

# Structural Basis for Control of Bacterial RNA Polymerase Pausing by a Riboswitch and its Ligand

Nils Walter (✉ [nwalter@umich.edu](mailto:nwalter@umich.edu))

University of Michigan - Ann Arbor <https://orcid.org/0000-0002-7301-1275>

Adrien Chauvier

University of Michigan-Ann Arbor <https://orcid.org/0000-0003-4473-6194>

Jason Porta

University of Michigan

Indrajit Deb

University of Michigan

Emily Ellinger

University of Michigan

Aaron Frank

University of Michigan

Melanie Ohi

University of Michigan

---

## Article

**Keywords:** Co-transcriptional folding, RNA polymerase pausing, riboswitch, cryo-electron microscopy

**Posted Date:** December 29th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1086698/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Nature Structural & Molecular Biology on June 1st, 2023. See the published version at <https://doi.org/10.1038/s41594-023-01002-x>.

# Abstract

Folding of nascent transcripts can be modulated by the proximal RNA polymerase (RNAP) that carries out their transcription, and vice versa. A pause of RNAP during transcription of a preQ1 riboswitch (que-ePEC) is stabilized by a previously characterized template consensus sequence and the ligand-free conformation of the nascent RNA. Ligand binding to the riboswitch induces RNAP pause release and downstream transcription termination, however, the mechanism by which riboswitch folding modulates pausing is unclear. Here, we report single-particle cryo-electron microscopy reconstructions of que-ePEC in ligand-free and ligand-bound states. In the absence of preQ1, the RNA transcript is in an unexpected hyper-translocated state, preventing downstream nucleotide incorporation. Strikingly, upon ligand binding the riboswitch rotates around its helical axis, expanding the surrounding RNAP exit channel and repositioning the transcript for elongation. Our study reveals the tight coupling by which small nascent RNA structures and their ligands can functionally regulate the macromolecular transcription machinery.

## Introduction

In bacteria, genetic information is transcribed by the five-subunit ( $\alpha_1\alpha_2\beta\beta'\omega$ ) protein RNA polymerase (RNAP) that forms a universal core architecture containing all basic transcription functions <sup>1</sup>. Transcription is divided into three phases: initiation, elongation and termination.

During transcription elongation, the incorporation of nucleotides into the RNA chain is coupled to the translocation of RNAP <sup>2</sup>. However, elongation is often temporarily interrupted due to an off-path state of the enzyme, which competes with the addition of the subsequent nucleotide. This so-called transcriptional pausing is universally involved in numerous biological processes including RNA folding <sup>3</sup>, transcription-translation coupling in bacteria <sup>4,5</sup>, transcription factors recruitment <sup>6</sup>, mRNA processing <sup>7,8</sup>, and is a prerequisite for transcription termination <sup>9</sup>. Transcriptional pausing is triggered when RNAP encounters a consensus DNA sequence ( $G_{-11} G_{-10} Y_{-1} G_{+1}$  in *E. coli* and  $G_{-11} G_{-10} Y_{-1} A_{+1}$  in *B. subtilis*) <sup>10,11</sup>, causing the enzyme to enter into an elemental paused state <sup>12</sup>.

The elemental paused elongation complex (ePEC) has been found to be further stabilized through at least three distinct mechanisms. Class I pauses are stabilized by the presence of a nascent RNA hairpin within the RNAP exit channel, whereas Class II pauses involve RNAP reverse translocation (or backtracking) along the DNA template <sup>13</sup>. In addition, these two transcriptional pause classes are sensitive to transcription factors that will further modulate pausing efficiency, such as N-utilization substance A (NusA) protein for Class I <sup>14</sup> and GreB for Class II pauses <sup>15</sup>. In contrast, while the consensus DNA sequence is also essential at Class III pauses, this class of transcriptional pauses was shown to be insensitive to transcription factors and instead responds to a conformational change of a nascent RNA motif termed a riboswitch, lodged within the RNAP exit channel (Figure 1A; Widom et al., 2018).

Riboswitches are structural RNA elements embedded in the 5'-untranslated regions (5'-UTR) of mRNAs that are predicted to regulate the expression of up to 4% of genes in certain bacteria <sup>17</sup>, making them an

attractive target for the design of new antibiotic drugs<sup>18,19</sup>. A typical riboswitch is composed of two interconnected domains: a conserved aptamer that binds a specific metabolite or ion termed ligand, followed by an expression platform that undergoes conformational changes upon ligand binding, altering the genetic expression level through the modulation of transcription termination or translation initiation<sup>20</sup>. In *Bacillus subtilis* (*Bsu*), the riboswitch localized upstream to the *queCDEF* operon binds the tRNA nucleotide precursor 7-methylamino-7-deazaguanine (preQ<sub>1</sub>) as a ligand and represents one of the smallest riboswitches identified to date<sup>21</sup>. This *Bsu* preQ<sub>1</sub> riboswitch operates at the transcriptional level, in which ligand binding stabilizes a H-type pseudoknot structure (the “docked” conformation) that favors the formation of a terminator hairpin in the expression platform for premature termination (Figure S1A), thus decreasing the expression level of genes involved in queuosine biosynthesis via a negative feedback loop<sup>22–25</sup>.

In previous work, we have identified a specific crosstalk between nascent riboswitch folding and RNAP processivity that is modulated by preQ<sub>1</sub> binding as a characteristic feature of the Class III pause<sup>16</sup>. RNAP pausing in the expression platform (the *que*-ePEC) is stabilized in the absence of ligand through interactions between the partially folded pseudoknot and the RNAP exit channel (Figure 1A). Ligand binding then disrupts these interactions to promote RNAP release from the paused state<sup>16</sup>. Although single molecule assays probing intramolecular riboswitch dynamics revealed that the presence of RNAP at the *que* pause significantly alters the ligand-dependent RNA conformational change, we lack a high-resolution mechanistic understanding of this new, likely widespread class of transcriptional pausing regulation mediated by coupling between nascent RNA and RNAP in cis. Here, we used single-particle cryo-electron microscopy (cryo-EM) to determine the ligand-free *que*-ePEC structure to 3.3 Å resolution and the corresponding preQ<sub>1</sub> bound *que*-ePEC structure to 3.8 Å resolution. A comparison of the two structures, in combination with biochemical assays, demonstrates that the ligand-free riboswitch in the exit channel induces retraction of the nascent RNA strand from the active site into a hyper-translocated state, explaining the inhibition of catalysis. Upon preQ<sub>1</sub> binding, the riboswitch twists, expanding the exit channel, inducing a counter-rotation of the RNAP swivel module, and jutting the RNA 3'-end back into the active site to release pausing. The pair of *que*-ePEC structures thus invokes a heretofore uncharacterized mechanism for the RNA-based control of transcriptional pausing and subsequent pause release that may apply to numerous other bacterial transcription elongation complexes (ECs) with promise for antibiotic drug screening.

## Results

### The In Vitro Assembled *que*-ePEC is Functional

*E. coli* RNAP ECs stalled at the *que* pause were formed using nucleic scaffolds comprising appropriate DNA and RNA oligonucleotides mimicking a transcription bubble (Figure 1B; Kyzer et al., 2007; Widom et al., 2018). Pause escape was measured upon rGTP addition to extend the RNA by two nucleotides (Figure 1B, C). In the presence of preQ<sub>1</sub>, pause dwell times decreased by 2-fold, consistent with our previous

observation of pause release upon ligand binding to the riboswitch during promoter-initiated transcription  
16

For structural studies, we similarly assembled the *que*-ePEC and purified it using size-exclusion chromatography (see STAR Methods). Addition of 8 mM CHAPSO into the cryo-EM buffer allowed the particles to adopt random orientations in the vitrified ice, and consistent with prior observations<sup>27</sup> the detergent did not interfere with preQ<sub>1</sub>-mediated pause release. Thus, conditions used during sample preparation are not interfering with ligand binding and riboswitch folding (Figure S1B and S1C). We conclude that the directly reconstituted *que*-ePEC affects reversible RNAP pausing, making it suitable for structural mechanistic studies via single particle cryo-EM.

## Initial Structure Determination Shows Similar Protein Conformations in the Absence and Presence of PreQ<sub>1</sub>

The cryo-EM reconstruction of the *que*-ePEC (see STAR Methods for details) revealed density attributable to the upstream and downstream DNA. Details of the active site, the RNA-DNA hybrid, and the riboswitch within the RNAP exit channel are also clearly visible in the density maps (Figure 1D, S2 and S3). 3D classification of the unmasked complexes led to two 3D classes for the *que*-ePEC in the absence of preQ<sub>1</sub>, however, the conformations of these two 3D classes were indistinguishable when overlaid and were therefore combined back into a single class that was refined to a nominal resolution of 3.3 Å (Figure S2). Despite the lack of distinct 3D classes, two discretely different structures were observed by application of 3D variability analysis (3DVA). One structure is in the non-swiveled position and has an open clamp. The second structure is swiveled and has a closed clamp conformation at the primary channel (Figure S4). Interestingly as well is the presence of density related to the riboswitch (Figure S4, right panel).

In the presence of preQ<sub>1</sub>, a single conformation was observed during 3D classification and refined to 3.8 Å (Figure S3) and unlike the sample in the absence of ligand, only a single 3DVA component was detected. Local resolution calculations indicated that much of the RNAP structures have local resolutions that span 3.0-4.5 Å and 3.0-5.5 Å in the absence and presence of preQ<sub>1</sub> respectively (Figures S2F and S3F). The core structure, which includes the RNA-DNA hybrid and the active site are the most stable regions with the best resolvability, while the RNA exit channel and the peripheral regions nearing the solvent interface of the complex are not as well resolved. This is particularly true for the β' sequence insertion 3 (β'-SI3) and β-Flap tip domains due to their solvent exposure and inherent flexibility<sup>28</sup>. The resulting two RNAP protein structures are very similar, with a RMSD of 1.6 Å, but this value represents a conformational average of the *que*-ePEC that we will further analyze in the following sections.

To bootstrap the process of model building for the *que*-ePEC structure we started by initially placing the *his*PEC RNAP<sup>28</sup> without the RNA stem loop. Some regions, including the hinge regions between helices and the nascent RNA emerging from the RNA exit channel needed to be rebuilt by hand.

### 3D Variability Analysis Reveals Global Conformational Changes within the Ligand-free *que*-ePEC

Even with reaching sub-4Å resolutions for the *que*-ePEC both in the absence and presence of preQ<sub>1</sub>, there were no obviously distinct conformations within the datasets using conventional 3D classification approaches. We wondered whether the apparent similarities between the *que*-ePEC structures could be due to a continuity of RNAP conformers within the datasets, rather than RNAP adopting static and distinct conformations as observed previously during transcription-translation coupling<sup>29</sup>. To test this idea, we submitted particles refined by both 2D classification and *ab initio* 3D classification to 3D variability analysis (3DVA) in *cryoSPARC*<sup>30</sup>. Possibly due to the limited number of available classified particles (<100,000) for the preQ<sub>1</sub>-bound *que*-ePEC dataset, 3DVA was unable to resolve conformational isomers. In contrast, the 142,410 classified particles of the ligand-free *que*-ePEC dataset were submitted to 3DVA analysis with a 5 Å resolution filter and resolved into three eigenvectors (or modes), each of which representing a 3D volume (Figure S4A).

The first mode shows conformational changes in the β'-SI3, β-Lobe, β-Flap helices, exit channel domains and the upstream DNA helix (Figure S4B, Movie S1). For this particular mode, a scissoring movement between the β'-SI3 and the β-lobe can also be observed (Movie S1). Additionally, the first mode shows movement of the exit channel domains, primarily the β'-Zinc Binding Domain (ZBD) and the β'-Dock, which are moving toward the clamp region (Figure 2). The second mode shows the most movement in the β'-SI3 domain and the upstream DNA helix, with less motion in the regions surrounding the RNAP primary channel (Movie S1). Unique to this mode is the appearance of the upstream helix as it grows out into the solvent region (Figure S4B). The third mode resolves the riboswitch within the RNAP exit channel with very little movement of the other parts of the *que*-ePEC (Movie S1); this singular motion starts without any noticeable riboswitch density and ends with density for a mature RNA transcript.

CryoSPARC was used to identify the full ranges of continuous motions along each variability dimension in our *que*-ePEC dataset. For this analysis, 20 structural intermediates were calculated (see STAR Methods for details). During the continuous motion, the last frame shows that RNAP is swiveled by ~3° relative to the initial frame about an axis centered near a hinge previously described in the *his*-PEC<sup>14,28</sup> (Figure 2). Simultaneously, the exit channel domains β'-ZBD and the β'-Dock are shifted away from the exit channel and toward the secondary DNA helix. In summary, the *que*-ePEC has large conformational freedom in the absence of ligand, possibly due to inherent dynamics of the RNA<sup>24</sup>, which is transduced to the RNAP swivel module motion stabilizing transcriptional pausing.

## Ligand Binding Induces Riboswitch Rotation Inside the RNA Exit Channel

The density associated with the P2 helix of the riboswitch was identifiable in the vicinity of the RNA exit channel as a double-helical stem structure (Figure 1D). However, since it is located in a solvent-accessible region and conformationally flexible, it was not possible to build a detailed de novo model of the entire nucleic acid. To this end, we took advantage of a previously obtained NMR structure of the same preQ<sub>1</sub>

riboswitch (Kang et al., 2009; PDB: 2L1V) and performed Molecular Dynamic Flexible Fitting (MDFF) simulations on the electron density associated with the riboswitch emerging from the RNAP exit channel (see STAR Methods for details). From 100 MDFF trajectories generated from each structure (minus and plus preQ<sub>1</sub>), the 20 best-fit models were selected based on their global cross-correlation coefficients (Figure S5A and Movies S2 and S3). Analysis of these models revealed that, upon preQ<sub>1</sub> binding, the riboswitch aptamer undergoes a significant twisting motion around the P1 helical axis within the RNAP exit channel. Specifically, ligand binding induces a ~42° rotation of the aptamer toward the β'-ZBD domain (Figure 3A), leading to a ~1.4 Å shift of the RNA 3'-end in the downstream direction of the active site (Figure 3B). Since we could not fully resolve the β-Flap tip due to its well-known flexibility due to its exposure to bulk solvent<sup>32</sup>, we asked whether its inclusion in the MDFF simulations would affect the observed RNA twisting motion. Additional models generated with the β-Flap tip included behave similarly (Figures S6 and S7), suggesting that the preQ<sub>1</sub>-induced aptamer rotation is a robust feature.

Overall, MDFF unveils a novel, ligand-induced twisting and insertion motion of the preQ<sub>1</sub> riboswitch relative to RNAP, which has not been visualized in previous structural analyses, performed with the isolated RNA<sup>22,23,25,31</sup>.

## The RNA Exit Channel Responds to the Ligand-Induced Riboswitch Rearrangement

As the newly synthesized transcript emerges from the RNAP main cleft, the first five nucleotides past the RNA:DNA hybrid (RNA residues C33, U34, A35, A36 and G37) reside in the RNAP exit channel and can potentially form duplex structures and modulate transcriptional pausing efficiency through RNA-protein interactions<sup>33,34</sup>. Additionally, the presence of a preformed positively charged surface in the RNAP exit channel complementary to an A-form helix has been suggested to constitute a path for the nascent RNA that guides the formation of RNA duplexes (Figure 4A; Kang et al., 2018). Therefore, we hypothesized that folding of the *que*-ePEC may directly affect the conformation of this sub-domain.

In the absence of preQ<sub>1</sub>, the nascent transcript is sterically hindered to form a stable RNA structure due to clashes between the β-Flap tip and the riboswitch P2 loop (RNA residues C14, U15, A16, C17, A18 and C19) and the 3'-end of the A stretch (RNA residues A27 to A32). Inside the RNAP exit channel the nucleotides C10, U11, A12 and G13 align with a path of positively charged residues involving arginine and lysine residues in the β'-ZBD specifically residue β'-R77 that could establish multiple hydrogen bonds with the nucleotide A16 (Figure 4B and Table S4). This structural observation provides molecular context for why mutation of positive residues in the β'-ZBD to alanine (R77A, K79A and R81A) significantly reduces the efficiency of the *que* pause in the absence of ligand and support the functional relevance of this interaction<sup>16</sup>

Upon ligand binding, due to the twist of the aptamer in the exit channel, the riboswitch shifts closer to the β'-ZBD domain and is in a position where it likely forms a hydrogen bond between β'-K79 and G4 (Figure 4C and Table S4). Moreover, K395 within the β'-Dock domain could now also form hydrogen bonds with

nucleotide and C33 nucleotides, and an additional contact could be formed between  $\beta'$ -K398, which is also conserved in eukaryotic RNAPII<sup>1</sup> and nucleotide C10 (Figure 4B and Table S4), suggesting that ligand binding not only stabilizes the pseudoknot but also increases the extent of RNA-protein interactions, accompanying the RNA twisting motion.

Conversely, analysis of the  $\beta$ -Flap tip,  $\beta'$ -Dock and ZBD reveals that RNAP undergoes significant movements as a function of the riboswitch docking state. In particular, the preQ<sub>1</sub>-induced twisting of the riboswitch pushes the  $\beta$ -Flap,  $\beta'$ -Dock and ZBD away from the riboswitch, effectively opening the RNAP exit channel to accommodate the nascent transcript (Figure 4C, D and Movie S4 and S5). Thus, ligand binding and the resulting folding of the nascent riboswitch have direct effects on the adjacent RNAP subdomains, providing a structural explanation for the inhibition of transcriptional pausing in the presence of preQ<sub>1</sub>.

## Distal Ligand Binding Modulates the RNAP Active Site to Release Hyper-Translocation and Pausing

We compared our two *que*-ePEC structures to the rigid core module of previously obtained EC structures<sup>28,35</sup>. In the absence of preQ<sub>1</sub>, the RNA-DNA hybrid is in an unconventional translocation state that most closely resembles the post-translocated *E. coli* RNAP complex (Kang et al., 2017; PDB: 6ALF). The RNA 3'-end is shifted  $\sim 1.6$  Å upstream from the *i* (or product) site into the *i-1* site relative to the post-translocated EC (Figure 5A). The template DNA (tDNA) still resides in a fully post-translocated state so that tDNA residue T16 remains in the *i* site while still pairing with A47 of the RNA 3'-end. The next tDNA nucleotide, C15, is not paired with the non-template DNA (ntDNA), while occupying the rNTP binding *i+1* site, but is too far from the RNA 3'-end to allow it to template the next incoming rNTP (Figure 5B; Weixlbaumer et al., 2013). To further test this model, we used exonuclease III (Exo III) footprinting of the upstream DNA border<sup>37</sup> and found that the tDNA in the *que*-ePEC indeed primarily resides in a post-translocated register (Figure S8A). Notably, the hyper-translocation of the RNA, but not the DNA, causes the hybrid to be weakened and tilted in a geometry previously suggested to occur during intrinsic or Rho-dependent termination<sup>38-40</sup>. This active site conformation is consistent with the functional disruption of the nucleotide addition cycle and provides a further structural explanation for the inhibition of catalysis in the absence of bound preQ<sub>1</sub>.

Strikingly, when the riboswitch is in the docked state, the RNA 3'-end – at an  $\sim 50$  Å distance from the bound preQ<sub>1</sub> – moves closer to the RNAP catalytic center and now occupies the *i*-site in a post-translocated register while remaining base-paired with tDNA residue T16 (Figure 5A and B). However, the tDNA is still pre-translocated because its C15 nucleotide remains base-paired with G25 in the ntDNA. A similar architecture of a half-translocated hybrid was observed in the *his*-PEC in the absence and presence of NusA transcription factor<sup>14,28</sup> supporting the notion that distinct functional pathways can lead to structurally related RNAP active sites.

In the absence of ligand, the geometry of the translocation state suggests a weak RNA-DNA hybrid that could lead to transcription termination<sup>41,42</sup>. To further probe the observed PEC translocation upon ligand binding to the riboswitch, we performed *in vitro* transcription of DNA templates in which the RNA-DNA hybrid in the context of the *que* pause was altered by replacing key nucleotides with rU:dA base-pairs to decrease its stability and favor termination of transcription (Figure 5C). Efficient termination was observed when the RNA-DNA hybrid is weak with 7 rUs out of the 9 total residues (variant 7U). As the stability of the RNA-DNA hybrid is strengthened by serially changing back these excess rU:dA base-pairs, a significant proportion of readthrough product is starting to be detected in the presence of preQ<sub>1</sub>, with the most efficient ligand-mediated anti-termination observed with the 4U variant (Figure 5D and S8B). These observations further support the notion that the ligand induces a forward RNA translocation, which stabilizes the RNA-DNA contacts with RNAP and prevents termination when the RNA-DNA hybrid is additionally weakened.

Together, our structural and functional observations support a mechanism wherein binding of preQ<sub>1</sub> triggers a series of riboswitch conformational changes relative to paused RNAP that modulates the translocation register and realigns the RNA 3'-end with the active site Mg<sup>2+</sup> (Figure 5A) to release the paused RNAP.

## Discussion

Understanding both the fundamental mechanism of transcriptional pausing and the dynamic interplay between nascent RNA folding and gene expression is essential due to their profound regulatory roles in bacteria and potential for antibiotic targeting<sup>43-46</sup>. In particular, the directional 5'-to-3' transcription of RNA is often paused to temporally and spatially program the co-transcriptional, hierarchical folding of RNA structures within or near the RNA exit channel<sup>47</sup>. We here report high-resolution cryo-EM structures of elongation complexes at the riboswitch-mediated Class III *que* pause site (*que*-ePEC) in the absence and presence of cognate ligand at 3.3 Å and 3.8 Å global resolutions, respectively, revealing how RNA and RNAP affect each other's conformation in a ligand dependent mechanism (Figure 4). Binding of the 252 Da preQ<sub>1</sub> ligand to the aptamer domain triggers a cascade reaction, initiated within the RNAP exit channel (Figure 3), that propagates downstream to the active site of the ~400 kDa RNAP (Figure 5) to ultimately release the enzyme from the paused state (Figure 6). With the riboswitch embedded within the RNAP exit channel, MDFF analysis reveals an intriguing RNA folding pathway (Movies S4 and S5) initiated by ligand binding that allosterically alters the global RNAP conformation (Figure 2 and 4). Our structural and biochemical studies rationalize how RNA arrests catalysis in the absence of preQ<sub>1</sub>, and how ligand-induced riboswitch folding reactivates RNAP from its paused state, thereby providing direct mechanistic insights into Class III transcriptional pausing regulation.

The RNAP exit channel is known to act as the gateway for regulatory processes that occur co-transcriptionally as a result of interactions with the translational machinery during transcription-translation coupling<sup>4,48</sup>, transcription factors<sup>14,49,50</sup>, and folding of RNA structures within it<sup>33</sup>. In

addition, leveraging its positive charges that align with the phosphates of A-form RNA, the RNAP exit channel forms a route that guides functional RNA duplex formation, acting as a basis for a regulatory connection with the downstream active site<sup>28</sup>.

In the absence of ligand, the 5'-segment of the P2 helix (specifically RNA residue C17), is found in close proximity to the  $\beta$ -Flap tip helix and threads toward the  $\beta'$ -ZBD domain (Figure 4), as predicted by molecular dynamic simulation<sup>16</sup>. Accordingly, both deletion of the  $\beta$ -Flap tip (residues 890-914) and point mutations in the RNA-binding region of the  $\beta'$ -ZBD (residues R77A, K79A and R81A) decrease the *que* pause efficiency substantially, supporting the functional relevance of this interaction<sup>16</sup>. The 3'-segment of P2 (in particular RNA residue U34) resides on the proximal face of the  $\beta'$ -Dock domain (contacting residue K395; Figure 4), in agreement with previous studies showing the equivalent interactions between the 5'-segment of the *his* paused RNA hairpin and RNAP<sup>28</sup>. Therefore, following translocation arrest at the elemental pause site, the ligand-free aptamer pseudoknot appears to constitute a physical barrier preventing RNA extrusion from the polymerase exit channel, thereby holding RNAP in the swiveled conformation (Figure 2). Conversely, this particular feature may also contribute to the fast response of the *que*-ePEC to ligand binding since the pseudoknot structure is held in place by RNAP for dynamic sensing of the preQ<sub>1</sub> ligand<sup>16,51,52</sup>.

Following ligand binding to the aptamer, the riboswitch rotates  $\sim 42^\circ$  along its P1 helical axis (Figure 3) and alters the *que*-ePEC at multiple architectural levels, leading to the release of transcriptional pausing. First, preQ<sub>1</sub>-induced stabilization of the P2 helix remodels RNAP, causing its RNA exit channel subdomains to shift away, effectively "opening" the path and providing clearance for the nascent RNA transcript (Figure 4 and Movies S4 and S5). To this end, the  $\beta$ -Flap moves away, expanding the channel's inner diameter, as has also been seen to allow for the accommodation of a more simple regulatory RNA hairpin<sup>32</sup>. In addition, the  $\beta'$ -Dock and ZBD shift, counteracting with the swivel module rotation, as proposed for an active elongation complex<sup>12,35</sup>. In the *que*-ePEC, this remodeling is expected to expose the RNAP binding surface for NusA domain S1 (Guo et al., 2018; Figure S9A, B). In support of this observation, we find that the *que*-ePEC is sensitive to NusA only in the presence of preQ<sub>1</sub> (Figure S9C, D).

Next, extensive contacts and physical proximity of the riboswitch L2 loop to the  $\beta'$ -Dock and ZBD rationalize the impact of riboswitch folding has on the global RNAP conformation. In the absence of preQ<sub>1</sub>, as P2's 3'-segment is abutted by the  $\beta'$ -Dock domain, stabilization of the preQ<sub>1</sub>-folded pseudoknot coupled with riboswitch twisting promotes pause release by pushing the swivel module toward the elongation-active non-swiveled conformation to operate through an induce-fit mechanism (Suddala et al., 2015; Figure 6). While this renders our *que*-ePEC structure in the docked riboswitch state similar to the canonical PEC at a Class I pause<sup>28</sup>, this anti-swiveling motion also occurs concomitantly with nucleotide addition and therefore constitutes a structural intermediate during transcription reactivation. In support of these observations, the  $\beta'$ -ZBD domain, to which most of the RNA contacts are detected in the presence of preQ<sub>1</sub>, previously has been implicated in stabilizing the EC<sup>53,54</sup> and in modulating transcription termination<sup>55-58</sup>.

Several mechanisms have been proposed for how RNAP translocates along the DNA template during transcription elongation<sup>2,59,60</sup>. An obstacle to this movement is presented by structural intermediates that are transient and difficult to capture because of their inherent dynamics. Such intermediates have been suggested to also occur in yeast RNAP II bound to  $\alpha$ -amanitin and in viral RNAP<sup>61,62</sup> and were recently observed in paused bacterial RNAP complexes. For example, an asymmetric movement of the RNA-DNA hybrid (half-translocation) explains the stabilization of RNAP pausing at hairpin-stabilized Class I pauses such as the canonical pause in the *his* operon<sup>14,28</sup>. Strikingly, even if the tDNA is post-translocated in the absence of ligand, the *que*-ePEC RNA is hyper-translocated, a geometry preventing subsequent nucleotide incorporation (Figure 5). Importantly, a weaker RNA-DNA hybrid converts the *que*-ePEC into a “pre-terminated” complex (Figure 5C,D and S8B) suggesting that the hyper-translocated intermediate is also present at the early steps of intrinsic or protein-mediated transcription termination<sup>50,63</sup>. During intrinsic termination, the asymmetric movement of the RNA would directly follow terminator hairpin nucleation in the hybrid-shearing mechanism<sup>9,42</sup>. However, further upstream melting of the hybrid in the *que*-ePEC is disfavored due to the lack of rU:dA base pairs in the RNA-DNA hybrid. This creates an energy barrier too high for pseudoknot stabilization at the edge of the RNA-DNA hybrid, leading to strong transcriptional pausing instead of transcription termination (Figure 5C and D).

To date, only RNA hairpin folding within the RNAP exit channel has been found to allosterically alter RNAP structure in the context of transcriptional pausing and termination<sup>33,64,65</sup>. For example, folding of an RNA hairpin within the RNAP exit channel could favor forward translocation of the enzyme by pulling out the nascent transcript in the RNA-DNA hybrid, which has been proposed to help rescue backtracked RNAPs<sup>47</sup>. The ligand-mediated regulation of Class III pauses operates through a distinct mechanism in which riboswitch folding in the ligand-bound conformation induces a forward RNA translocation, pushing the RNA 3'-end one nucleotide closer to the RNAP active site (Figure 5A and B). A similar transcription reactivation mechanism has been reported for the ATPase and DNA translocase RapA<sup>66</sup>, indicating that the *que*-ePEC pause release mechanism is broader and may have been operational as early as during the RNA World<sup>67</sup>.

The *que*-ePEC is still competent for ligand binding to the aptamer even if the *que* pause is situated in the expression platform, therefore, it may constitute a transcriptional checkpoint for late ligand binding events<sup>51</sup>. Because the riboswitch is very small and will be transcribed *in vivo* in a short amount of time, pause stabilization in the absence of preQ<sub>1</sub> will allow more time for the aptamer to recognize and bind low concentrations of the ligand, preventing premature transcription readthrough more efficiently<sup>68</sup>. This regulatory mechanism resembles that of a molecular fuse that helps promote an accurate gene regulatory response to a small molecular weight metabolite and can be compared to the burst of transcription by RNAP II when released from promoter-proximal pauses in response to environmental and cellular cues in eukaryotes<sup>69</sup>. Overall, the high-resolution structures and biochemical evidence obtained here reveal the adaptability of bacterial RNAP to funnel a variety of distinct molecular inputs into

achieving a desired gene regulatory output and promise to guide the development of novel antibacterial therapies against the transcription machinery.

## Declarations

### SUPPLEMENTAL INFORMATION

Supplemental information includes 9 figures, 4 tables and 5 movies and be found online.

### ACKNOWLEDGMENTS

The authors would like to thank Dr. Julia Widom for the initial design of the *que*-ePEC reconstitution, Dr. Irina Artsimovitch for helpful discussions, and UM BSI for support of the U-M cryo-EM facility. This work was supported by National Institutes of Health (NIH) R01 grants GM131922 and GM118524 to N.G.W and NIH S100D020011 and S100D030275 to M.D.O. A portion of the molecular graphics and analyses was performed with UCSF Chimera and ChimeraX developed by the Resource for Biocomputing, Visualization, and Informatics at UC-San Francisco, with support from NIH P41-GM103311.

### AUTHOR CONTRIBUTIONS

Conceptualization, A.C., M.D.O., N.G.W.; Methodology, A.C., J.P, I.D.; Investigation, A.C., J.P, I.D., E.E.; Writing - Original Draft, A.C., J.P; Writing - Review & Editing, A.C., J.P, I.D., E.E., M.D.O., A.T.F, N.G.W.; Supervision, A.T.F, M.D.O., N.G.W; Funding Acquisition, A.T.F., M.D.O., N.G.W.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

### CONTACT FOR REAGENT AND RESSOURCE SHARING

Further information and requests for resource and reagent should be directed and will be fulfilled by the Lead Contact, Nils G. Walter ([nwalter@umich.edu](mailto:nwalter@umich.edu)).

## References

1. Lane, W. J. & Darst, S. A. Molecular evolution of multisubunit RNA polymerases: structural analysis. *J. Mol. Biol.* **395**, 686–704 (2010).

2. Bar-Nahum, G. *et al.* A ratchet mechanism of transcription elongation and its control. *Cell* **120**, 183–193 (2005).
3. Pan, T. & Sosnick, T. RNA folding during transcription. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 161–175 (2006).
4. Chatterjee, S., Chauvier, A., Dandpat, S. S., Artsimovitch, I. & Walter, N. G. A translational riboswitch coordinates nascent transcription–translation coupling. *Proc. Natl. Acad. Sci.* **118**, (2021).
5. Proshkin, S., Rahmouni, A. R., Mironov, A. & Nudler, E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **328**, 504–508 (2010).
6. Artsimovitch, I. & Landick, R. The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. *Cell* **109**, 193–203 (2002).
7. Alexander, R. D., Innocente, S. A., Barrass, J. D. & Beggs, J. D. Splicing-dependent RNA polymerase pausing in yeast. *Mol. Cell* **40**, 582–593 (2010).
8. Mayer, A. *et al.* Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell* **161**, 541–554 (2015).
9. Gusarov, I. & Nudler, E. The mechanism of intrinsic transcription termination. *Mol. Cell* **3**, 495–504 (1999).
10. Larson, M. H. *et al.* A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science* **344**, 1042–1047 (2014).
11. Vvedenskaya, I. O. *et al.* Transcription. Interactions between RNA polymerase and the ‘core recognition element’ counteract pausing. *Science* **344**, 1285–1289 (2014).
12. Kang, J. Y., Mishanina, T. V., Landick, R. & Darst, S. A. Mechanisms of Transcriptional Pausing in Bacteria. *J. Mol. Biol.* **431**, 4007–4029 (2019).
13. Artsimovitch, I. & Landick, R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7090–7095 (2000).
14. Guo, X. *et al.* Structural Basis for NusA Stabilized Transcriptional Pausing. *Mol. Cell* **69**, 816–827.e4 (2018).
15. Abdelkareem, M. *et al.* Structural Basis of Transcription: RNA Polymerase Backtracking and Its Reactivation. *Mol. Cell* **75**, 298–309.e4 (2019).
16. Widom, J. R. *et al.* Ligand Modulates Cross-Coupling between Riboswitch Folding and Transcriptional Pausing. *Mol. Cell* **72**, 541–552.e6 (2018).

17. Sherwood, A. V. & Henkin, T. M. Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. *Annu. Rev. Microbiol.* **70**, 361–374 (2016).
18. Mulhbach, J. *et al.* Novel riboswitch ligand analogs as selective inhibitors of guanine-related metabolic pathways. *PLoS Pathog.* **6**, e1000865 (2010).
19. Wang, H. *et al.* Dual-Targeting Small-Molecule Inhibitors of the Staphylococcus aureus FMN Riboswitch Disrupt Riboflavin Homeostasis in an Infectious Setting. *Cell Chem. Biol.* **0**, (2017).
20. Bédard, A.-S. V., Hien, E. D. M. & Lafontaine, D. A. Riboswitch regulation mechanisms: RNA, metabolites and regulatory proteins. *Biochim. Biophys. Acta Gene Regul. Mech.* **1863**, 194501 (2020).
21. Roth, A. *et al.* A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. *Nat. Struct. Mol. Biol.* **14**, 308–317 (2007).
22. Rieder, U., Kreutz, C. & Micura, R. Folding of a transcriptionally acting preQ1 riboswitch. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 10804–10809 (2010).
23. Schroeder, G. M. *et al.* Analysis of a preQ1-I riboswitch in effector-free and bound states reveals a metabolite-programmed nucleobase-stacking spine that controls gene regulation. *Nucleic Acids Res.* **48**, 8146–8164 (2020).
24. Suddala, K. C. *et al.* Single transcriptional and translational preQ1 riboswitches adopt similar pre-folded ensembles that follow distinct folding pathways into the same ligand-bound structure. *Nucleic Acids Res.* **41**, 10462–10475 (2013).
25. Suddala, K. C., Wang, J., Hou, Q. & Walter, N. G. Mg(2+) shifts ligand-mediated folding of a riboswitch from induced-fit to conformational selection. *J. Am. Chem. Soc.* **137**, 14075–14083 (2015).
26. Kyzer, S., Ha, K. S., Landick, R. & Palangat, M. Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. *J. Biol. Chem.* **282**, 19020–19028 (2007).
27. Chen, J., Noble, A. J., Kang, J. Y. & Darst, S. Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *bioRxiv* 457267 (2018) doi:10.1101/457267.
28. Kang, J. Y. *et al.* RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. *Mol. Cell* **69**, 802-815.e1 (2018).
29. Webster, M. W. *et al.* Structural basis of transcription-translation coupling and collision in bacteria. *Science* **369**, 1355–1359 (2020).

30. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
31. Kang, M., Peterson, R. & Feigon, J. Structural Insights into riboswitch control of the biosynthesis of queuosine, a modified nucleotide found in the anticodon of tRNA. *Mol. Cell* **33**, 784–790 (2009).
32. Touloukhonov, I. & Landick, R. The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination. *Mol. Cell* **12**, 1125–1136 (2003).
33. Hein, P. P. *et al.* RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. *Nat. Struct. Mol. Biol.* **21**, 794–802 (2014).
34. Kolb, K. E., Hein, P. P. & Landick, R. Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH. *J. Biol. Chem.* **289**, 1151–1163 (2014).
35. Kang, J. Y. *et al.* Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex. *eLife* **6**, (2017).
36. Weixlbaumer, A., Leon, K., Landick, R. & Darst, S. A. Structural basis of transcriptional pausing in bacteria. *Cell* **152**, 431–441 (2013).
37. Nedialkov, Y. A., Nudler, E. & Burton, Z. F. RNA polymerase stalls in a post-translocated register and can hyper-translocate. *Transcription* **3**, 260–269 (2012).
38. Larson, M. H., Greenleaf, W. J., Landick, R. & Block, S. M. Applied force reveals mechanistic and energetic details of transcription termination. *Cell* **132**, 971–982 (2008).
39. Park, J.-S. & Roberts, J. W. Role of DNA bubble rewinding in enzymatic transcription termination. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4870–4875 (2006).
40. Santangelo, T. J. & Roberts, J. W. Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Mol. Cell* **14**, 117–126 (2004).
41. Kent, T., Kashkina, E., Anikin, M. & Temiakov, D. Maintenance of RNA-DNA hybrid length in bacterial RNA polymerases. *J. Biol. Chem.* **284**, 13497–13504 (2009).
42. Komissarova, N., Becker, J., Solter, S., Kireeva, M. & Kashlev, M. Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Mol. Cell* **10**, 1151–1162 (2002).
43. Bevilacqua, P. C., Ritchey, L. E., Su, Z. & Assmann, S. M. Genome-Wide Analysis of RNA Secondary Structure. *Annu. Rev. Genet.* **50**, 235–266 (2016).

44. Landick, R. Transcriptional Pausing as a Mediator of Bacterial Gene Regulation. *Annu. Rev. Microbiol.* (2021) doi:10.1146/annurev-micro-051721-043826.
45. Lewis, C. J. T., Pan, T. & Kalsotra, A. RNA modifications and structures cooperate to guide RNA-protein interactions. *Nat. Rev. Mol. Cell Biol.* **18**, 202–210 (2017).
46. Liu, X., Kraus, W. L. & Bai, X. Ready, pause, go: regulation of RNA polymerase II pausing and release by cellular signaling pathways. *Trends Biochem. Sci.* **40**, 516–525 (2015).
47. Zhang, J. & Landick, R. A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure. *Trends Biochem. Sci.* **41**, 293–310 (2016).
48. Wang, C. *et al.* Structural basis of transcription-translation coupling. *Science* **369**, 1359–1365 (2020).
49. Huang, Y.-H. *et al.* Structure-Based Mechanisms of a Molecular RNA Polymerase/Chaperone Machine Required for Ribosome Biosynthesis. *Mol. Cell* **79**, 1024-1036.e5 (2020).
50. Said, N. *et al.* Steps toward translocation-independent RNA polymerase inactivation by terminator ATPase  $\rho$ . *Science* **371**, (2021).
51. Chauvier, A. *et al.* Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation. *Nat. Commun.* **8**, 13892 (2017).
52. Suddala, K. C. *et al.* Hierarchical mechanism of amino acid sensing by the T-box riboswitch. *Nat. Commun.* **9**, 1896 (2018).
53. Komissarova, N. *et al.* Inhibition of a transcriptional pause by RNA anchoring to RNA polymerase. *Mol. Cell* **31**, 683–694 (2008).
54. Nudler, E., Gusarov, I., Avetissova, E., Kozlov, M. & Goldfarb, A. Spatial organization of transcription elongation complex in *Escherichia coli*. *Science* **281**, 424–428 (1998).
55. King, R. A., Markov, D., Sen, R., Severinov, K. & Weisberg, R. A. A conserved zinc binding domain in the largest subunit of DNA-dependent RNA polymerase modulates intrinsic transcription termination and antitermination but does not stabilize the elongation complex. *J. Mol. Biol.* **342**, 1143–1154 (2004).
56. Krupp, F. *et al.* Structural Basis for the Action of an All-Purpose Transcription Anti-termination Factor. *Mol. Cell* **74**, 143-157.e5 (2019).
57. Sen, R., King, R. A. & Weisberg, R. A. Modification of the properties of elongating RNA polymerase by persistent association with nascent antiterminator RNA. *Mol. Cell* **7**, 993–1001 (2001).
58. Sen, R., King, R. A., Mzhavia, N., Madsen, P. L. & Weisberg, R. A. Sequence-specific interaction of nascent antiterminator RNA with the zinc-finger motif of *Escherichia coli* RNA polymerase. *Mol. Microbiol.*

46, 215–222 (2002).

59. Belogurov, G. A. & Artsimovitch, I. The Mechanisms of Substrate Selection, Catalysis, and Translocation by the Elongating RNA Polymerase. *J. Mol. Biol.* **431**, 3975–4006 (2019).
60. Komissarova, N. & Kashlev, M. RNA polymerase switches between inactivated and activated states By translocating back and forth along the DNA and the RNA. *J. Biol. Chem.* **272**, 15329–15338 (1997).
61. Brueckner, F. & Cramer, P. Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nat. Struct. Mol. Biol.* **15**, 811–818 (2008).
62. Shu, B. & Gong, P. Structural basis of viral RNA-dependent RNA polymerase catalysis and translocation. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E4005-4014 (2016).
63. Ray-Soni, A., Bellecourt, M. J. & Landick, R. Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annu. Rev. Biochem.* **85**, 319–347 (2016).
64. Epshtein, V., Cardinale, C. J., Ruckenstein, A. E., Borukhov, S. & Nudler, E. An allosteric path to transcription termination. *Mol. Cell* **28**, 991–1001 (2007).
65. Touloukhonov, I., Artsimovitch, I. & Landick, R. Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* **292**, 730–733 (2001).
66. Liu, B., Zuo, Y. & Steitz, T. A. Structural basis for transcription reactivation by RapA. *Proc. Natl. Acad. Sci.* **112**, 2006–2010 (2015).
67. Breaker, R. R. Riboswitches and the RNA world. *Cold Spring Harb. Perspect. Biol.* **4**, (2012).
68. Wickiser, J. K., Winkler, W. C., Breaker, R. R. & Crothers, D. M. The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol. Cell* **18**, 49–60 (2005).
69. Larsson, A. J. M. *et al.* Genomic encoding of transcriptional burst kinetics. *Nature* **565**, 251–254 (2019).
70. Scheres, S. H. W. Amyloid structure determination in RELION-3.1. *Acta Crystallogr. Sect. Struct. Biol.* **76**, 94–101 (2020).
71. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, (2018).
72. Carragher, B. *et al.* Legimon: an automated system for acquisition of images from vitreous ice specimens. *J. Struct. Biol.* **132**, 33–45 (2000).
73. Zhang, F. *et al.* A Two-Phase Improved Correlation Method for Automatic Particle Selection in Cryo-EM. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **14**, 316–325 (2017).

74. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
75. Wagner, T. *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 218 (2019).
76. Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 531–544 (2018).
77. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
78. Davis, I. W., Murray, L. W., Richardson, J. S. & Richardson, D. C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615-619 (2004).
79. McGreevy, R., Teo, I., Singharoy, A. & Schulten, K. Advances in the molecular dynamics flexible fitting method for cryo-EM modeling. *Methods San Diego Calif* **100**, 50–60 (2016).
80. Trabuco, L. G., Villa, E., Mitra, K., Frank, J. & Schulten, K. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Struct. Lond. Engl. 1993* **16**, 673–683 (2008).
81. Trabuco, L. G., Villa, E., Schreiner, E., Harrison, C. B. & Schulten, K. Molecular dynamics flexible fitting: a practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods San Diego Calif* **49**, 174–180 (2009).
82. Best, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone  $\varphi$ ,  $\psi$  and side-chain  $\chi(1)$  and  $\chi(2)$  dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
83. Denning, E. J., Priyakumar, U. D., Nilsson, L. & Mackerell, A. D. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J. Comput. Chem.* **32**, 1929–1943 (2011).
84. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
85. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, 27–28 (1996).
86. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).

87. Landick, R., Wang, D. & Chan, C. L. Quantitative analysis of transcriptional pausing by Escherichia coli RNA polymerase: his leader pause site as paradigm. *Methods Enzymol.* **274**, 334–353 (1996).
88. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci. Publ. Protein Soc.* **30**, 70–82 (2021).
89. Vassylyev, D. G. *et al.* Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**, 163–168 (2007).

## Method Details

### RNAP Expression and Purification

*E. coli* core RNAP bearing an AviTag biotinylation tag on the C-terminus of the  $\beta'$  subunit was prepared as described previously<sup>33</sup>. Glycerol was added to the purified RNAP to 25% (v/v) prior to flash-frozen in liquid nitrogen. Aliquots were stored at -80°C until use.

### Preparation of Paused Elongation Complex for Cryo-EM

The *que*-ePEC RNA was *in vitro* transcribed using T7 RNAP transcription and gel purified before use. Synthetic DNA oligonucleotides were obtained from IDT (Integrated DNA Technologies). The nucleic scaffold (tDNA, ntDNA and RNA) was annealed at a 1:1:1 ratio in a buffer containing 100 mM KCl and 50 mM Tris-HCl pH 7.5 (90°C for 2 min, 37°C for 10 min and RT for 10 min). RNAP core was mixed to the nucleic scaffold at a molar ratio of 1:3 in assembly buffer (100 mM KCl, 50 mM Tris-HCl pH 7.5 and 1 mM MgCl<sub>2</sub>) and incubated for 15 min at 37°C. CHAPSO (Sigma Aldrich, CAS: 82473-24-3) was added, and the complex was concentrated by centrifugal filtration (Amicon – 100 kDa column) to 5 – 6 mg/mL RNAP concentration before grid preparation. PreQ<sub>1</sub> ligand, when present, was added to 10  $\mu$ M final concentration and incubated with the complex for 5 min at RT.

### Cryo-EM Data Collection and Processing

For both the (-) and (+) preQ<sub>1</sub> samples, 3.5  $\mu$ l of the complex were applied to a glow discharged C-flat 400-mesh Au grid (Ted Pella). The sample was vitrified by plunge freezing in a liquid ethane slurry using a Thermo Fisher Vitrobot at 4 °C and 100% humidity. All images for each dataset were collected on a Titan Krios electron microscope (Thermo Fisher) equipped with a K2 Summit direct electron detector (Gatan) operated at 300 keV and a nominal pixel size of 1.01 Å per pixel. Images were acquired using Legicon software<sup>72</sup>. The total exposure time was 8 s and frames were recorded every 0.2 s, giving an accumulated dose of 62 e<sup>-</sup> Å<sup>-2</sup> using a defocus range of -0.5  $\mu$ m to -3.5  $\mu$ m.

Raw movie frames pertaining to each dataset were dose-weighted and corrected for beam-induced drift using MotionCor2<sup>73</sup>. The contrast transfer function (CTF) parameters were determined using CTFIND4<sup>74</sup>. All image processing was done in RELION 3.0 and cryoSPARC<sup>30,71</sup>. Following the determination of the CTF parameters, non-template-based particle picking was done in crYOLO<sup>75</sup>. Particles were extracted into 300 px<sup>2</sup> boxes and imported into cryoSPARC for 2D classification. These particles were used to generate an initial 3D reconstruction *ab initio* into two classes to further separate out 'junk' particles. Prior to 3D refinement, per-particle drift correction was carried out. The *ab initio* 3D volume was refined against the particle data while correcting for the CTF higher-order aberrations using the homogenous refinement procedure in cryoSPARC<sup>30</sup>. Following homogenous refinement, the data was further refined using the non-uniform refinement procedure. Gold standard global FSC calculations gave resolutions of 3.65 Å and 3.9 Å in the absence and presence of preQ<sub>1</sub> respectively. These consensus volumes were subjected to 3D classification without image alignment and without applying a mask in RELION 3.0<sup>71</sup>. After this round of 3D classification, identical classes were recombined and subjected to a second round of alignment-free classification (see Figure S2 and S3). The best 3D classes for each dataset were then refined with FSC resolutions of 3.3 Å and 3.8 Å in the absence and presence of preQ<sub>1</sub>, respectively.

To better resolve the area around the RNAP exit channel, signal subtraction for each particle was carried out using a soft mask of the areas outside of this region. The subtracted particles were then subjected to alignment-free focused classification and refinement.

## Model Building, Refinement and Validation

The protein coordinates from the post-translocated elongation complex<sup>35</sup> were used as a starting point for coordinate fitting and refinement. The coordinates were first rigidly fit into the EM density map while defining each RNAP subunit as its own rigid body using the real-space refinement procedure in PHENIX<sup>76</sup>. Nucleic acid was built into the density in Coot<sup>77</sup>. All-atom refinement of both protein and nucleic acids was carried out while imposing both secondary structure and Ramachandran restraints. Finally, the fitted coordinates were validated in MolProbity<sup>78</sup>.

## Molecular Dynamics Flexible Fitting (MDFF) Simulations

The molecular dynamics flexible fitting (MDFF) is a simulation approach to flexibly fit atomic coordinates into EM maps<sup>79–81</sup>.

The MDFF integrates the EM density map as a potential so that high density areas in the map are minima on the potential energy surface. To achieve this, guiding forces are applied to the atoms in a molecular

system that are proportional to the gradient of the EM map potential. Specifically, the EM map potential, which is defined on 3D grid, is given by:

$$U_{EM}(R) = \sum_j w_j V_{EM}(r_j),$$

where  $j$  runs over the atoms in the system and

$$V_{EM}(r) = \begin{cases} \xi \left[ 1 - \frac{\Phi(r) - \Phi_{thr}}{\Phi_{max} - \Phi_{thr}} \right] & \text{if } \Phi(r) \geq \Phi_{thr}, \\ \xi & \text{if } \Phi(r) < \Phi_{thr}. \end{cases}$$

Here  $w_j$  are atomwise weights,  $\xi$  is a force scaling,  $\Phi(r)$  is the EM density at position  $r$ ,  $\Phi_{max}$  is the maximum value of the EM density map, and  $\Phi_{thr}$  is a density threshold. The density threshold serves to eliminate EM data corresponding to the solvent contribution to the map. The actual MDFF guiding forces that bring the structure in correspondence with the EM density map is given by:

$$f_i^{EM} = - \frac{\partial}{\partial r_i} U_{EM}(R) = -w_i \frac{\partial}{\partial r_i} V_{EM}(r_i)$$

The initial structures of the *que*-ePEC contained the template DNA and a 10-residues long downstream riboswitch RNA forming the RNA-DNA hybrid but lacked the complete coordinates for the residues corresponding to the preQ<sub>1</sub> riboswitch. Complete models of the *que*-ePEC in the absence and presence of preQ<sub>1</sub> were generated by attaching a structural model of the 37-nt preQ<sub>1</sub> riboswitch to the 5' end of the RNAs present in the initial structure. The structural model of the riboswitch corresponded to conformer 1 in the NMR structure of a Class I preQ<sub>1</sub> riboswitch aptamer bound to its cognate ligand (Kang et al., 2009; PDB: 2L1V). In the case of the model in the absence of preQ<sub>1</sub>, the ligand present in the NMR structure was removed. Missing hydrogens and terminal patches were added to the initial models using the CHARMM-GUI web server. The resulting structures were energy minimized in vacuum with 100 steps of steepest descent (SD) and 500 steps of adopted basis Newton-Raphson (ABNR) method with a gradient tolerance of 0.01 to remove the initial clashes. A cutoff of 16 Å was used to generate the non-bonded list. The non-bonded list was updated heuristically. Switching function was used to treat the Lennard-Jones and electrostatic interactions within the complexes. During minimization, the heavy atoms in protein backbone (C, O, N, and Ca), nucleic acid backbone (P, O1P, O2P, O5', C5', C4', C3' and O3'), and the ligand were harmonically restrained with a force constant of 1.0 kcal/mol/Å<sup>2</sup> and the heavy atoms in the protein side-chain and the sugar and base of the nucleic acid residues were harmonically restrained using a force constant 0.1 kcal/mol/Å<sup>2</sup>. During CHARMM energy minimization and subsequent MDFF simulations (see

below), we used the CHARMM36 force field<sup>82,83</sup> for the protein and nucleic acid components of the Paused Elongation Complex, and the CHARMM general force field (CGenFF)<sup>84</sup> for the preQ<sub>1</sub> ligand.

The energy minimized coordinates of the initial models, along with the cryo-EM density maps into which the initial structures were already docked, were used to set up the MDFF simulations. The input files for the MDFF simulations were generated using the MDFF plugin within visual molecular dynamics (VMD) software<sup>85</sup>. These files included restraints parameters that we used to maintain the secondary structure of the individual components in the complex and prevent overfitting during the MDFF simulations. The MDFF simulations were run in vacuum for 2 ns with a time step of 1 fs using NAMD<sup>86</sup>. Before the production run, 1000 steps of energy minimizations were performed. The temperature was maintained using the Langevin thermostat at 300K. During the initial MDFF simulations, the coordinates of the protein components, template DNA, and 10-residue long downstream riboswitch RNA that forms the RNA-DNA hybrid in the initial structures were held fixed. Only the upstream 37 residues of the RNA were free to move during the simulations. For all the MDFF simulations, atomwise weights  $w_j$  were set to atomic masses, the scaling factor ( $\xi$ ) was set to 0.3, and  $\Phi_{\text{thr}}$  is a density threshold. For both conditions (-/+ preQ<sub>1</sub>), 100 independent MDFF trajectories were generated.

## Pausing Assays

A 148-nucleotides DNA template including the preQ<sub>1</sub> riboswitch from *B. subtilis* under the control of the T7A1 promoter was generated using an overlapping PCR strategy. In addition, 25 nucleotides not found in the wild-type sequence were inserted after the promoter sequence in order to generate a 25-nucleotide stretch in which the RNA transcript lacks any uracil residues (EC-25) except for the +2-position dependent of the ApU dinucleotide used to initiate the transcription. Transcription templates for *in vitro* transcription were generated by PCR using the “T7A1-PreQ1-RNA0p (1)” forward oligonucleotide and the complementing reverse oligonucleotides (Table S1). Oligonucleotides used in this study are also listed in the Key Resources Table.

Halted complexes were prepared in transcription buffer (20 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM MgCl<sub>2</sub>, 14 mM 2-mercaptoethanol, 0.1 mM EDTA) containing 25  $\mu$ M ATP/CTP mix, 50 nM  $\alpha^{32}$ P-GTP (3000 Ci/mmol), 10  $\mu$ M ApU dinucleotide primer (Trilink, O-31004), and 50 nM DNA template. 100 nM *E. coli* RNAP holoenzyme (New England Biolabs, M0551S) was added to the reaction mixture and incubated for 10 min at 37°C. The reaction mixture was then passed through a G50 (GE Healthcare, 27533001) gel filtration column to remove any free nucleotides. To complete the transcription reaction, a mixture containing all four rNTPs (25  $\mu$ M for time pausing experiments and 100  $\mu$ M for termination assays) was added concomitantly with heparin (450  $\mu$ g/mL) to prevent the re-initiation of transcription. preQ<sub>1</sub> (when present) was added to 10  $\mu$ M or ranged from 100 nM to 250  $\mu$ M for the T50 determination in Figure S5B. The reaction mixture was incubated at 37°C, and aliquots were quenched at the desired times into an equal volume of loading buffer (95% formamide, 1 mM EDTA, 0.1% SDS, 0.2% bromophenol blue, 0.2%

xylene cyanol). Reaction aliquots were denatured before loading 5  $\mu$ L each onto a denaturing 8 M urea, 6% polyacrylamide sequencing gel. The gel was dried and exposed to a phosphor screen (typically overnight), which was then scanned on an Amersham Typhoon PhosphorImager (GE Lifesciences). Gel images were analyzed with ImageLab (Bio-Rad) software.

### **Time pausing analysis**

The half-life of transcriptional pausing was determined by calculating the fraction of the RNA pause species compared with the total amount of RNA for each time point, which was analyzed with pseudo-first-order kinetics to extract the half-life<sup>87</sup>. For each determination we subtracted the background signal. Error bars in transcription quantification represent the standard deviation of the mean from at least two independent replicates.

### **Exonuclease III footprinting**

The *que*-ePEC was assembled to 50 nM final RNAP concentration as described previously; the RNAP bears an AviTag biotinylation tag on the C-terminus of the  $\beta'$  subunit for subsequent immobilization on magnetic streptavidin beads<sup>16</sup>. Depending on whether mapping was from upstream or downstream, the tDNA (upstream border mapping) or ntDNA (downstream border mapping) was 5'-<sup>32</sup>P-labeled for detection on gels. *E. coli* Exo III was added to 4 units/ $\mu$ L in the same buffer as for Cryo-EM data acquisition. Control experiments using 5'-<sup>32</sup>P-labeled RNA were performed separately. The reaction mixture was incubated at RT, and aliquots were quenched at the desired times into an equal volume of loading buffer (95% formamide, 1 mM EDTA, 0.1% SDS, 0.2% bromophenol blue, 0.2% xylene cyanol). Reaction aliquots were denatured before loading 5  $\mu$ L each onto a denaturing 8 M urea, 6% polyacrylamide sequencing gel. The gel was dried and exposed to a phosphor screen (typically overnight), which was then scanned on an Amersham Typhoon PhosphorImager (GE Lifesciences). Gel images were analyzed with ImageLab (Bio-Rad) software.

### **DATA AND CODE AVAILABILITY**

The cryo-EM volumes have been deposited in the Electron Microscopy Data Bank

## **Figures**

Figure 1

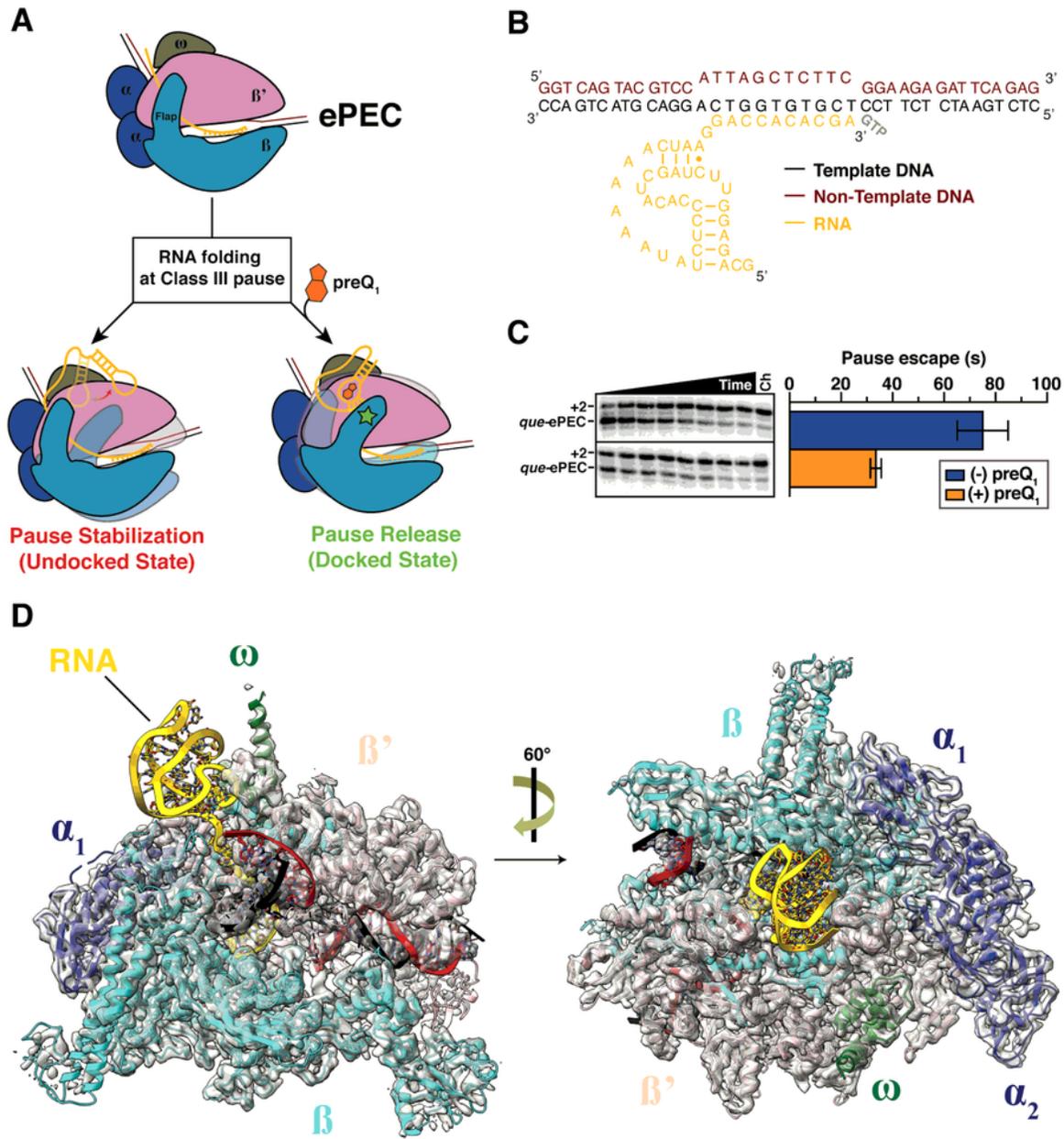


Figure 1

**Mechanism of transcriptional pausing at Class III pause and structure of the *que*-ePEC** (A) Schematic illustration of transcriptional pausing regulation through a riboswitch conformational change at a Class III pause. Upon encountering a consensus pause sequence, RNAP enters an offline elemental paused state (ePEC) that can be either stabilized (undocked state) or released through a riboswitch conformational change (docked state) upon preQ<sub>1</sub> ligand binding. Green star represents potential contact

between the RNA and RNAP  $\beta$ -flap domain **(B)** Nucleic acid scaffold used for cryo-EM data acquisition. Sequence and secondary structure of the *Bsu* preQ<sub>1</sub> riboswitch used in this study are shown. **(C)** Lifetimes of the *que* pause determined in the absence and presence of preQ<sub>1</sub>. *que*-ePEC was assembled under conditions similar to those of cryo-EM data collection and the rate of pause escape was determined after addition of the next templated rNTP (rGTP). Samples were taken 15, 30, 45, 60, 90, 120, 240 and 480 s after the addition of 5  $\mu$ M rGTP. **(D)** Overall fold and cryo-EM density of the *que*-ePEC obtained in the absence of preQ<sub>1</sub>. The 3.3 Å resolution cryo-EM map is rendered as a transparent surface and the refined model of the *que*-ePEC is colored as labeled. The RNAP backbone is represented as a ribbon diagram. The RNA transcript is colored gold, tDNA is colored black and ntDNA is colored dark red.

Figure 2

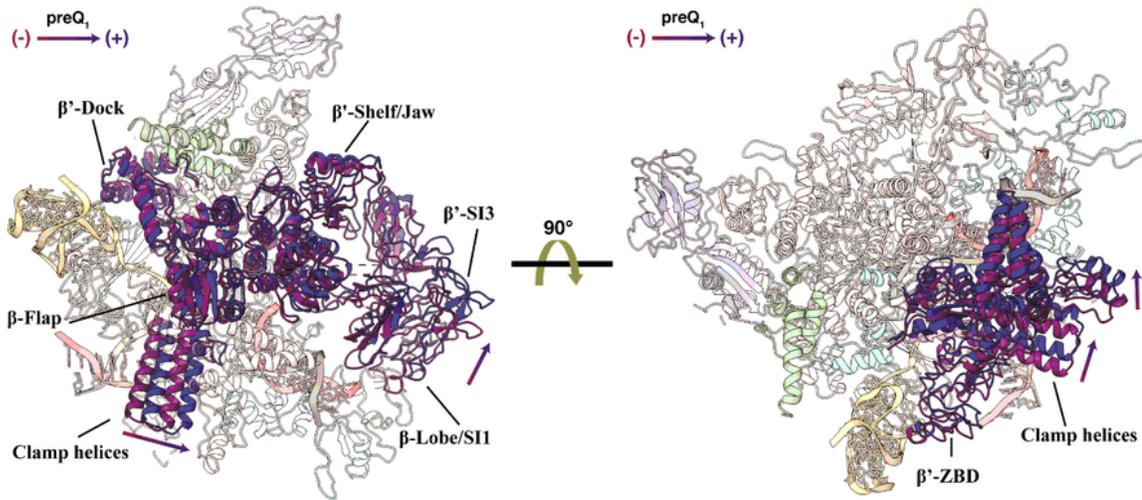
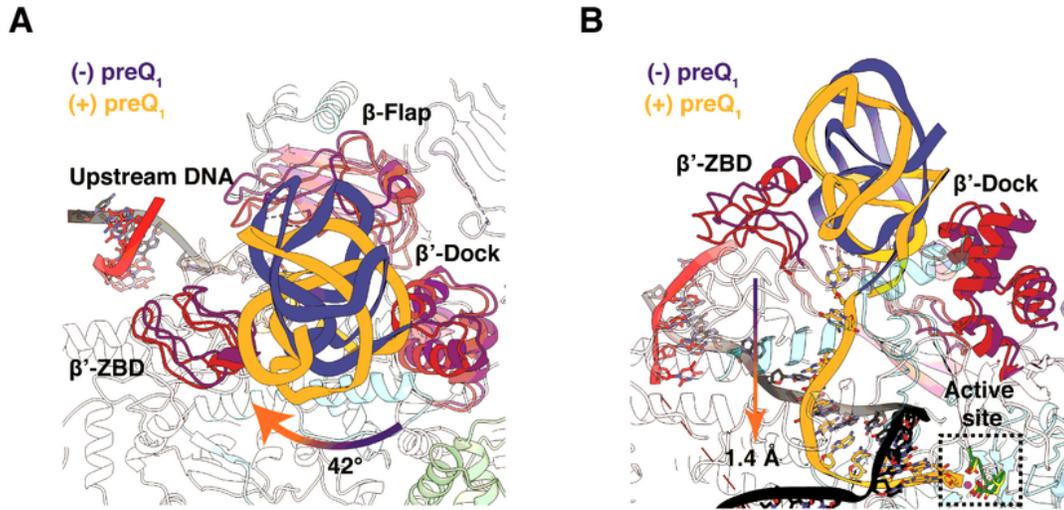


Figure 2

**Binding of preQ<sub>1</sub> to *que-ePEC* RNAP induces swivel module rotation and β-clamp closing. (A)** Least-squares super-positioning of the *que-ePEC* in the absence and presence of preQ<sub>1</sub>. The arrows at the β'-SI3 domain and clamp helices represent the directionality of the transition upon preQ<sub>1</sub> binding. The non-swiveled preQ<sub>1</sub>-bound *que-ePEC* with the closed-clamp domain is colored blue and the swiveled *que-ePEC*

structure in the absence of preQ<sub>1</sub> is colored purple. The ntDNA is colored red, the tDNA is colored black and the RNA nucleotides are colored orange. The protein backbone is rendered as cartoon ribbons and the nucleic acid is rendered as both sticks and cartoons. Coordinates were aligned with the RNAP structural core module in ChimeraX<sup>88</sup>.

**Figure 3**



**Figure 3**

**Rotation of the riboswitch upon ligand binding expands the RNA exit channel to accommodate the nascent transcript. (A)** RNAP exit channel subdomains in close proximity to the emerging riboswitch in the absence (blue) and presence (orange) of preQ<sub>1</sub> ligand. The arrow depicts the 42° rotation of the riboswitch upon ligand binding. Key RNAP structural subdomains are colored purple and red in the absence and presence of preQ<sub>1</sub>, respectively. **(B)** Rotated view of the RNAP exit channel showing the widening of the exit channel domains and forward translocation of the transcript toward the active site (dashed black box). The catalytic triad residues D460, D462 and D464 (green) and the magnesium ion (magenta) of the active site are indicated <sup>89</sup>. The arrow depicts the direction of the RNA translocation in the presence of preQ<sub>1</sub>.

Figure 4

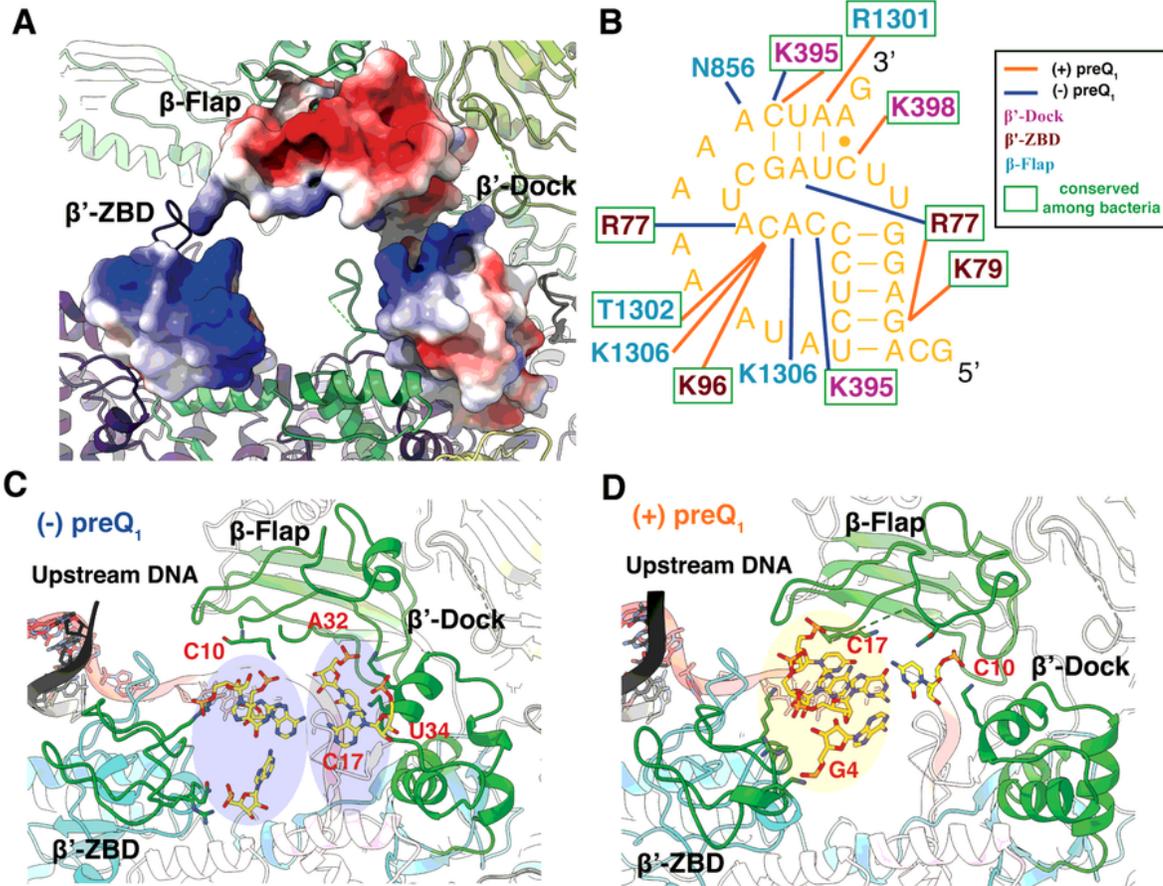
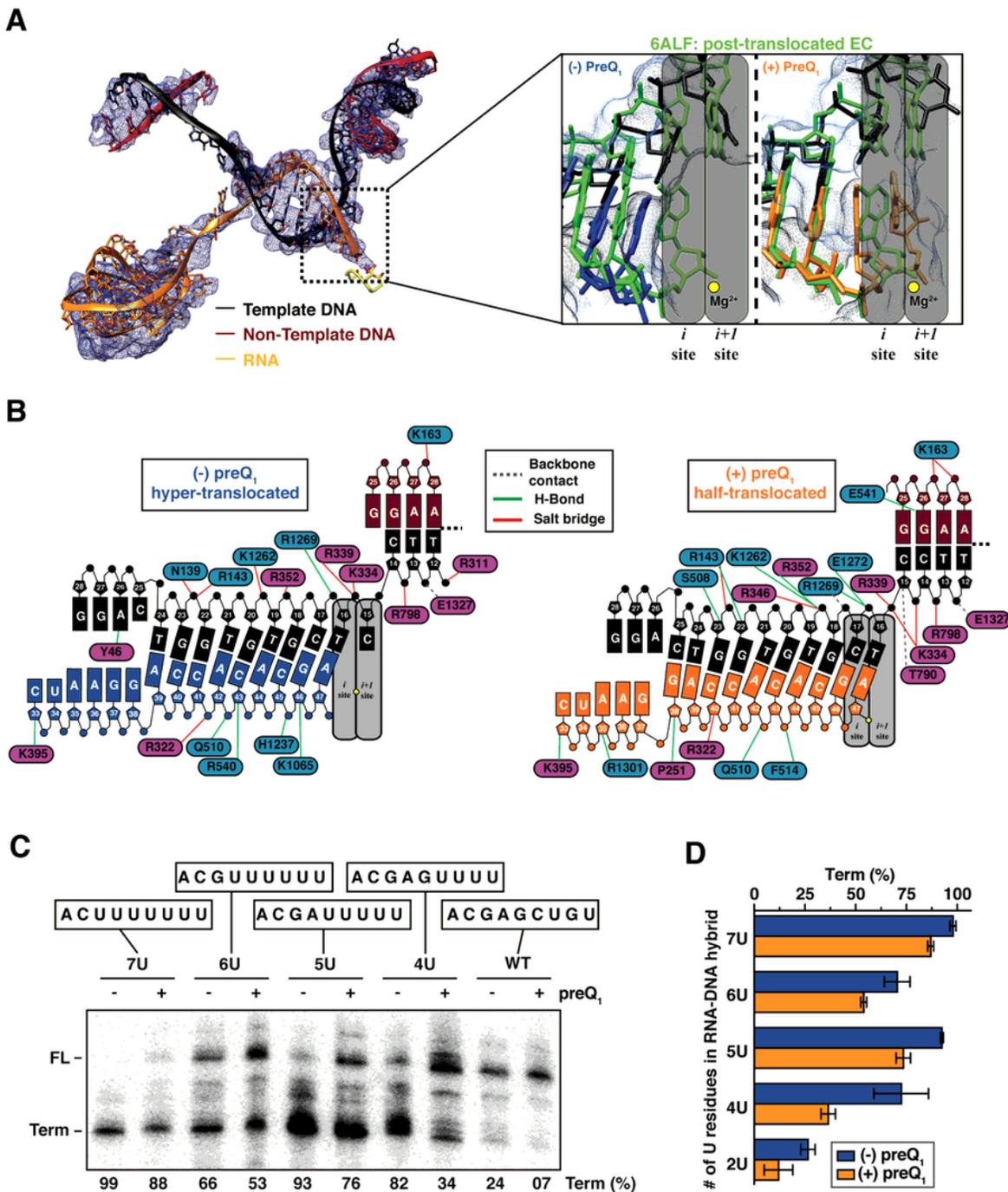


Figure 4

**Key RNA:protein interactions within the RNA exit channel. (A)** View from the outside of the RNAP exit channel with the surface of RNAP colored by electrostatic charge (-5 red to +5 blue). **(B)** Map of RNAP amino acid side chains that are in close contact with the riboswitch pseudoknot. Blue, interactions in the absence of preQ<sub>1</sub>; orange in the presence of preQ<sub>1</sub>. Green boxes highlight residues conserved among bacteria <sup>1</sup>. **(C)** Close contacts between The RNAP exit channel nucleotides and the 5'-side of the

pseudoknot in the absence of preQ<sub>1</sub>. The RNAP exit channel domains are colored in green and the RNA residues in yellow. Shaded ovals indicate the regions of proximity between the riboswitch and RNAP domains and key nucleotides are indicated in red (see also Table S4). **(D)** preQ<sub>1</sub> binding and subsequent steric hindrance shift the riboswitch closer to the β'-ZBD. Shaded oval indicates the region of proximity between the riboswitch and RNAP domains and key nucleotides are indicated in red (see also Table S4).

**Figure 5**



**Figure 5**

**Transcription reactivation occurs through RNA forward translocation in the presence of ligand. (A)** Nucleic acid scaffold and density map (blue mesh) for the RNA-DNA hybrid is indicated on the left. The RNA transcript is colored gold, tDNA is colored black and ntDNA is colored dark red. On the right the RNAP translocation state is compared to the cross-linked EC (PDB: 6ALF), which is colored green. The *que*-ePEC is colored in blue in the absence of preQ<sub>1</sub> and orange in the presence of ligand. **(B)** RNAP:DNA and RNAP:RNA contacts in the absence (left) and presence (right) of preQ<sub>1</sub> ligand. **(C)** *In vitro* transcription assay of mutants altering the uracil content in the RNA-DNA hybrid at the *que*-ePEC performed in the absence and presence of preQ<sub>1</sub>. Sequences of the RNA-DNA hybrid for each mutant are indicated on the top. **(D)** Percentage of terminated product relative to the full-length for each construct tested. Error bars are SD (stand deviation) of the mean from independent replicates.

Figure 6

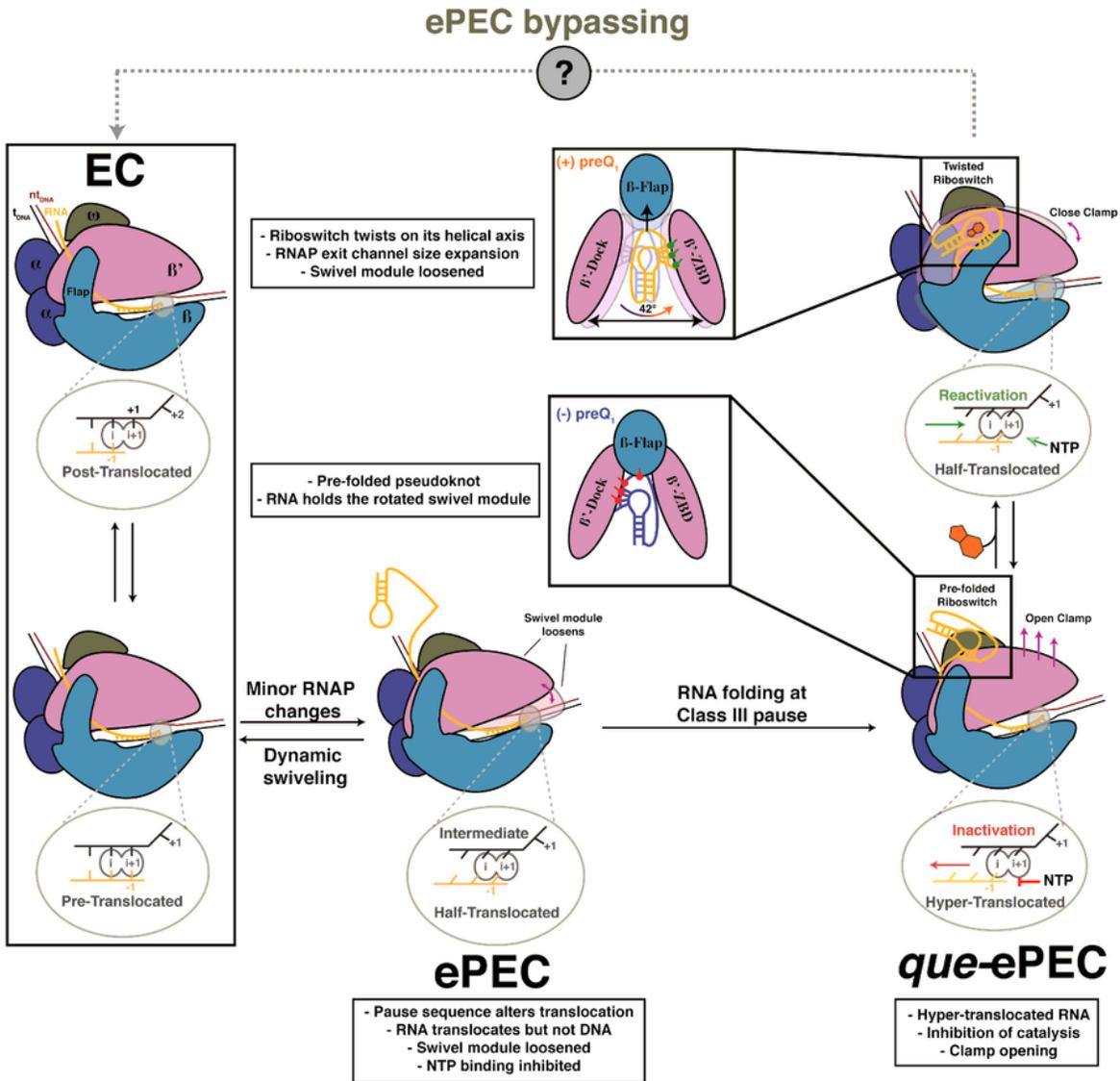


Figure 6

**Model for RNAP entering and release from the Class III pause as a function of ligand binding to the riboswitch.** RNAP can convert to an ePEC once encountering a consensus pause sequence that is stabilized by riboswitch folding. Binding of preQ<sub>1</sub> ligand induces pseudoknot stabilization to release RNAP from the paused state. The docked state (induced by ligand binding) leads to riboswitch rotation

within the RNAP exit channel, ultimately leading to RNA exit channel expansion to accommodate the nascent transcript. Active site schematics are shown as oval insets.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ChauvierPortaNSMBSupplementaryInformation.docx](#)
- [Movie1.mp4](#)
- [Movie2.mp4](#)
- [Movie3.mp4](#)
- [Movie4.mp4](#)
- [Movie5.mp4](#)