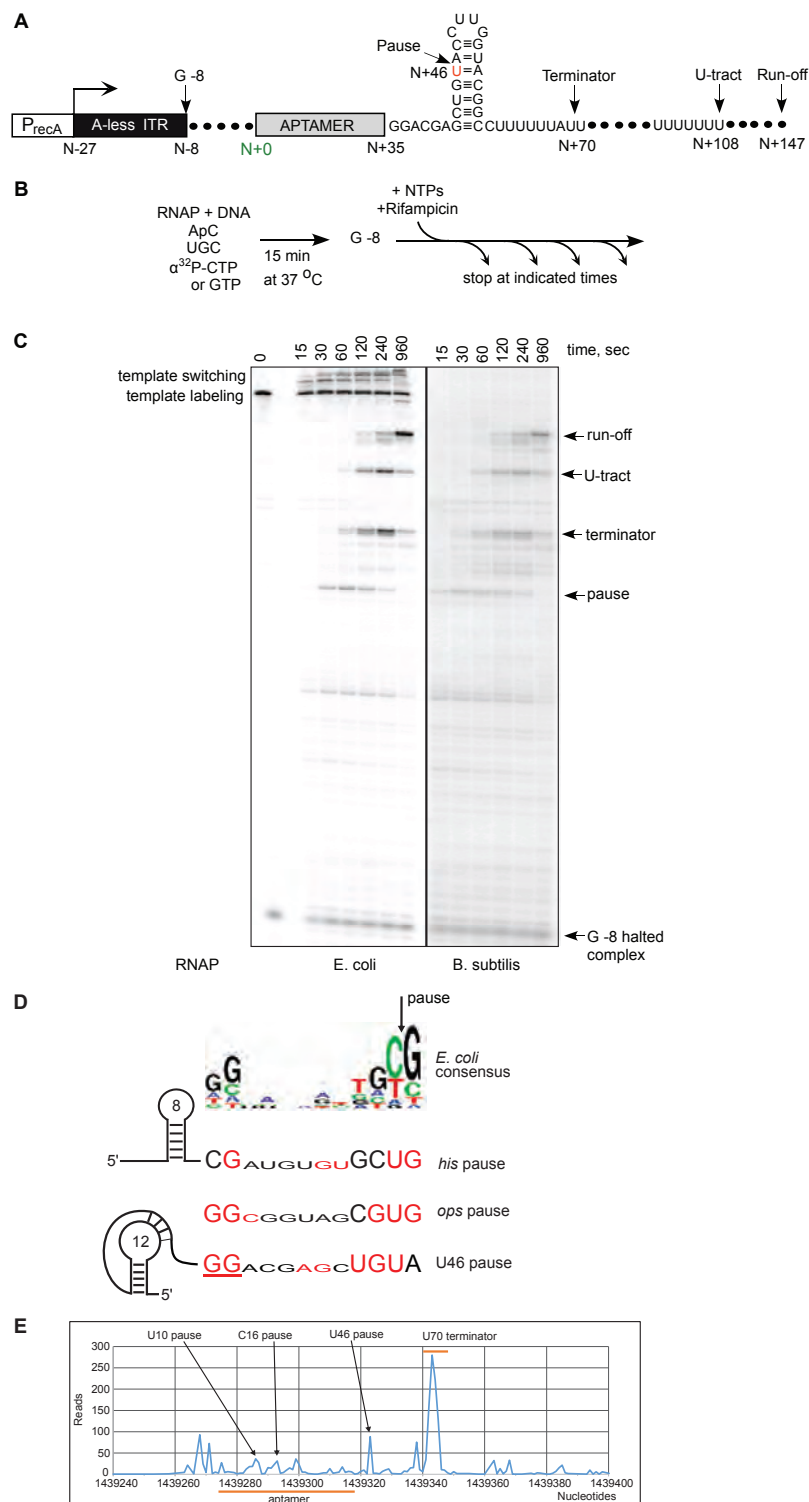
Figure S1:

Figure S1. Comparison of *Eco* and *Bsu* RNAPs Transcribing a Template Encoding the PreQ$_1$ Riboswitch, related to all figures

A) The RNA produced from a strong *recA* promoter contains a 20 nucleotide-long initially transcribed region (ITR) devoid of A residues, the preQ$_1$ riboswitch aptamer domain and the expression platform. The numbering scheme used throughout the paper is shown, with nucleotide N+0 being the first nucleotide of the aptamer used for smFRET. Four major RNA species are indicated: the *que* pause (U46), terminator (U70), U-tract arrest site (U108) and runoff product (U149).

B) Schematic of a single-round transcription assay. Halted ^{32}P-labeled G -8 elongation complexes were formed by incubating *Eco* or *Bsu* RNAP holoenzyme (40-80 nM) and linear PCR-generated template with starting substrates (1 µM CTP, 5 µM GTP and UTP, 10 µCi $[\alpha^{32}P]$-CTP or $[\alpha^{32}P]$-GTP, 3000 Ci/mmol and 100 µM ApC primer) for 15 minutes at 37 °C. Elongation was re-started by the addition of all NTPs to 25 µM and 20 µg/mL rifampicin, and was stopped at the indicated times by quenching aliquots into urea loading buffer.

C) A representative denaturing urea-polyacrylamide gel with the positions of major RNA species indicated shows that *Eco* and *Bsu* RNAPs produce the same intermediate and final species in roughly the same proportions. We note that *Bsu* RNAP (a generous gift from Tina Henkin) does not exhibit the template labeling and switching activities characteristic of the *Eco* enzyme. The vertical black line indicates the location where the lower-intensity *Bsu* lanes were spliced next to the *Eco* lanes after an extended exposure.

D) Comparison of the *que* pause sequence to the consensus pause sequence from *Eco* and the well studied *his* and *ops* pause sequences. The *his* pause additionally requires a hairpin in the nascent RNA. Red bases in all sequences are those that match the *Eco* consensus; underlined bases are those for which mutations were investigated (see Figures 6 and S6).

E) Mapped reads from the *Bsu* data set published in Larson et al., *Science* 6187, 1042-1047 (2014). The three pauses observed in our experiments (U10, C16 and U46) are also observed *in vivo*.
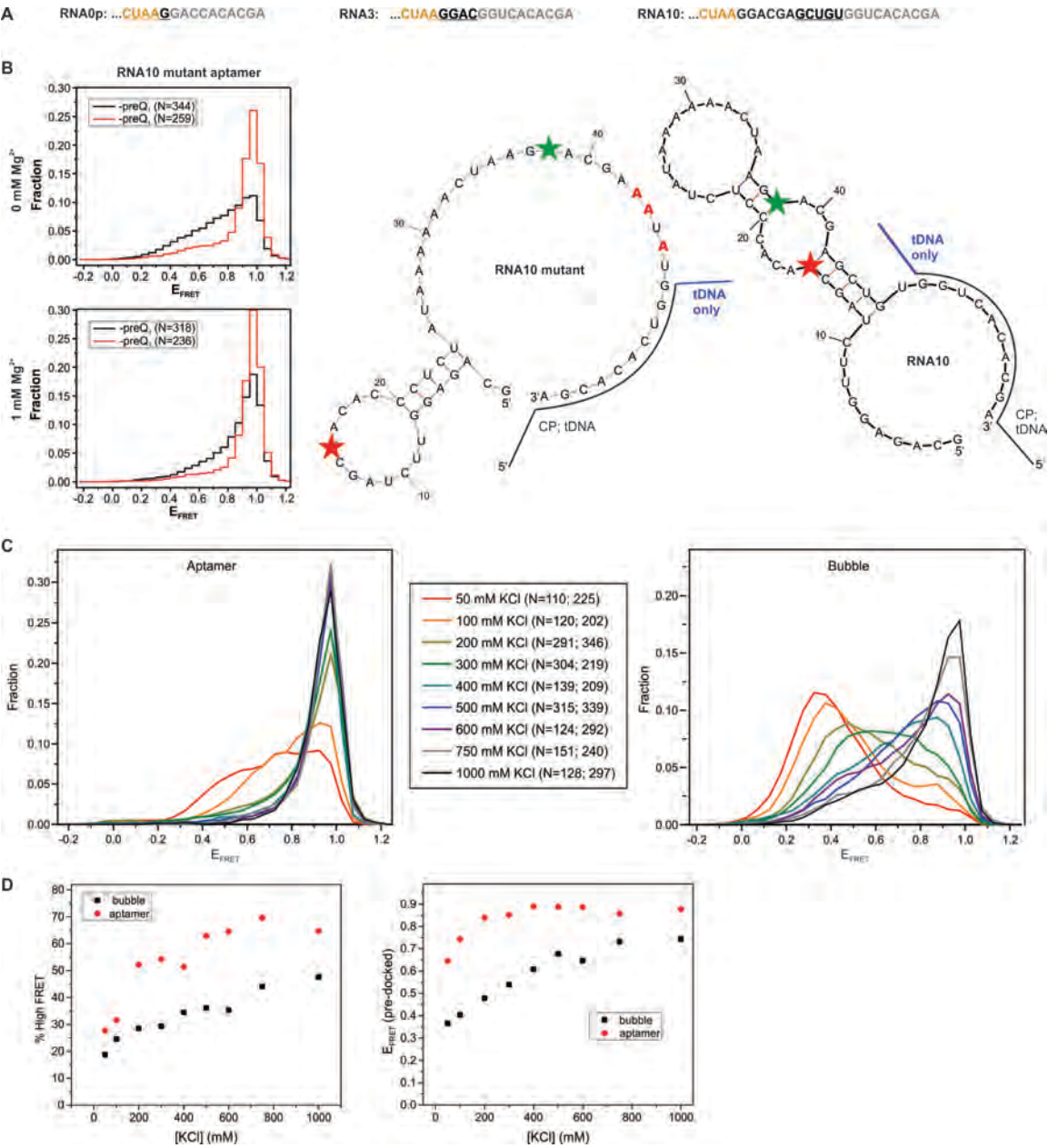
Figure S2:

A   RNA0p: ...**CUAAG**GACCACACGA   RNA3: ...**CUAAGGAC**GGUCACACGA   RNA10: ...**CUAA**GGACGA**GCUGU**GGUCACACGA

Figure S2. Additional smFRET Data, related to Figures 2 and 3

A) RNA sequences used in smFRET experiments. Ellipses indicate the first 32 nucleotides of the aptamer sequence in Table S1. Orange bases make up the 3' segment of P2, and gray bases make up the RNA:LNA or RNA:DNA hybrid in aptamer or bubble/EC immobilization, respectively. Underlined bases fall within the RNAP exit channel under EC immobilization.

B) Right: mFold-predicted structures of RNA10, requiring that bases involved in LNA or DNA binding, plus one additional base adjacent to them, remain single-stranded. The LNA capture probe is shown schematically as a black line, and the additional segment present in the tDNA is shown as a blue line. RNA10 adopts a non-native conformation that places the fluorophores in close proximity, explaining the observation of a high-FRET state that is not sensitive to $preQ_1$ (Figures 2A and 3A). Three mutations (red) were made to generate a version of RNA10 that was predicted to adopt the native structure. This mutant responds to $preQ_1$ (left).

C) Additional salt can offset the effects of the DNA bubble. The FRET histogram of RNA3 under aptamer immobilization changes little at KCl concentrations higher than 100 mM (left). The FRET histogram under bubble immobilization continues to shift over the entire range of 50-1000 mM KCl (right). ~400 mM KCl increases the E_{FRET} of the pre-docked conformation and the population of docked conformation to the levels seen in the aptamer at 100 mM KCl. Histograms were recorded in the presence of 1 mM $MgCl_2$.

D) Quantification of the percentage of population in the docked conformation (left) and the E_{FRET} of the pre-docked conformation (right) as a function of KCl concentration, obtained by fitting each histogram in (C) to two Gaussians.
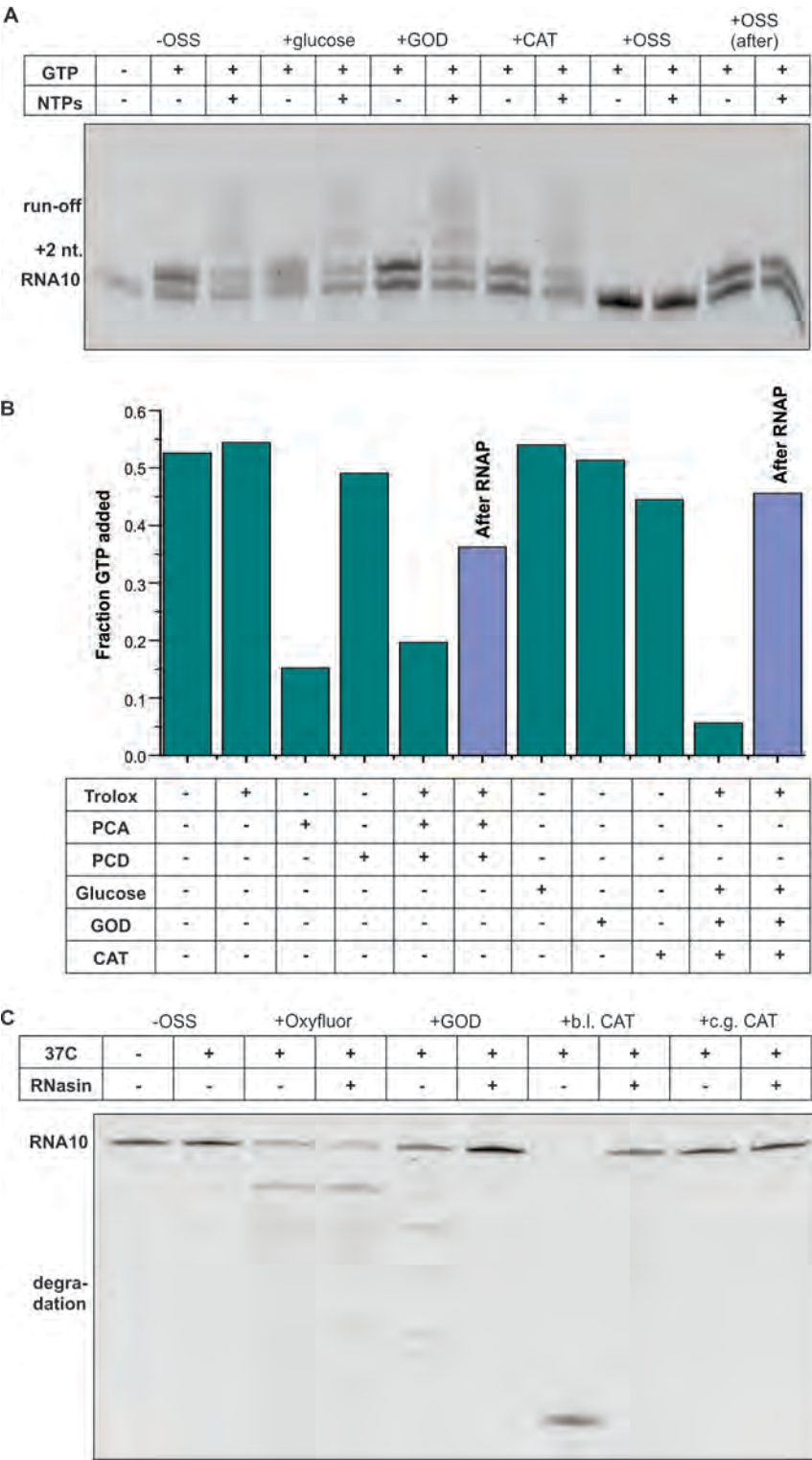
Figure S3:

Figure S3. Optimization of Oxygen-Scavenging System (OSS) for Imaging of ECs, related to Figures 3 and 4

A) Example gel showing that the entire glucose oxidase/catalase/glucose (GOD/CAT) OSS inhibits elongation, even though no component has an effect individually. Elongation is rescued if the OSS is added after incubating with RNAP (right-most two lanes), suggesting that the OSS interferes with RNAP binding.

B) Quantification of the effects of GOD/CAT and protocatechuic acid/protocatechuate 3,4-dioxygenase (PCA/PCD) OSS systems on elongation. Adding OSS after RNAP rescues elongation to a larger extent with the GOD/CAT OSS.

C) Optimization of OSS to avoid RNA degradation. Another widespread OSS component, oxyfluor, leads to RNA degradation that is not prevented RNase inhibitor (RNasin). Glucose oxidase from *Aspergillus niger* leads to some degradation that is prevented by RNasin. Catalase from bovine liver (b.l.) exhibits very high RNase activity that is largely inhibited by RNasin. Catalase from *Corynebacterium glutamicum* (c.g.) exhibits minimal RNase activity.
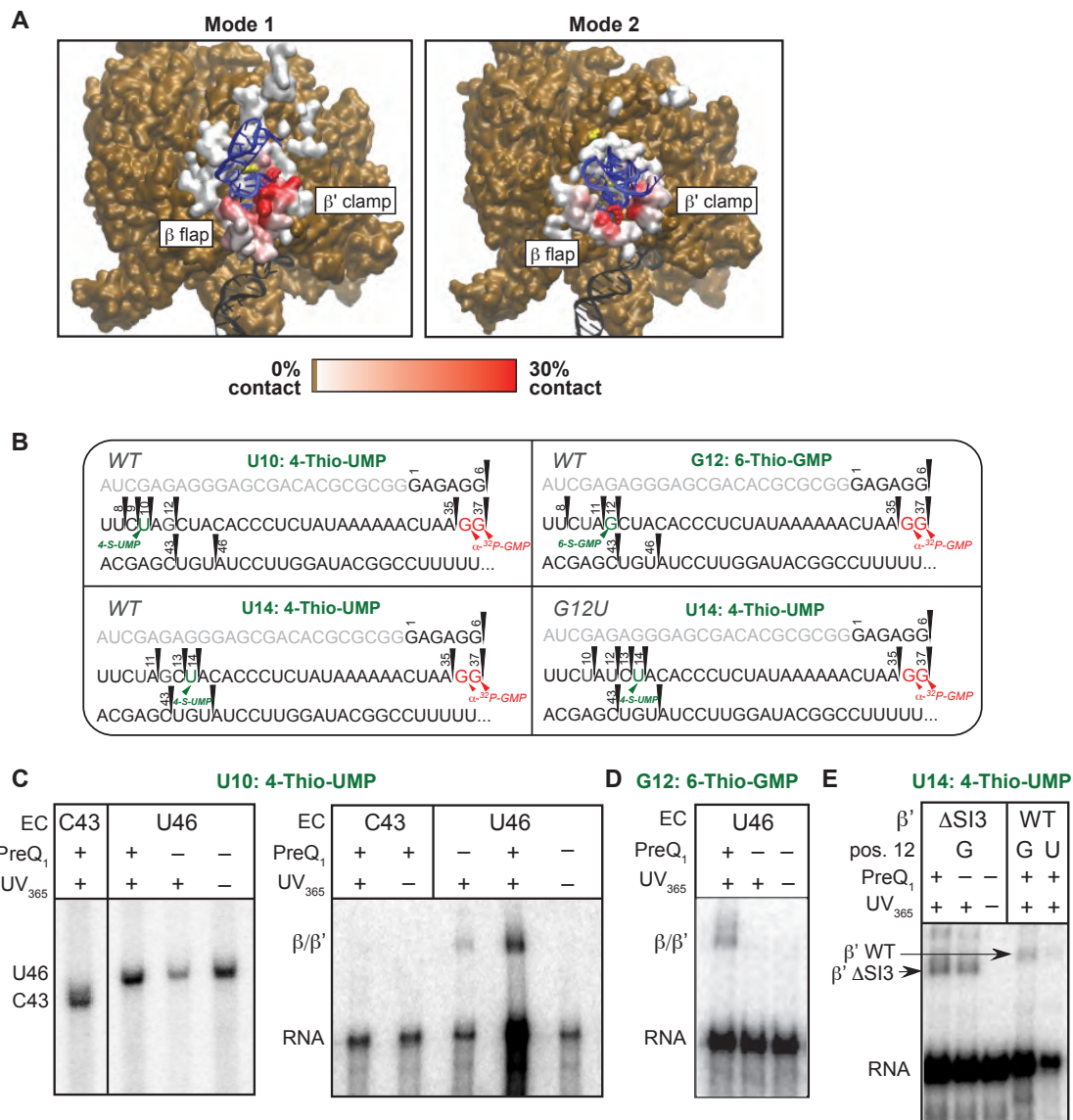
Figure S4:

A

Mode 1 Mode 2

β' clamp β' clamp

β flap β flap

0% contact — 30% contact

B

WT U10: 4-Thio-UMP
AUCGAGAGGGAGCGACACGCGCGG GAGAGG
UUCUAGCUACACCCUCUAUAAAAAACUAA GG
4-S-UMP α-³²P-GMP
ACGAGCUGUAUCCUUGGAUACGGCCUUUUU...

WT G12: 6-Thio-GMP
AUCGAGAGGGAGCGACACGCGCGG GAGAGG
UUCUAGCUACACCCUCUAUAAAAAACUAA GG
6-S-GMP α-³²P-GMP
ACGAGCUGUAUCCUUGGAUACGGCCUUUUU...

WT U14: 4-Thio-UMP
AUCGAGAGGGAGCGACACGCGCGG GAGAGG
UUCUAGCUACACCCUCUAUAAAAAACUAA GG
4-S-UMP α-³²P-GMP
ACGAGCUGUAUCCUUGGAUACGGCCUUUUU...

G12U U14: 4-Thio-UMP
AUCGAGAGGGAGCGACACGCGCGG GAGAGG
UUCUAUCUACACCCUCUAUAAAAAACUAA GG
4-S-UMP α-³²P-GMP
ACGAGCUGUAUCCUUGGAUACGGCCUUUUU...

C U10: 4-Thio-UMP

EC	C43	U46				EC	C43		U46		
PreQ₁	+	+	−	−		PreQ₁	+	+	−	+	−
UV₃₆₅	+	+	+	−		UV₃₆₅	+	−	+	+	−

U46
C43

β/β'
RNA

D G12: 6-Thio-GMP

EC	U46		
PreQ₁	+	−	−
UV₃₆₅	+	+	−

β/β'
RNA

E U14: 4-Thio-UMP

β'	ΔSI3			WT	
pos. 12	G			G	U
PreQ₁	+	−	−	+	+
UV₃₆₅	+	+	−	+	+

β' WT
β' ΔSI3
RNA

Figure S4. Cross-linking of the Riboswitch to RNAP, related to Figure 5

A) MD simulation snapshots illustrate two binding modes in which P2 interacts with RNAP. In Mode 1, observed in all simulations, U14 interacts with residues near Y68 and L78 on the β' subunit. In Mode 2, observed in all simulations but particularly predominant in simulations of ΔFT RNAP, U14 additionally interacts with D912 and E892 (for WT RNAP) or I870 and K914 (for ΔFT RNAP) on the β subunit. PreQ$_1$ is shown in yellow. "% contact" refers to the percent of all simulation frames during which a heavy atom on U14 was within 8 Å of a given heavy atom on RNAP.

B) The sequence of walking steps used to generate ECs for the cross-linking experiments reported here and in Figure 5. The numbering system is explained in Figure S1. Black arrowheads indicate RNAP stalling points during the walk. Cross-linkable analogs were incorporated at positions indicated in green, and radiolabeled GMP at positions indicated in red. In panels B, C and D, where indicated, 40 µM preQ$_1$ was added for 5 min at 37 °C, followed by exposure to 365 nM UV light for 10 minutes on ice. The cross-linked ECs were analyzed on 4-12% Bis-Tris polyacrylamide gels.

C) ECs were walked to C43 or U46 following incorporation of 4-S-UMP at position 10. The vertical black line indicates the location where an empty lane was deleted. RNA-to-β/β' cross-links are observed at U46, while at C43, the 3' leg of P2 is not yet accessible and cross-linking is minimal.

D) U46 ECs with 6-S-GMP cross-linkable analogue at position 12. A Bis-Tris gel shows RNAP-RNA cross-linking is observed only in the presence of preQ$_1$.

E) U46 ECs containing 4-thio-UMP at position 14 were prepared with wild-type RNAP (WT) or the enzyme with a large deletion (residues 943-1130) in the β' subunit (ΔSI3). Comparison of the cross-linked species obtained with WT and ΔSI3 RNAPs (arrows) indicates that U14 is cross-linked to β'. Cross-linking is greatly reduced by the P2-weakening mutation G12U.
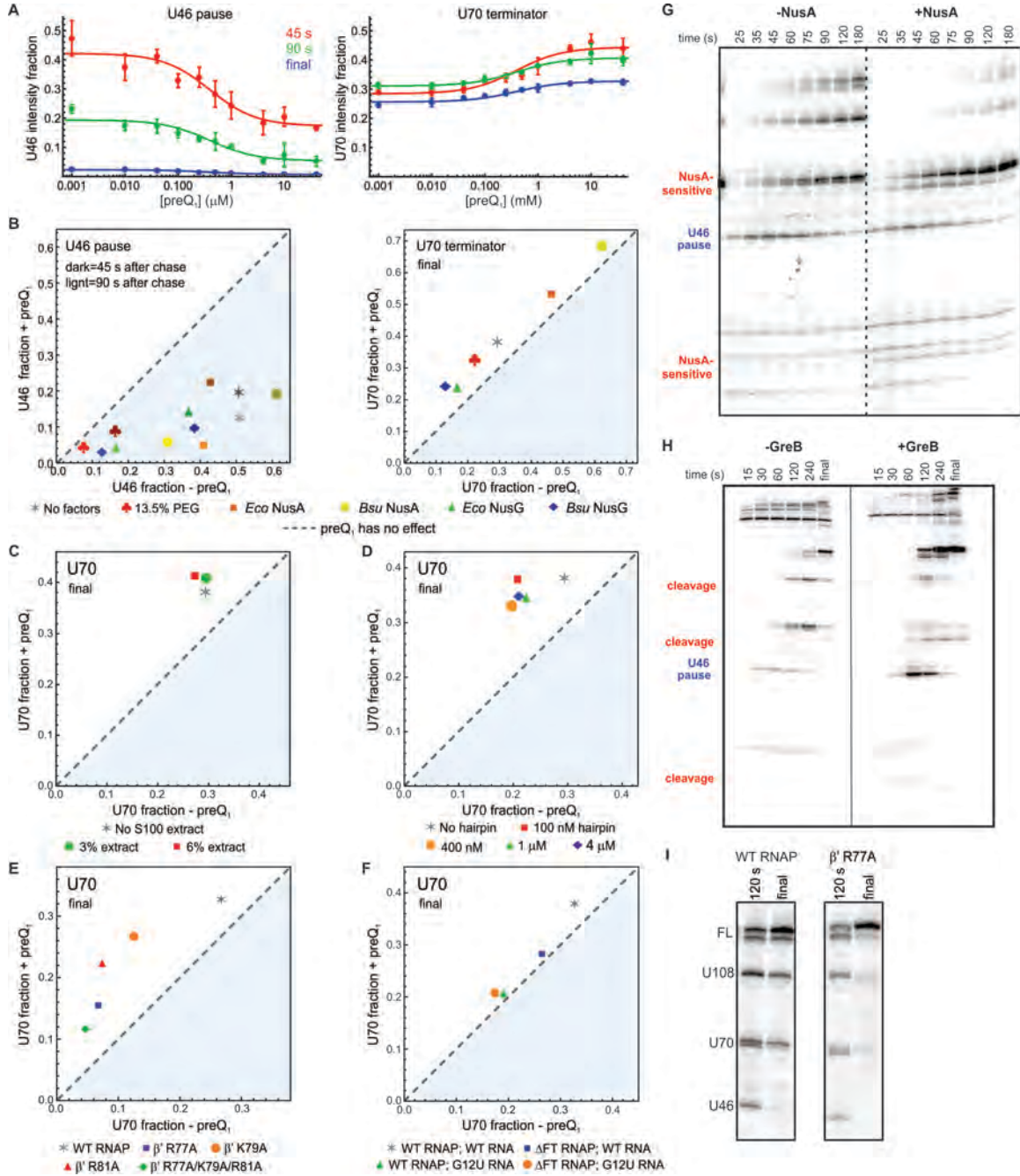
Figure S5:

Figure S5. Effects of Various Factors on Transcriptional Pausing and Termination, related to Figure 6.

A) Quantification of preQ$_1$ titrations in which G-8 halted complexes were chased with 25 µM NTPs and preQ$_1$ at the indicated concentration. Aliquots were taken 45 and 90 seconds after chasing, and a final aliquot was taken 5 minutes after addition of NTPs to 200 µM.

B) Effects of transcription factors and macromolecular crowding on pausing (left) and termination (right). NusA factors from *Eco* and *Bsu* do not have a significant effect on the *que* pause (see also panel G). NusG factors and macromolecular crowding using poly(ethylene glycol) both decrease pausing. NusA factors increase termination at U70 and NusG factors decrease it, but none increase the effect of preQ$_1$ on termination.

C) Adding *Eco* S100 extract does not significantly change termination efficiency or the magnitude of the effect of preQ$_1$.

D) Adding an RNA hairpin with the same sequence as P1 and L1/L2 of the riboswitch (to mimic the existence of already-transcribed RNA) decreases termination efficiency in the absence of preQ$_1$ but does not significantly change the magnitude of the effect of preQ$_1$.

E) Point mutations in the β' subunit near the exit channel decrease transcription termination at U70, particularly R77A and R77A/K79A/R81A triple mutations.

F) Compared to WT RNAP, ΔFT RNAP significantly decreases termination on templates encoding the WT riboswitch. ΔFT RNAP has no effect on termination on the template encoding the G12U riboswitch variant.

G) NusA has no effect on the *que* pause while it is active at other positions on the template, suggesting a mechanism distinct from canonical hairpin-stabilized pausing. The dashed line in the middle indicates the division between samples without (left) and with (right) NusA.

H) At the *que* pause, GreB does not promote cleavage of the RNA into shorter products while it does at other positions on the template, indicating that RNAP is not backtracked at the pause. The solid line in the middle indicates where two non-adjacent portions of the gel were spliced together.

I) R77A (shown) and triple RNAP mutants exhibit significantly decreased termination at U70 and U108, and the position of termination around U108 is "smeared out".
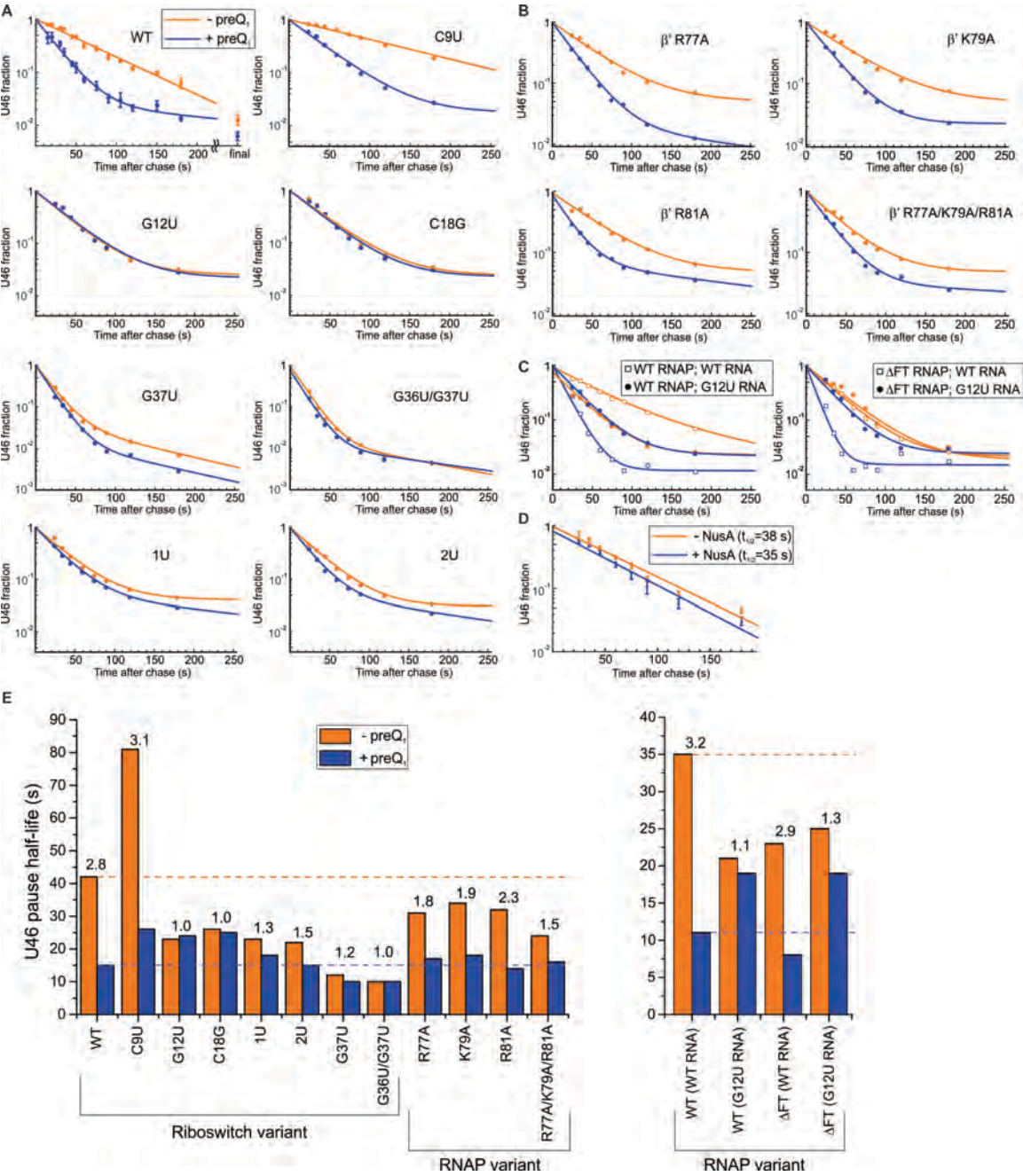
Figure S6:

Figure S6. Transcription Time Courses for Riboswitch and RNAP Variants, related to Figure 6

A) Fraction of complexes at the *que* pause site as a function of reaction time for WT (reproduced from Figure 5) and all riboswitch mutants.

B) Fraction of complexes at the *que* pause site as a function of reaction time for RNAP with point mutations in the β' subunit (transcribing the WT riboswitch).

C) Fraction of complexes at the *que* pause as a function of reaction time for WT and ΔFT RNAPs on WT and G12U riboswitch templates. The templates used in these experiments contained a Gal1 promoter to facilitate initiation by ΔFT RNAP (all other experiments used a RecA promoter).

D) Fraction of complexes at the *que* pause site as a function of reaction time for the WT riboswitch with and without NusA. The template used in these experiments was purified by spermine precipitation.

E) Quantification of pause half-life based on the fits shown above. For double-exponential fits, the half-lives plotted here are those of the vast majority population (≥73% in all cases). Results in the plot on the left were obtained using templates with a RecA promoter, and those on the right using a Gal1 promoter. Fit parameters are presented in Tables S5-S6.

Figure S7:

Variant	Schematic	Direct effect	Nature of pause; absence of preQ$_1$	Nature of pause; presence of preQ$_1$	Analogous effect on *his* pause
WT			pseudoknot-stabilized	pseudoknot-inhibited	
C9U		strengthens pseudoknot	more strongly pseudoknot-stabilized	pseudoknot-inhibited	
G12U		weakens pseudoknot	elemental	elemental	elemental pause in absence of hairpin/duplex
C18G		eliminates preQ$_1$ binding	elemental	elemental	
1U, 2U		moves pseudoknot further from RNAP	elemental	weakly pseudoknot-inhibited	insertion of 2 bases weakens pause
G36U, G36U/G37U		changes consensus to anti-consensus	weaker elemental	weaker elemental	
RNAP point mutants		disrupts electrostatic interactions with exit channel	weakly pseudoknot stabilized	pseudoknot-inhibited	
ΔFT RNAP; WT RNA		deletes flap-tip	elemental	pseudoknot-inhibited	weakens pause; smaller effect of hairpin
ΔFT RNAP; G12U RNA		deletes flap-tip and weakens pseudoknot	elemental	elemental	FT does not affect elemental pause

Figure S7. Summary of Transcription Assay Results, related to all figures

For each riboswitch and RNAP variant tested, the properties of the *que* pause and its relationship to the *his* pause are summarized.

Supplemental Tables

WT Aptamer RNA	G**C**A GAG GUU CUA GC(5-N-U) ACA CCC UCU AUA AAA AAC UAA G(G[2'Dy547])A C
EC RNA 0p	pCAC ACG A
Splint LNA 0p	+T+CG +TGT +GGT CCT TAG TTT TTT ATA GAG
EC RNA 3	pGGU CAC ACG A
Splint DNA 3	TCG TGT GAC CGT CCT TAG TTT TTT ATA GAG
EC RNA 10	pGAG CUG UGG UCA CAC GA
Splint DNA 10	TCG TGT GAC CAC AGC TCG TCC TTA GTT TTT TAT AGA G
EC RNA 10 mutant	pGA**A** A**U**A UGG UCA CAC GA
Splint DNA 10 mutant	TCG TGT GAC CA**T** A**TT** TCG TCC TTA GTT TTT TAT AGA G
tDNA for RNA0p	CTC TGA ATC TCT TCC TCG TGT GGT CAG GAC GTA CTG ACC
tDNA for RNA3, RNA10	CTC TGA ATC TCT TCC TCG TGT GAC CAG GAC GTA CTG ACC
biotin-ntDNA	GGT CAG TAC GTC CAT TAG CTC TTC GGA AGA CAT TCA GAG-C18-biotin
ntDNA	GGT CAG TAC GTC CAT TAG CTC TTC GGA AGA GAT TCA GAG
CP LNA for RNA0p	biotin-TTTTTT+CG+TG+TGA+CC
CP LNA for RNA3, RNA10	biotin-TTTTTT+CG+TG+TGG+TC
RNA hairpin for transcription	AGA GGU UCU AGC UAC ACC CUC U

Table S1: Sequences of Oligonucleotides, related to Figures 1, 2, 3 and 6. Sequences of oligonucleotides used for smFRET experiments, written in 5' to 3' direction. Abbreviations are as follows: 5-N-U (5-aminoallyl uracil), G[2'Dy647] (guanosine labeled with Dy547 at the 2' position), p (5' phosphate), C18 (18-carbon ethylene glycol spacer), + (following nucleotide is LNA). Bold text indicates the sites of mutations.

Forward PCR primer	GTA AAA CGA CGG CCA GT
Reverse PCR primer	ATT ACG AAT TCG CGA GAG ACA TCG
C9U SDM primer	AGA GGT T**T**T AGC TAC ACC CTC TAT AAA AAA C
G12U SDM primer	AGA GGT TCT A**T**C TAC ACC CTC TAT AAA AAA C
C18G SDM primer	GGT TCT AGC TAC A**G**C CTC TAT AAA AAA CTA AGG
G37U SDM primer	CCT CTA TAA AAA ACT AAG **T**AC GAG CTG TAT CCT TGG
G36U/G37U SDM primer	CCT CTA TAA AAA ACT AA**T** **T**AC GAG CTG TAT CCT TGG
+1U SDM primer	TAA AAA ACT AA**T** GGA CGA GCT GTA TCC
+2U SDM primer	TAA AAA ACT AA**T** **T**GG ACG AGC TGT ATC C
Forward Gal1 primer 1	CACTAATTTATTCCATGTCACACTTTTCGCATCTTTTT TATGCTATAATTATTTCACCGCTCTGCGTTCCGTG
Forward Gal1 primer 2	CACTAATTTATTCCATGTCACACTTTTCGCATCTTTTT TATGC

Table S2: Sequences of Primers, related to Figure 6. Bold text indicates the sites of mutations.

	Mean BIC			Mean BIC'		
Experimental condition	1 st.	2 st.	3 st.	1 st.	2 st.	3 st.
RNA0p apt. -Mg -preQ$_1$	-1266	**-1293**	-1167	**0.05**	0.09	0.37
RNA0p apt. -Mg +preQ$_1$	-681	**-1045**	-1023	0.66	**0.07**	0.24
RNA0p apt. +Mg -preQ$_1$	-946	**-956**	-946	**0.05**	0.10	0.20
RNA0p apt. +Mg +preQ$_1$	-1103	**-1317**	-1248	0.37	**0.05**	0.34
RNA0p EC +Mg -preQ$_1$	-238	**-272**	-236	**0.11**	0.13	0.38
RNA0p EC +Mg -preQ$_1$	-515	**-589**	-555	0.17	**0.11**	0.31
RNA3 EC +Mg -preQ$_1$	-567	**-573**	-547	0.29	**0.18**	0.29
RNA3 EC +Mg +preQ$_1$	-1068	**-1145**	-1108	**0.10**	0.12	0.30

Table S3: Metrics for FRET Model Selection, related to Figures 2, 3 and 4. For a selection of experimental conditions, traces were analyzed by HMM using 1, 2 and 3 state ("st.") models. The BIC and BIC' (described in Methods) were computed for each trace, and their means across the entire data set are reported here. Entries in bold indicate the model with the optimal BIC or BIC' for a given data set.

	Average number of base pairs of P2 formed			
RNA species	-preQ$_1$	+preQ$_1$	-preQ$_1$, slippery exit channel	+preQ$_1$, slippery exit channel
RNA0p	1.4	1.8	2.0	2.1
RNA1	1.8	2.0	2.5	2.5
RNA2	1.9	2.0	2.7	2.7
RNA3	2.4	2.5	2.7	2.8

Table S4: P2 Base pairs Predicted by MD, related to Figure 5. Molecular dynamics simulations predict the average number of base pairs of P2 that are formed in ECs containing RNA0p and inserts of 1-3 nucleotides.

Template	a1 (fraction paused)	τ1 (s) [pause half-life]	a2	τ2 (s)	R^2
WT-PQ1	1.00	42			0.98
WT +PQ1	0.97	15	0.03	207	0.99
C9U-PQ1	1.00	81			0.97
C9U +PQ1	0.98	26	0.02	>1000	0.99
G12U –PQ1	0.97	23	0.03	>1000	0.98
G12U +PQ1	0.98	24	0.02	>1000	0.98
C18G –PQ1	0.98	26	0.02	>1000	0.97
C18G +PQ1	0.98	25	0.02	>1000	0.97
G37U –PQ1	0.95	12	0.05	68	0.99
G37U +PQ1	0.98	10	0.02	71	0.99
G36U/G37U –PQ1	0.98	10	0.02	85	0.99
G36U/G37U +PQ1	0.76	10	0.01	113	0.98
1U –PQ1	0.96	23	0.04	>1000	0.98
1U +PQ1	0.94	18	0.06	180	1.00
2U –PQ1	0.97	22	0.03	>1000	1.00
2U +PQ1	0.93	15	0.07	114	0.99
WT-NusA[1]	0.96	38			0.96
WT+NusA[1]	0.81	35			0.96

[1]Using DNA template purified by spermine precipitation

Table S5: Fitting Results for Transcription Time Courses of Riboswitch Variants, related to Figure 6. Results of single- or double-exponential fits to the pausing time courses shown in Figure S6.

Table S6:

RNAP variant	a1 (fraction paused)	τ1 (s) [pause half-life]	a2	τ2 (s)	R^2
R77A RNAP -PQ1	0.95	31	0.05	>1000	0.98
R77A RNAP +PQ1	0.98	17	0.02	172	1.00
K79A -PQ1	0.95	34	0.05	>1000	0.98
K79A +PQ1	0.98	18	0.02	>1000	1.00
R81A -PQ1	0.89	32	0.05	>1000	0.99
R81A +PQ1	0.80	14	0.07	185	1.00
Triple -PQ1	0.95	25	0.05	>1000	0.99
Triple -PQ1	0.97	17	0.03	444	0.99
WT RNAP/Gal1 WT template -PQ1	0.83	35	0.12	132	1.00
WT RNAP/Gal1 WT template +PQ1	0.99	11	0.01	>1000	0.98
WT RNAP/Gal1 G12U -PQ1	0.73	21	0.02	>1000	0.99
WT RNAP/Gal1 G12U +PQ1	0.97	19	0.03	>1000	0.99
ΔFT RNAP/Gal1 WT -PQ1	0.98	23	0.02	>1000	0.99
ΔFT RNAP/Gal1 WT +PQ1	0.98	8	0.02	>1000	0.86
ΔFT RNAP/Gal1 G12U -PQ1	0.98	25	0.02	>1000	0.93
ΔFT RNAP/Gal1 G12U +PQ1	0.97	19	0.03	>1000	0.97

Table S6: Fitting Results for Transcription Time Courses of RNAP Variants, related to Figure 6. Results of double-exponential fits to the pausing time courses shown in Figure S6. "Gal1" indicates that the DNA template contained the Gal1 promoter; all other assays were done with templates containing the RecA promoter.