# Supplementary Materials for:

# Direct Observation of Cotranscriptional Folding in an Adenine Riboswitch

Kirsten L. Frieda[1], Steven M. Block[2,3]*

[1]Biophysics Program, Stanford University, Stanford, CA 94305, USA.
[2]Department of Applied Physics, Stanford University, Stanford, CA 94305, USA.
[3]Department of Biology, Stanford University, Stanford, CA 94305, USA.

*To whom correspondence should be addressed. E-mail: sblock@stanford.edu (S.M.B.)

**CONTENTS:**



## Materials and Methods

### Riboswitch DNA template and RNA transcript

The designs of the DNA template and corresponding RNA transcript were similar to those described previously (*14*); the DNA template used for these experiments is diagramed in Fig. S1. A 117 bp DNA fragment containing the *pbuE* adenine riboswitch was first cloned into plasmid pALB3 at a position 31 bp downstream of the T7A1 promoter. The DNA transcription template was amplified by PCR from this cloned plasmid as a 809 bp region containing the T7A1 promoter and riboswitch sequence. The template ended with a 5′-terminal biotin-labeled base located 36 bp from U8 of the riboswitch terminator U-tract. This biotin tag was used to create a transcriptional roadblock at the template terminus by incubating the template with excess streptavidin protein (ProZyme), followed by excess biotin, and then purifying the DNA.

This DNA template design placed a spacer of 36 nt between the transcriptional start site of the T7A1 promoter and any native riboswitch-flanking sequence, and 40 nt from the

transcription start site to the beginning of the riboswitch sequence (operationally defined as the 5′ end of P1, the first helix involved in aptamer formation). The spacer sequence between the start site and the riboswitch provided a short region where RNAP could be transcriptionally initiated and then elongationally stalled (by depriving it of one of the four nucleotides) before reaching the riboswitch sequence. Once elongation was re-started in assays, the completed transcript based on this template contained 102 nt from the start of P1 to U8 of the terminator U-tract. Assuming that RNAP has a ~30 bp footprint (*26, 27*), the separation between an RNAP molecule that ultimately reached the transcriptional roadblock and any upstream nucleotides involved in the terminator element was designed to be at least 10 nt, so that the folding of a fully-transcribed riboswitch would be unencumbered by the presence of RNAP. (Experimentally, the position where RNAP was found to arrest at the roadblock agreed closely with our design estimate, as discussed in the main text and reported in Table S2.) Working with short RNA transcripts lacking extended flanking regions proved useful because these provide an isolated view of riboswitch elements without interference from other, potentially non-native, cross-hybridization events that might have led to additional folds. Using short transcripts also improved instrumental resolution over previous work that applied controlled loads to RNA transcripts (*20, 21*) by decreasing the elastic compliance, which reduces Brownian noise in the measured extension.

**Trapping Assay Assembly**

Biotinylated RNAP molecules were initiated on the DNA template and then stalled by the omission of UTP at the downstream T residue, after transcribing 29 nt, as described previously (*28*). Transcriptionally stalled elongation complexes (ECs) were purified free of NTPs by a size-exclusion column, and could be restarted by subsequent addition of NTPs. Stalled ECs were attached to avidin-coated polystyrene beads (730 nm diam.) by a roughly stoichiometric

incubation (~0.3 nM ECs and beads) for ~1 hr at room temperature (RT). A 3,057 bp dsDNA handle with one biotin end-label and one sticky end, complementary to the nascent transcript (described previously (*20*) and below), was incubated at ~50-fold molar excess with avidin-coated polystyrene beads (600 nm diam.) for ~1 hr at RT. These beads were then washed free of unbound handle by centrifugation. To assemble completed dumbbells, EC-bound beads and handle-coated beads were combined, roughly stoichiometrically, and incubated for ~1 hr at RT (*21*), so that the 5′ cohesive end of the DNA handles could hybridize with the initial RNA transcripts to form tethers (Fig. 1B). The product of this final incubation was then diluted ~15-fold into transcription buffer (50 mM Hepes pH 8.0, 130 mM KCl, 4 mM $Mg(OAc)_2$, 0.1 mM EDTA, 1 mM DTT) containing an RNase-free oxygen-scavenging system consisting of glucose oxidase (Calbiochem) and glucose (Sigma). The mix was placed in a flow chamber on a microscope slide, and transcription was restarted by the addition of 1 mM NTPs in the presence or absence of adenine. Typically, a dumbbell was captured and maintained in the twin optical traps as a buffer containing NTPs was introduced into the flow channel by capillary action. This flow was carefully limited, in such a way that beads did not get ejected from the optical traps by viscous drag forces associated with the flow. This approach allowed the tether length to be monitored continuously as RNAP resumed transcription, with only a few seconds of a settling transient as the buffer was first introduced (and before transcription began).

**DNA Handle Considerations**

A sticky-ended dsDNA handle was used to connect the initial RNA transcript to one of the beads of the dumbbell. The cohesive end of this handle consisted of a 5′ overhang of ssDNA complementary to the 5′ end of the RNA transcript. Handles with differing lengths of overhangs were selected for hybridization to the initial transcript, in order to sequester different lengths of

nonnative, 5′ sequence. The three handle versions used were: xB handle, complementary to the first 36 nt of the RNA transcript; mxB handle, complementary to the first 36 nt of the RNA transcript, except for a mismatched bubble from nt 26-30; and E5′ handle, complementary to the first 25 nt of the RNA transcript.

The longer xB and mxB handles were specifically designed to block the first 36 nt of transcribed sequence from participating in any folding interactions, leaving just 4 native-sequence nt free before the synthesis of the riboswitch P1 stem. (The nucleotides of the mismatched bubble created with the mxB handle hybrid were either not accessible for base pairing or did not form energetically favorable pairings, or both.) As expected, no differences in riboswitch folding were observed between assays employing xB or mxB handles: these two handles were therefore functionally equivalent, and are referred to collectively as the (m)xB handle. The shorter E5′ handle blocked only the first 25 nt of the RNA transcript, leaving 15 nucleotides of mostly nonnative sequence upstream of the riboswitch P1 stem, potentially available for base pairing. Based on our measurements, some of those nucleotides could participate weakly in forming a large, looped structure (the "E5′LF" fold) which involves a portion of the aptamer sequence and forms under certain conditions at low force. (See the section on *Equilibrium Riboswitch Conformation*.) For forces above ~7.5 pN, where this extra fold is disfavored, the choice of handle was immaterial. However, for low force data, where the E5′LF folding event was potentially undesirable (e.g., for low-force estimates of *TE*), the (m)xB handle was used exclusively.

**Optical Trap**

The dual-beam optical trapping apparatus has been previously described in detail (*29*) and summarized (*14*). One trap was steered under computer control, while the position of the

bead in the other trap was measured. Force was typically applied using a feedback-based force clamp. Data were sampled at 2 kHz (although rates up to 20 kHz were used in some cases), then filtered at half the sampling rate with an 8-pole Bessel filter (Krohn-Hite). For presentation, data were median-smoothed in 15-50 ms windows. All measurements were taken at a sample temperature of 25°C, and the room temperature was controlled within ±0.2°C.

**Equilibrium Riboswitch Conformation**

Structures observed to fold in the riboswitch RNA are reported in Table S1. The assignments were verified by blocking the formation of the structures in question with short DNA oligomers complementary to their RNA sequence, as well as comparing the measured changes in extension with those expected from the structure predictions and comparing against previously reported values, where available (*14*). Folding transitions were analyzed as two-state systems, as described previously (*30*). In the completely folded riboswitch RNA at equilibrium, two structures were observed: P2, which unfolds ~10 pN, and the terminator hairpin, which unfolds at ~13 pN. These structures were always observed whether the tether was probed at constant force or in force-extension curves. Even with adenine present, the riboswitch aptamer structure was never observed to exist post-transcriptionally. The dominance of the terminator hairpin was anticipated given the overwhelmingly greater energetic stability of this hairpin (~32 kcal/mol) compared with the unbound aptamer (~12 kcal/mol; aptamer secondary structure only) (*15*) or with the bound aptamer (18 ± 2 kcal/mol) (*14*). Based on a Boltzman equilibrium, the terminator is expected to be $10^{10}$-fold more likely to form.

We found that an additional fold involving the riboswitch RNA was produced in templates that coded for a few additional non-native nucleotides located at the 5′ end of the riboswitch sequence. This version of the construct, referred to here as the "extra 5′ riboswitch"

version, was generated by the use of the shorter E5′ DNA handle (see *DNA Handle Considerations*), which, compared to the longer handles used in most of or experiments, blocks fewer nucleotides of the initial transcript after hybridization to the nascent RNA. Some of these unblocked nucleotides are able to pair with portions of the aptamer sequence to form an additional structure that opens at low force, designated the "E5′LF" fold. This fold involves the aptamer sequence around P2, and likely extends the P2 stem (with a possible rearrangement of the P2 nucleotides) by ~18-19 nt 5′ of P2 and ~7-10 nt 3′ of P2. The E5′LF fold was observed to occur at equilibrium when using the E5′ handle, in addition to the P2 and terminator hairpin folds, and it could not form when the aptamer was folded.

For aptamer-only RNA, the folding force for the collapse of the aptamer to its adenine-bound form is reported in Table S1. The collapse occurs from an ensemble of partially-folded states: these were operationally grouped into a single effective, not-fully-folded state, without trying to distinguish among them. This approach was helpful in supplying an estimate of the force where the adenine-bound aptamer tends to form. The partially-folded states represent conformations with P2 and P3 folded much of the time, though P3 unfolding is possible (at slightly higher force), and the nearly-folded, adenine-competent state reported previously (*14*) likely dominates (at slightly lower force). The transition distance is based on the observed mix of partially-folded states, and the associated number of nucleotides therefore depends somewhat on the model of what is folding. Here, the distance is correct, within error, for folding from the P2- and P3-folded state. Our measurements of the aptamer RNA properties were consistent with those reported previously (*14*).

**Additional Sample Transcription Traces**

Additional examples of transcription records are supplied in Figures S2-S6. The force exerted on the RNA affects the cotranscriptional folding transitions that are permissible, and structural elements tend to fold only if their unfolding forces are greater than the tension applied. Figure S2 displays a representative sample of records taken over the range of forces explored in this study.

**TEs and Intrinsic Terminator Fit**

Termination efficiency (*TE*) was defined as the fraction of elongating complexes dissociating at the terminator position: $TE = N_T/(N_T + N_R)$, where $N_T$ is the number of events that terminate and $N_R$ is the number that run through the termination signal. At zero and low force, the amount of modulation in switching that we observed in the absence compared with the presence of saturating adenine ($\Delta TE \sim 40\%$) was consistent with changes in *TE* reported previously in assays of riboswitches *in vitro* (*7, 18*). Somewhat greater switching modulation may be achieved *in vivo*, perhaps with the help of additional transcription factors like NusA and GreA (*7, 18, 22*).

A quantitative model that accurately describes the termination efficiency of an intrinsic terminator as a function of load was developed previously (*21*): this model accounts for both the stabilizing effect of the RNA:DNA hybrid located inside RNAP and the destabilizing effect of terminator hairpin formation. Increasing the applied load reduces the hybrid stability but simultaneously biases against terminator hairpin folding. At lower forces, it is the latter effect that initially dominates, resulting in a decreasing *TE* with load. Here, the load response of the *TE* for the riboswitch in the absence of adenine (Fig. 2B) is similar to that of an intrinsic terminator,

and it was well fit by the model. The model parameters are $E_{hybrid}$, $\delta_{hybrid}$, $E_{stem}$, $\delta_{stem}$, which represent the energies ($E$) and distance parameters ($\delta$) associated with either the RNA:DNA hybrid or the hairpin stem (*21*). $E_{hybrid}$ and $\delta_{hybrid}$ were held constant within the range previously determined for the *his* terminator, which also has a U-tract composed entirely of U. The fit values for $E_{stem}$ and $\delta_{stem}$ (6.4 ± 0.7 $kT$ and 3.4 ± 0.6 nm, respectively) are similar to those reported for other terminators studied (*21*). The terminator stem distance parameter, $\delta_{stem}$ , is slightly larger than that identified previously for terminators with GC closing base pairs (2.4 ± 0.2 nm), and may reflect the importance of closing additional bases in the hairpin stem base, because the riboswitch terminator has a weak final UA base pair and a greater fraction of UA base pairs overall.

The model was fit without zero-force data because it was previously shown that the unloaded *TE* can be depressed by structural interference from upstream sequences (*21*). However, we also performed two gel-based measurements of the unloaded *TE* without adenine (Fig. 2B, blue and gray open symbols). The model estimate agrees well with our bulk results for the isolated terminator hairpin, as anticipated from past work (*21*). However, the presence of the upstream aptamer sequence only weakly depressed the intrinsic *TE* of the terminator (~10%): in the absence of adenine binding, the aptamer sequence is an ineffective competitor of termination.

**Folding Model: Theoretical Transcription Trace Prediction**

We modeled folding riboswitch trajectories based on a highly simplified picture, but one that nevertheless managed to capture many of the essential elements. Our model assumes a constant rate of RNA synthesis, although instantaneous transcriptional rates are known to vary and can also include transcriptional pauses. We furthermore assume that any structural elements that can fold under a given tension will do so in their entirety, and promptly, once synthesized

and free of the RNAP footprint. This model also takes into account physical details affecting extension, including the force-dependent stretching of RNA and its helix diameter.

Parameters necessary for estimating the expected RNA extensions for the folding/unfolding of structural elements include the following: the RNA helix width (*H*), the extension associated with an A-form helix base, taken to be 2.2 nm (*31*); epsilon (ε), a small extension correction (based on the fact that the first few nucleotides transcribed are not single-stranded, but hybridized to the DNA handle) (*31*), and gamma (Γ), the load-dependent conversion factor between nanometers and nucleotides, determined from a modified Marko-Siggia (worm-like chain) model of RNA (*32*). The mechanical parameters for ssRNA were the same as used previously (*14*): 1.0 nm persistence length , 0.59 nm/nt contour length, and an elastic modulus of 1600 pN/nm. For the handle correction, ε, the constants for the RNA:DNA hybrid were assumed to be similar to those of dsRNA, with an A-form contour length of 0.28 nm/nt (*31*). To calculate extensional changes, all free nucleotides were assumed to contribute to the extension (according to Γ), whereas nucleotides sequestered inside any folded structures did not contribute, except by an amount associated with the width of a stem helix base.

The predictions of this simplified model are compared with data in Fig. 3B, for a record acquired at 5.8 pN. The model agrees reasonably well with the actual positions and length changes observed for various cotranscriptional events. By construction, the model depicts folding events as being much more abrupt than they are observed to be experimentally. In reality, the sharpness of a folding event is set by several factors. First, transcripts continue to elongate even as folding takes place. Second, hairpins begin folding as soon as energetically favorable, even if not all the stem nucleotides are yet available (this is probably more significant for long

hairpins than for short ones like the P2 and P3 helices). Third, noise is present along with a measurement response time set by the time window for any signal averaging.

**Folding Model: Cotranscriptional Folding Event Characterization**

We used the extension change ($\Delta x$) and time interval ($\Delta t$) associated with different events to pinpoint the nucleotide positions where events occurred and their sizes. As a matter of definition, the nucleotide position of an event identifies the last nucleotide transcribed at the polymerase active site when the event occurred, and it is measured relative to the restart position of transcription. The parameters, $H$, $\Gamma$, and $\varepsilon$, as well as an estimate of transcription rate (approximated by averaging two separate estimates of the slope in non-folding regions), were used to back-calculate event positions. For each single-molecule record, event positions were estimated based on ($\Delta x$, $\Delta t$) relative to the transcription start by averaging two different determinations: the first was based on the event time, $\Delta t$, and the second on the event position, $\Delta x$. Event fold sizes were calculated based on ($\Delta x$, $\Delta t$) while extension decreased for the fold.

Cotranscriptional events were characterized at various forces (see *Additional Sample Transcription Traces*); data are summarized in Table S2. We note that the apparent folding size of the terminator hairpin measured in run-through records corresponded closely to its full length, including base pairs involving the first Us of the U-tract. By contrast, the folding size of this element in records that terminated was only sufficient to account for the topmost ~14 bp of the terminator hairpin stem. This result may be explained because the remaining ~13 nt of RNA are likely sequestered within the RNAP footprint prior to complete terminator hairpin formation and coincident termination near U7.

The measured aptamer folding size was somewhat variable, exhibiting a range of apparent folding distances (~5-10 nm). The likely explanation for this is that aptamer folding

occurs from an ensemble of initial states, e.g., from a conformation with P2 and P3 folded separately, or from a conformation with P2-P3 forming a loop-loop interaction in the adenine-competent state.

The chief sources of error in these single-molecule measurements are associated with inaccuracies in the force and position, as discussed in (*14*). A primary contribution to such inaccuracies stems from small variations in bead size that affect position determinations (directly) and force estimates (indirectly, via the force-displacement relationship, and directly via the trap stiffness). Measuring transcription events on many separate bead-based dumbbells helps to reduce this error. Additional uncertainties arise from systematic errors, for example those associated with trap stiffness calibration. Overall, we estimated the systematic error in force at ±7%, and errors in the calculated nucleotide positions and fold sizes at ±2%.

**Kinetic Switch**

The time window for the binding of adenine to the riboswitch is expected to extend, maximally, from the time that the aptamer sequence becomes available to fold until the moment that the terminator hairpin potentially forms. Re-expressed in terms of nucleotides, this amounts to ~27 nt, because the distance from the 3′-end of aptamer P1 to U8 of the terminator U-tract is 39 nt, and an additional ~12 nt of transcription are necessary beyond the completion of P1 RNA to free the aptamer sequence from the RNAP footprint so that it can fold (*23, 24*). Assuming a steady transcription rate of ~18 nt/s (similar to that observed at 5.8 pN load), the duration of this time window becomes 1.5 s, similar to a previous estimate (*16*). This time is consistent with our observations, and short compared with estimates of the adenine-bound aptamer lifetime: its lifetime averages ~5-25 s, based on various estimates of the aptamer adenine off-rate and

adenine-bound lifetime at low force (*14, 16*), though it is possible that the off-rate increases as transcription continues beyond the aptamer.

In this study, saturating adenine concentrations (300 μM) were used, since the goal was to ensure switching behavior. Because the switch is found to be kinetically, and not thermodynamically, controlled by folding of the aptamer, the observable riboswitch behavior is not expected to be very sensitive to adenine concentration. A drop in the adenine level of nearly 100-fold, which takes it below the intracellular level of ~30 μM (*33*), should still yield measureable binding within a roughly 2-s binding window, given the high on-rate for adenine (e.g., reported previously as $2.6 \times 10^5$ $M^{-1}$ $s^{-1}$ (*16*) or $8 \pm 1 \times 10^4$ $M^{-1}$ $s^{-1}$ (*14*)). In this context, we note that gel-based experiments with this riboswitch supported a kinetic mechanism for its action and a low, μM-range concentration regime for modulating *TE* under the experimental conditions (*18*).

As discussed (main text), we did not observe significant pausing at the U-stretches before the terminator-associated U-tract of the *pbuE* riboswitch. Various factors could affect pause lifetimes. Force, particularly for the loads applied here, likely exerts a negligible influence on pausing, because (a) the forces were low (below ~13 pN) and (b) applied in a direction that assists transcription, and (c) force mainly affects pauses that involve backtracking, and not ubiquitous pauses (*28, 34*). However, decreasing NTP concentrations below physiological levels (mM NTPs) is known to make certain pauses more prominent, as well as lower overall transcription rates. Working at an NTP concentration of 20 μM, a recent study of the *pbuE* riboswitch reported a transcriptional pause near the end of the 6U series, before the terminator U-tract. Similarly, a study of the FMN riboswitch reported long pauses in assays conducted with 200 μM NTPs, but pause lifetimes were reduced at a higher NTP concentration of 500 μM, as

expected (*7*). In principle, the use of the cognate RNAP from *B. subtilis* (rather than from *E. coli*) to transcribe the *pbuE* riboswitch might affect pausing. However, this issue was examined in a previous study that found little difference with the choice of RNAP, with both yielding similar *TE*s and pausing (*18*). If anything, *B. subtilis* tends to display less pausing behavior than *E. coli* (i.e., shorter lifetimes and less efficient recognition), not more (*7, 22*). It is unclear if transcriptional pausing plays any role in the switching behavior of riboswitches *in vivo*, at physiological NTP levels.

**Folded Aptamer Dwell Times**

The adenine-bound aptamer can remain stable for seconds in a nascent transcript. It is intriguing to consider how much of this time the aptamer remains folded despite the presence of the terminator hairpin, which is energetically favored (this is a measure of the kinetic competitiveness of the aptamer). Operationally, we took transcription of U8 in the U-tract as the reference position where the termination decision was made and the terminator hairpin was formed during transcription. (Appropriately, the few records that initially displayed a signature for aptamer folding, but subsequently terminated, all showed the aptamer later unfolding before reaching the U8 reference point.) This position was estimated in each record in two ways: one based on the transcription rate, and the other based on the expected extension, as described earlier. The time that the aptamer remained folded following U8 transcription could then be determined. Based on our data, this dwell time averaged 2.6 s (Fig. S7), which represents the time that the bound aptamer structure can kinetically trap the RNA out of equilibrium before the formation of the more stable terminator hairpin.

Only a very small extension change was anticipated between the adenine-bound aptamer conformation and the final, equilibrium conformation with both the terminator hairpin and P2

folded. To clarify this transition, the persistence of the adenine-bound aptamer was studied using the "extra 5′ riboswitch" version created by using the shorter E5′ handle (see *Equilibrium Riboswitch Conformation*). This handle variant leads to the formation of the E5′LF fold, which, in effect, amplifies any extension change in the final configuration, due to the possibility of E5′LF formation. The E5′LF fold cannot form when the aptamer is folded (because it pairs with some of the aptamer sequence), but does so once the terminator hairpin structure dominates. This version of the riboswitch assay therefore makes it especially evident when the aptamer unfolds to give way to the terminator structure (Fig. S6).

The time that the bound aptamer persists in the presence of the terminator hairpin sequence is roughly exponentially distributed (Fig. S7), which is appropriate for an (essentially) one-step unfolding process. The lifetime of the adenine-bound fold is expected to be similarly distributed in RNA coding for the aptamer sequence alone. However, the bound lifetime has been reported to be longer for the aptamer alone (*14, 16*) than the lifetime in the context of the complete riboswitch observed here. Comparing the folded lifetime distribution of the aptamer alone ($k_{off}$ ~0.09 ± 0.01 $s^{-1}$) with that of the aptamer in competition with the terminator hairpin ($k_{off}$ ~0.34 ± 0.04 $s^{-1}$) illustrates this point (Fig. S7B). The difference in unfolding rates is consistent with our observation that many more aptamer folding events unfold in the first few seconds following U8 transcription than in a comparable time interval before terminator transcription. Before the terminator is present, the unfolding rate is mainly governed by the lower off-rate of adenine, but after terminator transcription, it is significantly higher, and likely enhanced by structural competition with the terminator. In summary, the adenine-bound aptamer successfully traps the riboswitch transcript out of equilibrium, but as a consequence it persists for less time than it would without competition.

**Bulk Termination Assays**

Transcriptionally-stalled RNAP complexes, pre-initiated on the riboswitch DNA template, and similar to those used in the single-molecule assays, were also used for bulk termination assays. Initial transcripts were labeled with α-$^{32}$P-GTP. Elongation proceeded under the identical assay conditions as those used in the single-molecule assays, including the addition of 1 mM NTPs, the same transcription buffer, the exclusion or addition of adenine (up to 300 μM), and the same assay temperature (25°C). For comparison with single-molecule results, a DNA oligomer complementary to the same RNA transcript region blocked by the DNA handles used in the dumbbell assay was added prior to NTP addition. Similarly, to estimate the *TE* of an isolated terminator hairpin by blocking the upstream RNA, a 44-nt oligomer complementary to most of the aptamer (until 4 nt before the sequence that participates in the terminator hairpin) was added before re-starting transcription. Blocking oligos were added at large molar excess (~10-100 μM). Termination and run-through products were separated by native or denaturing 15% PAGE, visualized with a PhosphorImager, and quantified using image analysis routines (Igor Pro; ImageQuant).

# Supplementary Figures and Tables

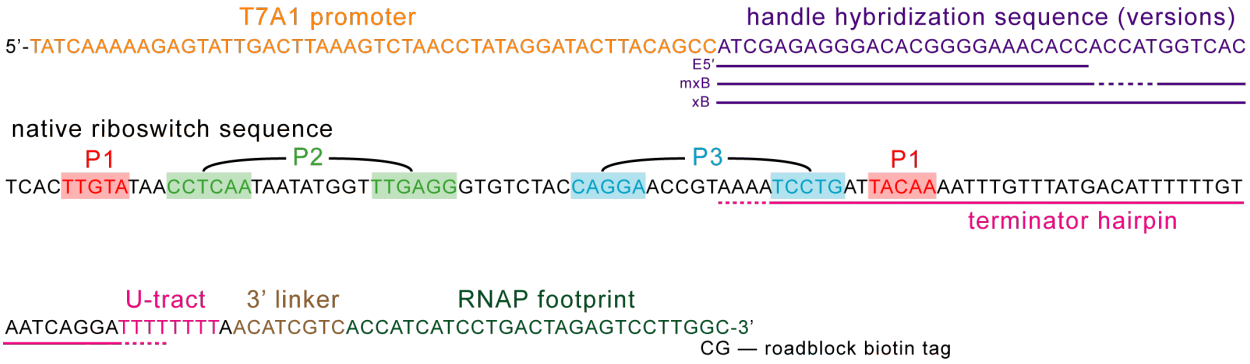


**Fig. S1.** DNA template used for RNA transcription, showing the sequence of the nontranscribed promoter followed by the transcribed region, with relevant sequence elements marked. (The upstream, nontranscribed portion of the DNA template is not shown.) The template first codes for a transcript that hybridizes to the DNA handle (used for trapping assay construction), and then codes for the complete *pbuE* riboswitch (including both aptamer and terminator hairpin sequences) followed by a short linker before the final position of the footprint of RNAP when stalled by the roadblock. The extents of three different DNA handles used are shown; the dashed line indicates a mismatched region.

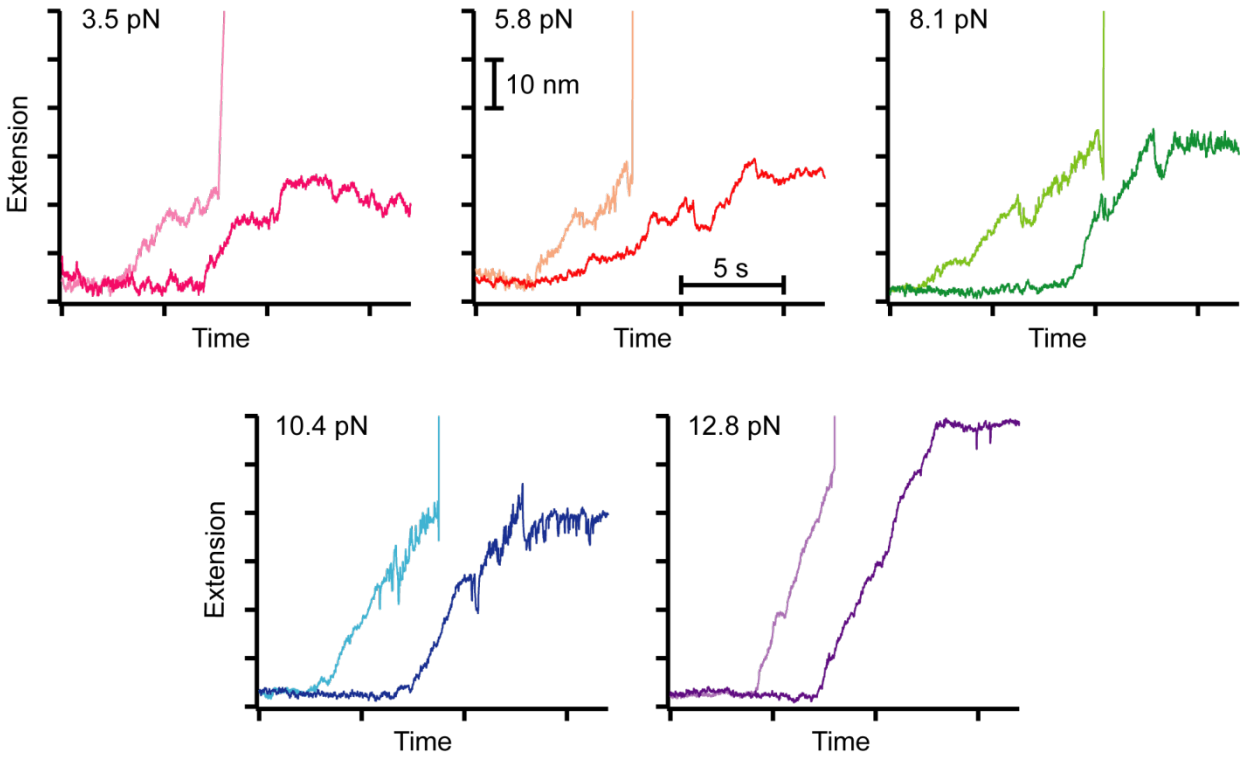


**Fig. S2.** Representative samples of cotranscriptional records collected under a variety of forces. An example of a terminating and a run-through trace are shown for each applied load; a sharp upward deflection in the extension signal line distinguishes the termination event, corresponding to tether breakage. Data were collected under saturating adenine conditions, although the presence of adenine did not affect the appearance of records at higher forces, above ~7 pN.

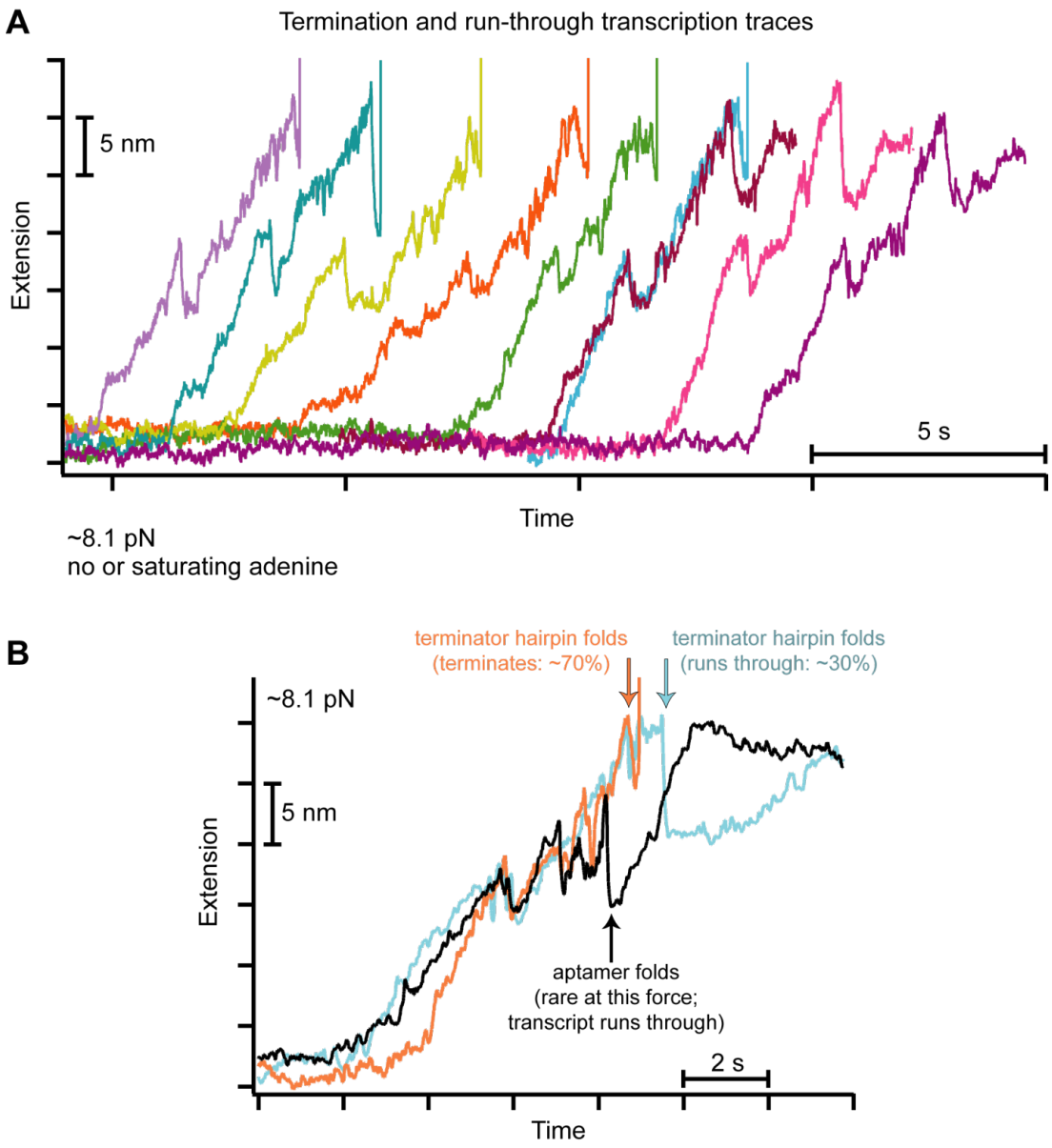


**Fig. S3.** Representative records obtained at ~8.1 pN load in the presence or absence of adenine. (**A**) At high loads, traces appeared similar regardless of the transcription outcome or the presence of a saturating level of adenine. P2 formation constitutes the first major folding event, followed later by terminator hairpin formation and immediate termination, or (more rarely) run-through to the roadblock. Termination events are distinguished by a sharp upward deflection in the extension signal at the end of a trace, corresponding to tether breakage. Run-through records displayed here were cut off shortly after transcription was arrested at the roadblock. Traces are offset horizontally for clarity. One pair of traces displaying different transcriptional outcomes (maroon, blue) is superimposed for comparison purposes. (**B**) A single record (black) of adenine-bound aptamer formation was obtained at this relatively high force, leading to run-through: it is superimposed against two typical traces collected at this force that terminated (~70% probability overall; orange) or ran through (~30% probability overall; blue).

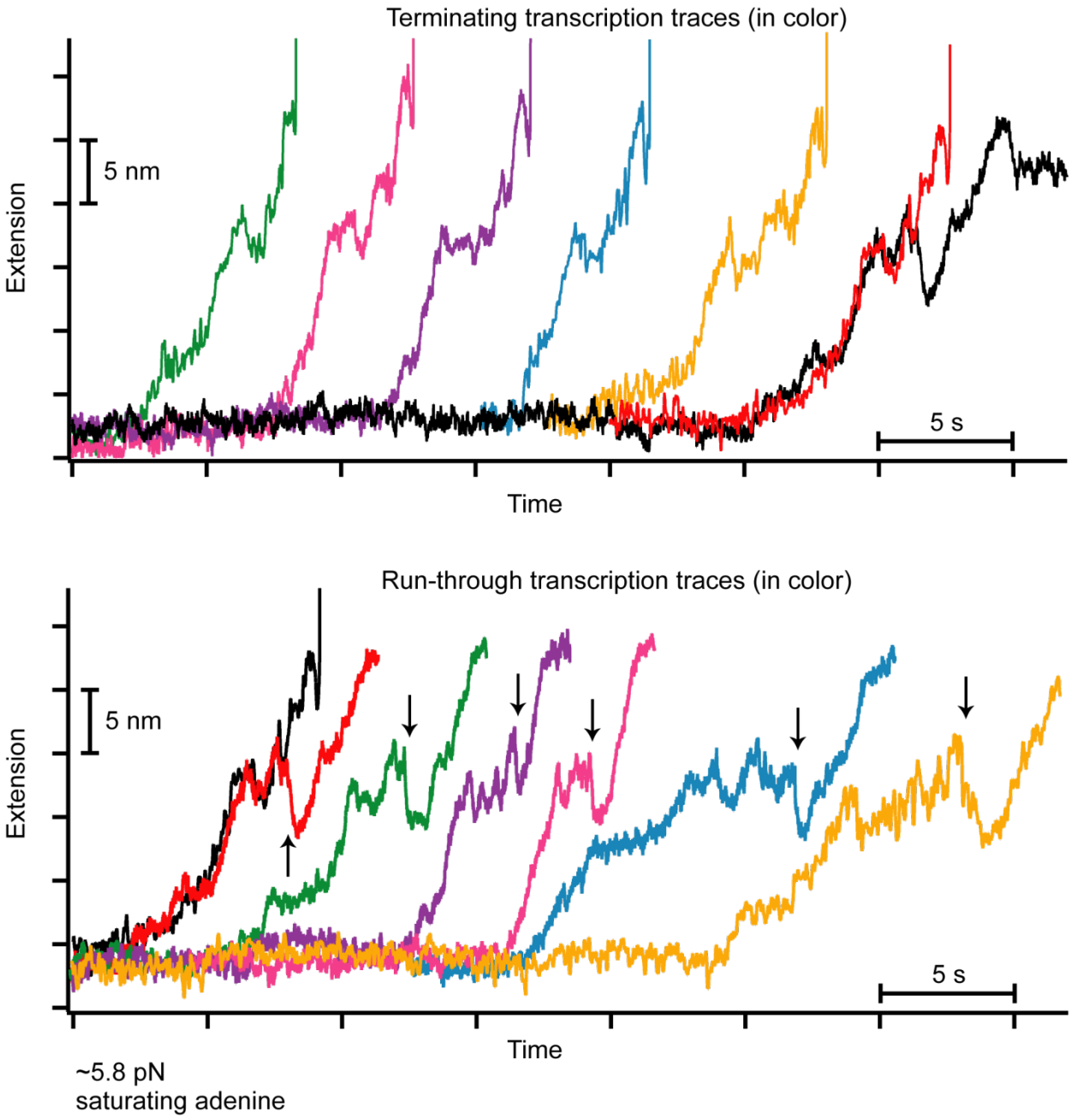


**Fig. S4.** Representative records obtained at ~5.8 pN load and saturating adenine conditions, where termination (top) and run-through (bottom) traces are generally distinct. P2 formation constitutes the first major folding event, occurring about midway through the records, and was a common feature in all traces. Termination events are distinguished by a sharp upward deflection in the extension signal at the end of a trace, corresponding to tether breakage. Run-through records displayed here (bottom) were cut off after transcription was arrested at the roadblock. Aptamer folding event in records that carry these are indicated (black arrows); once folded, the adenine-bound aptamers were stable for the remainder of the record. Traces are offset horizontally for clarity. One pair of traces displaying different transcriptional outcomes is superimposed in each panel for comparison purposes (black and red; see also Fig. 3C).

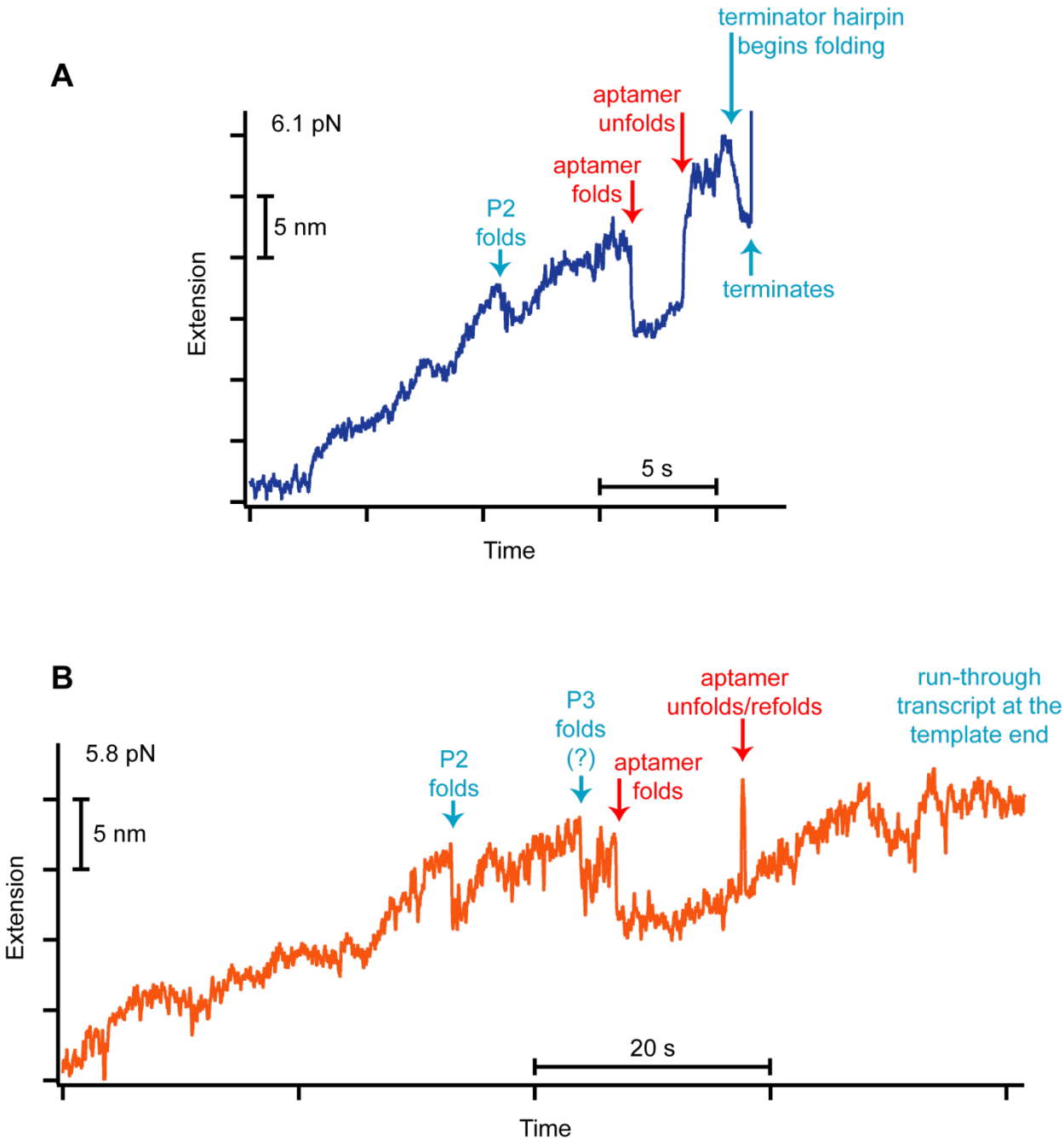


**Fig. S5.** Unusual (rare) examples of cotranscriptional aptamer folding behavior, obtained at ~5.8 pN and saturating adenine. (**A**) In this relatively slow transcriptional record, the adenine-bound aptamer formed only transiently, unfolding before the position of the terminator was reached, leading to hairpin folding and termination. (**B**) In this very slow transcriptional record, the adenine-bound aptamer was observed to fold twice: once following aptamer transcription, and again following a brief unfolding event (red arrows), resulting in transcriptional run-through.

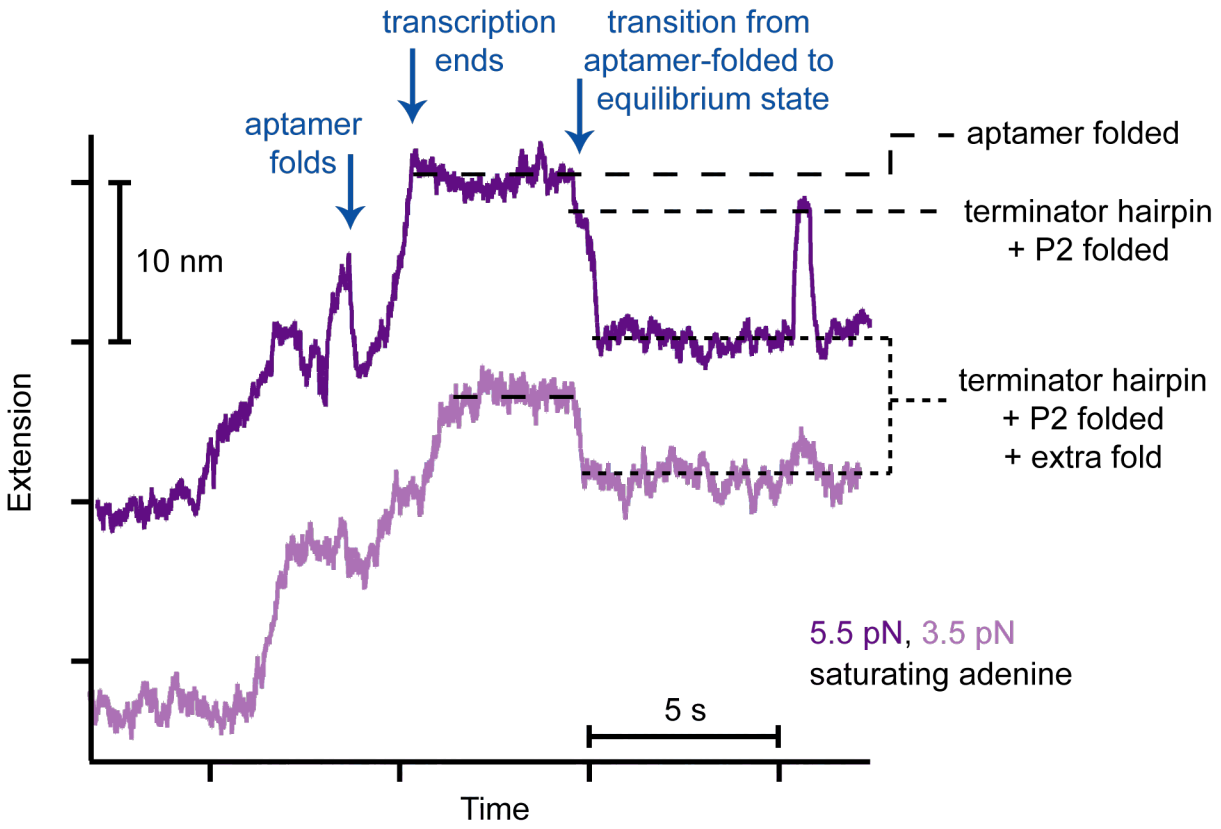


**Fig. S6.** The aptamer kinetically traps RNA structure and delays the transition to the equilibrium conformation with the terminator hairpin folded. The adenine-bound aptamer was observed to persist for several seconds following complete terminator transcription. Ultimately, the structure rearranged and the terminator hairpin conformation formed permanently. The transition between competing structures was particularly evident in a subset of records acquired with the shorter E5′ handle, which permits an additional, low-force fold, E5′LF, at equilibrium when the terminator hairpin is present (see SOM text). Once the aptamer has unfolded, this extra feature folds and unfolds at 5.5 pN, and is always folded at 3.5 pN.

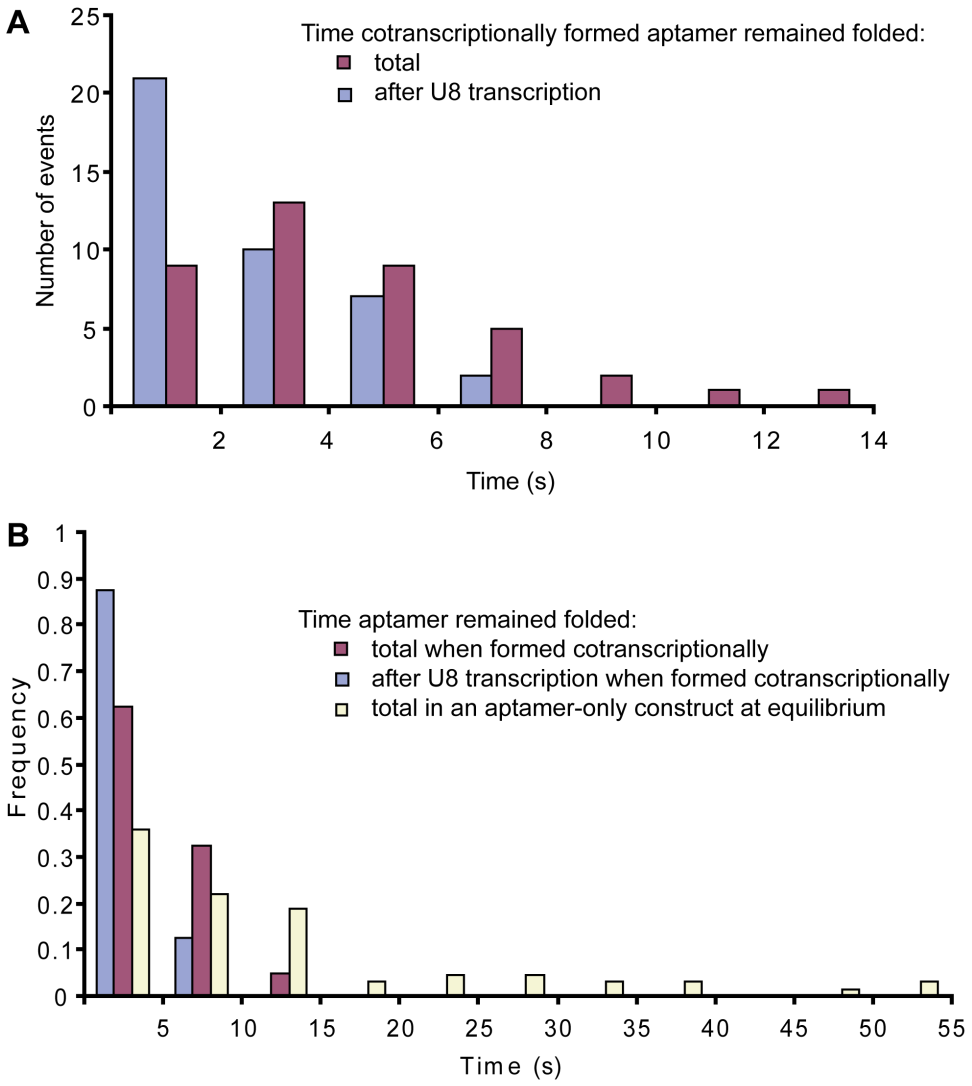


**Fig. S7.** Lifetime distributions for the folded aptamer. Data were acquired under ~5.8 pN load with adenine present. (**A**) Time that the adenine-bound aptamer persisted in a folded state after forming cotranscriptionally in run-through traces. Time was measured either as the total time spent in the folded state (maroon) or the time spent in the folded state after transcription of U8 in the terminator U-tract (blue), which corresponds to the lifetime of aptamer stability beyond the transcription of the competing terminator. (**B**) The aptamer lifetime measured cotranscriptionally (maroon, blue; data replotted from panel (A)) was markedly less than the aptamer lifetime measured at the same force and adenine concentration, but in the absence of any competing RNA structures (yellow).

**Table S1.** Extension changes and associated unfolding forces for post-transcription, equilibrium folding events in the full riboswitch or aptamer-only RNA. The $F_{1/2}$ value is the force under which equal time is spent in folded and unfolded states. Distances (in nm), measured near $F_{1/2}$, were converted to nucleotides (nt) based upon the load-dependent extension of RNA. Measurements for each transition are reported for at least 6 molecules in the presence of adenine. Uncertainties (std. dev.) are statistical errors and do not include systematic errors.

| **Folding event** | **Δx (nm)** | **Δx (nt)** | **$F_{1/2}$ (pN)** |
|---|---|---|---|
| Terminator hairpin folding (full riboswitch) | 18.3 ± 0.5 | 49 ± 1 | 13.1 ± 0.5 |
| P2 folding (full riboswitch) | 5.9 ± 0.2 | 21 ± 1 | 10.1 ± 0.3 |
| E5′LF folding ^ (full riboswitch) | 8.3 ± 0.5 | 25 ± 1 | 6.7 ± 0.5 |
| Aptamer folding * (aptamer only) | 9.1 ± 1.2 | 20 ± 3 | 7.4 ± 0.7 |

^ The E5′LF folding event was only observed in a version of the riboswitch construct with extra 5′ nucleotides, since some of these nucleotides participate in this fold (see SOM text).

* Aptamer folding describes the collapse of the RNA from an undifferentiated mix of partially-folded states to a completely-folded state. This information is useful for estimating the force where the aptamer can readily fold on its own.

**Table S2.** Positions and sizes of cotranscriptional folding events. Values are given in nucleotides (nt), with positions measured relative to the first template nucleotide transcribed (a 5′-U) upon restarting RNAP. Absolute nucleotide positions, based on the transcript sequence and secondary structure predictions, are given for reference in boldface. Results (mean ± SEM) are indicated for the different loads tested; overall results represent weighted averages of measurements. (Data at 3.5 pN, the lowest force collected, were too noisy to parameterize and were not included here.) The statistical error (±SEM; shown) was ~±1 nt, with an additional systematic error (not included) estimated around ±2%. T and R specify termination and run-through traces, respectively (see Fig. 4).

| **Event** | **Event position or size observed at the specified force (reported as a number of nucleotides (nt))** | | | | |
|---|---|---|---|---|---|
| | **5.8 pN** | **8.1 pN** | **10.4 pN** | **12.8 pN** | **Overall** |
| **First nt transcribed (U)** | | | | | **1** |
| **Last nt of the P2 stem** | | | | | **40** |
| # of nt transcribed when P2 folds: | | | | | |
| | T&R: 51 ± 1 | T&R: 53 ± 1 | — | — | 52 ± 1 |
| # of nt folding in P2 (observed): | | | | | |
| | T&R: 21 ± 1 | T&R: 22 ± 1 | — | — | 21 ± 1 |
| **# of nt in P2 (structure prediction)** | | | | | **21** |
| **Last nt of the P1 aptamer stem (all sequence of the aptamer has been transcribed)** | | | | | **74** |
| **Middle nt of the terminator hairpin loop (only the first half of the terminator transcribed)** | | | | | **84** |
| Average # of nt transcribed when the adenine-bound aptamer folds: | | | | | |
| | T&R: 90 ± 1 (few T traces) | — | — | — | 90 ± 1 |
| # of nt transcribed when the terminator hairpin begins folding: | | | | | |
| | T&R: 106 ± 1 | T&R: 106 ± 1 | T: 110 ± 1<br>R: 117 ± 1 | may form after terminator transcription | variable |
| **Position of terminator 8U tract** | | | | | **106-113** |
| # of nt transcribed when termination occurs: | | | | | |
| | T: 112 ± 1 | T: 109 ± 1 | T: 111 ± 1 | T: 113 ± 1 | 112 ± 1 |
| # nt folding in the terminator hairpin (observed): | | | | | |
| | T: 30 ± 1<br>R: 51 ± 3 | T: 29 ± 1<br>R: 44 ± 1 | T: 30 ± 2<br>R: 49 ± 1 | —<br>R: 50 ± 1 | T: 30 ± 1<br>R: 49 ± 1 |
| **# of nt in the terminator hairpin (structure prediction)** | | | | | **51** |
| # of nt transcribed when the run-through RNAP is stalled at the roadblock: | | | | | |
| | R: 131 ± 1 | R: 134 ± 2 | R: 135 ± 1 | R: 136 ± 1 | 134 ± 1 |
| **Last possible nt (but template DNA nt is attached to the roadblock)** | | | | | **149** |

**References and Notes**


1. S. L. Heilman-Miller, S. A. Woodson, *RNA* **9**, 722 (2003).
2. T. Pan, X. Fang, T. Sosnick, *Journal of Molecular Biology* **286**, 721 (1999).
3. T. Pan, I. Artsimovitch, X. W. Fang, R. Landick, T. R. Sosnick, *Proc Natl Acad Sci USA* **96**, 9545 (1999).
4. T. N. Wong, T. R. Sosnick, T. Pan, *Proc Natl Acad Sci USA* **104**, 17995 (2007).
5. A. Xayaphoummine, V. Viasnoff, S. Harlepp, H. Isambert, *Nucleic Acids Research* **35**, 614 (2007).
6. E. M. Mahen, P. Y. Watson, J. W. Cottrell, M. J. Fedor, *PLoS Biology* **8**, e1000307 (2010).
7. J. K. Wickiser, W. C. Winkler, R. R. Breaker, D. M. Crothers, *Molecular Cell* **18**, 49 (2005).
8. G. Nechooshtan, M. Elgrably-Weiss, A. Sheaffer, E. Westhof, S. Altuvia, *Genes Dev* **23**, 2650 (Nov 15, 2009).
9. G. A. Perdrizet, 2nd, I. Artsimovitch, R. Furman, T. R. Sosnick, T. Pan, *Proc Natl Acad Sci USA* **109**, 3323 (Feb 28, 2012).
10. T. Pan, T. Sosnick, *Annu. Rev. Biophys. Biomol. Struct.* **35**, 161 (2006).
11. T. M. Henkin, *Genes Dev* **22**, 3383 (2008).
12. A. Roth, R. R. Breaker, *Annu Rev Biochem* **78**, 305 (2009).
13. M. Mandal, R. R. Breaker, *Nature Structural & Molecular Biology* **11**, 29 (2004).
14. W. J. Greenleaf, K. L. Frieda, D. A. N. Foster, M. T. Woodside, S. M. Block, *Science* **319**, 630 (2008).
15. J. F. Lemay, J. C. Penedo, R. Tremblay, D. M. J. Lilley, D. A. Lafontaine, *Chemistry & Biology* **13**, 857 (2006).
16. J. K. Wickiser, M. T. Cheah, R. R. Breaker, D. M. Crothers, *Biochemistry* **44**, 13404 (2005).
17. R. Rieder, K. Lang, D. Graber, R. Micura, *ChemBioChem* **8**, 896 (2007).
18. J. F. Lemay *et al.*, *PLoS Genetics* **7**, e1001278 (2011).
19. Materials and methods are available as supplementary materials on *Science* Online.
20. R. V. Dalal *et al.*, *Molecular Cell* **23**, 231 (2006).
21. M. H. Larson, W. J. Greenleaf, R. Landick, S. M. Block, *Cell* **132**, 971 (2008).
22. I. Artsimovitch, V. Svetlov, L. Anthony, R. R. Burgess, R. Landick, *Journal of Bacteriology* **182**, 6027 (2000).
23. J. A. Monforte, J. D. Kahn, J. E. Hearst, *Biochemistry* **29**, 7882 (1990).
24. N. Komissarova, M. Kashlev, *Proc Natl Acad Sci USA* **95**, 14699 (1998).
25. Y. d'Aubenton Carafa, E. Brody, C. Thermes, *Journal of Molecular Biology* **216**, 835 (1990).
26. S. A. Darst, *Current Opinion in Structural Biology* **11**, 155 (2001).
27. S. J. Greive, P. H. von Hippel, *Nature Reviews Molecular Cell Biology* **6**, 221 (2005).
28. K. C. Neuman, E. A. Abbondanzieri, R. Landick, J. Gelles, S. M. Block, *Cell* **115**, 437 (2003).
29. E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, S. M. Block, *Nature* **438**, 460 (2005).
30. M. T. Woodside *et al.*, *Proc Natl Acad Sci USA* **103**, 6190 (2006).
31. W. Saenger, *Principles of nucleic acid structure*. (Springer-Verlag New York, 1984).

32. M. D. Wang, H. Yin, R. Landick, J. Gelles, S. M. Block, *Biophysical Journal* **72**, 1335 (1997).
33. P. Nygaard, H. H. Saxild, *Journal of Bacteriology* **187**, 791 (2005).
34. K. M. Herbert *et al.*, *Journal of Molecular Biology* **399**, 17 (2010).