Cotranscriptional folding of a riboswitch at nucleotide resolution

Kyle E Watters[1,6,7], Eric J Strobel[1,6,7], Angela M Yu[1–3], John T Lis[4] & Julius B Lucks[1,5]

RNAs can begin to fold immediately as they emerge from RNA polymerase. During cotranscriptional folding, interactions between nascent RNAs and ligands are able to direct the formation of alternative RNA structures, a feature exploited by noncoding RNAs called riboswitches to make gene-regulatory decisions. Despite their importance, cotranscriptional folding pathways have yet to be uncovered with sufficient resolution to reveal how cotranscriptional folding governs RNA structure and function. To access cotranscriptional folding at nucleotide resolution, we extended selective 2′-hydroxyl acylation analyzed by primer-extension sequencing (SHAPE-seq) to measure structural information of nascent RNAs during transcription. Using cotranscriptional SHAPE-seq, we determined how the cotranscriptional folding pathway of the *Bacillus cereus crcB* fluoride riboswitch undergoes a ligand-dependent bifurcation that delays or promotes terminator formation via a series of coordinated structural transitions. Our results directly link cotranscriptional RNA folding to a genetic decision and establish a framework for cotranscriptional analysis of RNA structure at nucleotide resolution.

As nascent RNA molecules exit from RNA polymerase (RNAP), they transition through intermediate structural states that ultimately determine RNA structure and function[1–4]. Because RNA folding generally occurs faster than transcription, the 5′-to-3′ polarity of RNA synthesis directs an order of folding, or cotranscriptional 'folding pathway', that sets the structural stage for many types of interactions that govern cellular processes such as transcription, translation, and macromolecular assembly[5–7].

Cotranscriptional folding has been predicted to be particularly important for bacterial riboswitches[8], a class of regulatory RNAs that control gene expression as a function of specific ligand concentration. Riboswitches contain two functional domains that may structurally overlap: a ligand-binding aptamer and an expression platform that makes regulatory decisions according to the structural state of the aptamer[8,9]. For riboswitches that regulate transcription, ligand binding must influence folding pathways within a short time window in order to commit the riboswitch to one of two mutually exclusive pathways, either promoting or preventing intrinsic terminator-hairpin formation[10–14]. A number of structural studies have revealed the details of the RNA-ligand interactions of many aptamers[8], and biochemical[15,16] and biophysical[13,14] studies using actively transcribing RNAP have indicated distinct RNA structural transitions during transcription. However, a complete nucleotide-resolution understanding of how ligand binding influences the folding pathway of an entire riboswitch and enables it to regulate gene expression remains lacking.

Here we introduce cotranscriptional SHAPE-seq, a method that couples *in vitro* RNAP arrest with high-throughput structure probing to characterize the structures of nascent RNA transcripts at single-nucleotide resolution (**Fig. 1**). Cotranscriptional SHAPE-seq begins with the *in vitro* transcription of a DNA template library that directs the synthesis of each intermediate length of a target RNA (**Fig. 1a**). Each template contains an EcoRI site at the 3′ end that, when bound by the catalytically inactive EcoRI E111Q mutant (Gln111)[17], establishes a transcription roadblock that halts RNAP 14 nt upstream of the EcoRI-binding site[18]. Initiation of single-round transcription from this template library generates halted elongation complexes at all intermediate lengths of the target RNA that, after 30 s of transcription, are rapidly modified with the fast-acting SHAPE reagent benzoyl cyanide (BzCN) (reaction half-life of 250 ms)[19].

After chemical modification, RNAs are extracted and processed for paired-end sequencing according to our previously developed SHAPE-seq v2.1 protocol[20]. Each paired-end read encodes the location of the halted RNAP (nascent RNA 3′ end) and the SHAPE modification position (**Fig. 1b**). Reads are bioinformatically binned on the basis of RNAP position and are used to calculate a SHAPE-seq reactivity spectrum for each intermediate length of the RNA[21]. These reactivities represent flexibilities for every nucleotide within each nascent-RNA-transcript length. High reactivities are indicative of unpaired bases, and low reactivities indicate bases that

are potentially involved in base-pairing, stacking, or ligand interactions[20,22]. Comparison of reactivities at different points during transcription allows for the identification of structural-rearrangement events as transcription proceeds.

RESULTS

Cotranscriptional folding of the *Escherichia coli* SRP RNA

To validate cotranscriptional SHAPE-seq, we first examined the signal recognition particle (SRP) RNA from *E. coli*. The final folded form of the SRP RNA is an extended helical structure containing interspersed inner loops[23] (**Fig. 2a**). Biochemical studies have suggested that the 5′ end forms a labile hairpin structure early during transcription that rearranges into the extended helix only after the 3′ end is synthesized[1]. To determine whether we could observe this rearrangement, we used cotranscriptional SHAPE-seq to obtain a matrix of reactivity spectra for the intermediate lengths of the nascent SRP RNA transcripts (**Fig. 2b**).

Over the course of transcription, changes in nucleotide reactivity patterns (**Fig. 2c**) suggested a series of structural transitions corresponding to the early formation of a stem-loop that ultimately rearranged into the elongated SRP RNA helical structure (**Fig. 2d**). Formation of the early stem-loop structure was apparent as a cluster of highly reactive nucleotides across positions 11–18 in the loop. The pattern of high reactivity persisted until the SRP RNA reached a length of 117 nt, when the sharp drop in reactivity at these positions indicated rearrangement into the elongated helical structure. We observed similar transitions when intermediate SRP RNA fragments were refolded at equilibrium before probing (**Supplementary Fig. 1a**), although the cotranscriptional transitions occurred later, owing to the 14-nt RNAP footprint protecting the RNA 3′ ends[24] (**Supplementary Fig. 1b**).

To further validate cotranscriptional SHAPE-seq, we repeated our analysis of the SRP RNA by using dimethyl sulfate (DMS), an RNA-structure probe that preferentially modifies unstructured A and C bases[25] (**Supplementary Fig. 2**). Overall, the resulting reactivity matrix was consistent with the BzCN data, and we observed similar structural transitions in both data sets. However, the DMS data set showed weaker reactivities at G and U bases, as expected, thus limiting the overall resolution of important structural transitions as a result of the lack of A and C residues in key regions of the SRP RNA sequence.

Ligand-mediated folding of the *B. cereus* fluoride riboswitch

On the basis of our SRP RNA results, we expected cotranscriptional SHAPE-seq to possess the resolution necessary to reveal how alternative folding pathways are controlled during ligand-mediated transcription regulation by a riboswitch. To confirm this capability, we examined the *B. cereus crcB* fluoride riboswitch, which controls transcription by preventing the formation of an intrinsic terminator hairpin in the presence of fluoride[26]. Covariation and equilibrium structural analyses have suggested how the fluoride-bound and unbound forms of the aptamer domain may interact with the downstream expression-platform sequence[26,27] (**Fig. 3a**). However, the specific mechanism by which fluoride binding directs or prevents the folding of the intrinsic terminator during transcription has yet to be elucidated.

To probe the 'off' (terminated) and 'on' (antiterminated) structural states of the *B. cereus* fluoride riboswitch, we generated cotranscriptional SHAPE-seq reactivity matrices in the presence of either 0 mM or 10 mM NaF, respectively (**Fig. 3** and **Supplementary Fig. 3a**). Comparison of the matrices revealed ligand-independent similarities in the initial aptamer folding, which were followed by a bifurcation in the folding pathway that accompanied fluoride binding and fluoride-directed antitermination (**Figs. 3c** and **4a**). Early in transcription, the *B. cereus* fluoride riboswitch folds into two hairpins before the formation of the aptamer, regardless of fluoride concentration. The first hairpin formed within the first ~40 nt of transcription and comprised the P1 stem and a highly reactive loop between nucleotides 11 and 16 (**Supplementary Fig. 4a–c**). The second hairpin formed shortly thereafter and comprised the P3 helix and a loop that exhibited a highly reactive position at U34 and low or moderate reactivities elsewhere (**Supplementary Fig. 4d–f**).

Folding of the P1 and P3 hairpins sets the stage for folding of the aptamer. At a transcript length of 58 nt, we observed a fluoride-independent drop in the reactivity values at nucleotides 12–16 in the P1 loop, a result indicating the formation of pseudoknot PK1 between nucleotides 12–17 and 42–47 as the latter emerged from the RNA-exit channel (**Fig. 4a** and **Supplementary Fig. 5a**). Once PK1 is formed, the aptamer is complete[26,27], thus demonstrating that it first enters a preorganized state independent of ligand, as has been observed for other aptamers[28,29]. From that point, the fate of the aptamer structure is directed by the presence or absence of fluoride.
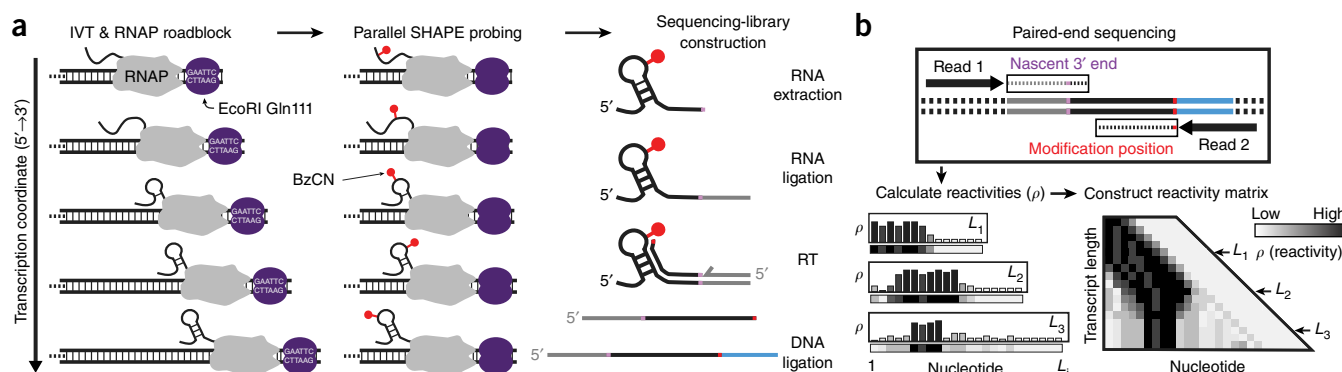

Figure 1 Cotranscriptional SHAPE-seq overview. (**a**) A set of templates are generated, each containing an *E. coli* promoter, a variable-length RNA template, and an EcoRI Gln111 roadblock site. Single-round *in vitro* transcription (IVT) with *E. coli* RNAP is performed with a template library containing a roadblock site at each intermediate transcript length; this is followed by simultaneous SHAPE probing of the arrested complexes and preparation for sequencing. RT, reverse transcription. (**b**) Paired-end sequencing reveals the SHAPE modification position and the 3′ end of each nascent RNA transcript. Reads are binned according to transcript length and are used to calculate SHAPE reactivity profiles that are stacked to generate the reactivity matrix. Increases or decreases in reactivity between transcript lengths ($L_{1–3}$; rows) at particular nucleotides (columns) of this matrix reveal cotranscriptional folding events.
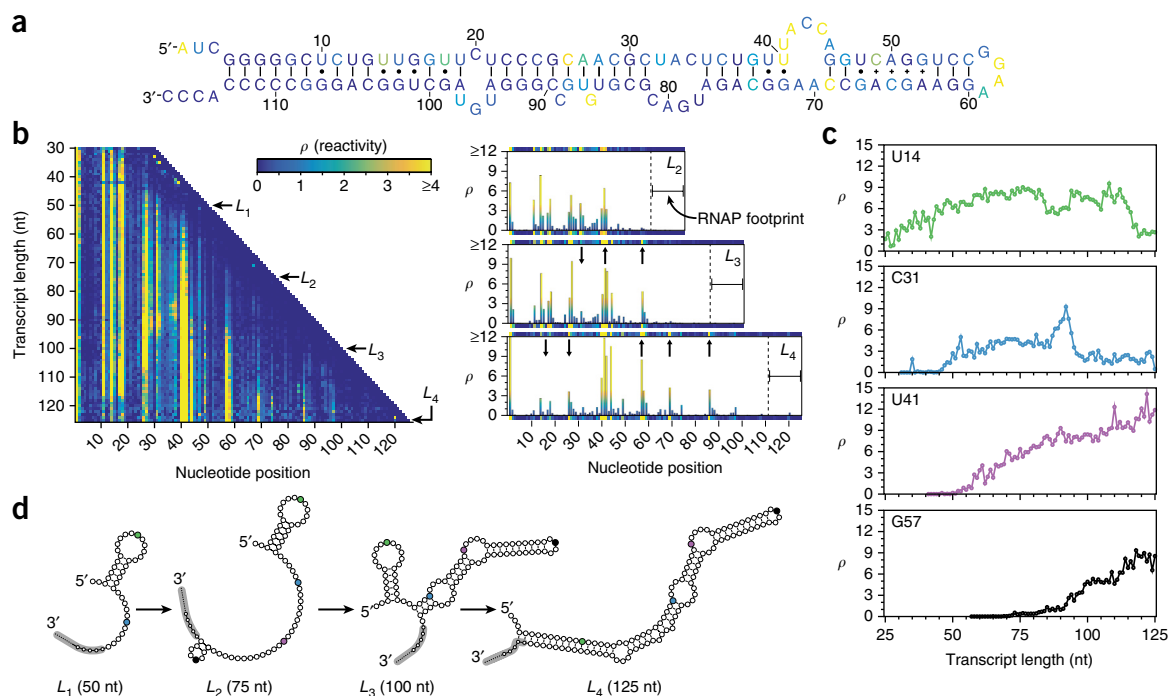
Figure 2 SRP RNA cotranscriptional folding. (**a**) Secondary structure of the final SRP RNA fold, colored according to reactivity intensity at length L_4 (125 nt) and drawn according to a crystal structure determined in ref. 23. (**b**) Cotranscriptional SHAPE-seq reactivity matrix for the SRP RNA folding pathway (left). Lengths L_1, L_2, L_3, and L_4 correspond to 50, 75, 100, and 125 nt, respectively. Selected bar charts and corresponding matrix rows above and below (right) display reactivities for L_2 to L_4. Reactivity changes for $L_2{\rightarrow}L_3$ and $L_3{\rightarrow}L_4$ are marked with arrows. (**c**) Reactivity values for positions U14, C31, U41, and G57 over the course of transcription. U14 undergoes a loop-to-helix transition at a transcript length of 117 nt. Similarly, C31 becomes paired at a transcript length of 96 nt. Plot colors correspond to the marked base positions in **d**, outlining a proposed folding pathway of the SRP RNA that is consistent with these transitions. The 14-nt RNAP footprint[24] for each length is indicated in gray, and small circles indicate the ~5 nt in the RNA-exit channel. Results shown in **b** and **c** are derived from one representative of three independent experiments (additional replicates in **Supplementary Fig. 9a**). Source data for **b** and **c** are available online.

The first steps in the fluoride-dependent bifurcation of the folding pathway involve fluoride-mediated aptamer stabilization. In the presence of fluoride, the P1 loop reactivities continued to decrease until a transcript length of 69 nt, thus suggesting that fluoride binding stabilizes the pseudoknot (**Fig. 4a** and **Supplementary Fig. 5a**). Stabilization of the fluoride-binding pocket also requires a long-range noncanonical base pair between U38 and A10 (**Fig. 3a** and refs. 26,27), the latter of which is paired in the P1 stem before PK1 formation. In the presence of fluoride, we first observed an increase in A10 reactivity, thus indicating that PK1 formation disrupts its base pair within the P1 stem. A10 then transitioned to low reactivity at a transcript length of 58 nt, thereby indicating the formation of a long-range interaction with U38 after PK1 formation (**Fig. 4a**). In the absence of fluoride, the sustained high reactivity at A10 suggested that its interaction with U38 is not favored without ligand. Our observation of fluoride-induced aptamer stabilization was further supported by distinct reactivity changes in nucleotides 22–27, which join the P1 and P3 helices but do not participate in any pairing interactions in the analogous *Thermotoga petrophila* aptamer domain[27]. Specifically, nucleotides 24, 25, and 27 displayed lower reactivities in the presence of fluoride (**Supplementary Fig. 5b**) after a transcript length of 58 nt. In contrast, A22 underwent a dramatic reactivity spike upon PK1 formation when fluoride was present but underwent only a modest increase in the absence of fluoride (**Fig. 4a**), thus revealing that A22 hyper-reactivity is a strong indicator of aptamer state. Together, these results support a model in which the pseudoknot forms the basis of the aptamer, which can undergo further coordinated restructuring

after fluoride binding and consequently form a more stable structure with additional interactions.

After aptamer formation, the riboswitch follows one of two ligand-dependent folding trajectories that direct transcription termination or antitermination. When RNAP reaches nucleotide 69, the loop of the terminator hairpin begins to emerge from the RNA-exit channel (**Fig. 4a**). Without fluoride, the terminator hairpin nucleated at nucleotide 77, as observed as a decrease in reactivity in the upper terminator stem (nucleotides 52–55) (**Fig. 4a** and **Supplementary Fig. 5c**). Increased reactivity in the P1 loop (nucleotides 12–16) (**Supplementary Fig. 5a**) and decreased reactivity at A22 occurred concurrently, thus indicating that PK1 opens and consequently dissolves the metastable aptamer (**Fig. 4a**).

In the presence of fluoride, the stabilized aptamer promoted antitermination in two ways: (i) disfavoring complete terminator formation by sequestering part of the terminator hairpin (**Fig. 3a**) and (ii) delaying terminator-hairpin nucleation until nucleotide 88, after RNAP had transcribed past the poly(U) sequence (**Fig. 4a** and **Supplementary Fig. 5c**). When the terminator hairpin did begin to form, high reactivities at U48 and nucleotides 69–74 indicated that only the top half of the terminator formed (**Figs. 3** and **4a**), and the ribosome-binding site (RBS; nucleotides 67–72) was left accessible for translation of the downstream fluoride transporter[26] (**Supplementary Fig. 5d**). Thus, fluoride binding fundamentally alters the cotranscriptional folding pathway of the *B. cereus* fluoride riboswitch by stabilizing an RNA structure that promotes transcription via antitermination and translation by RBS exposure.
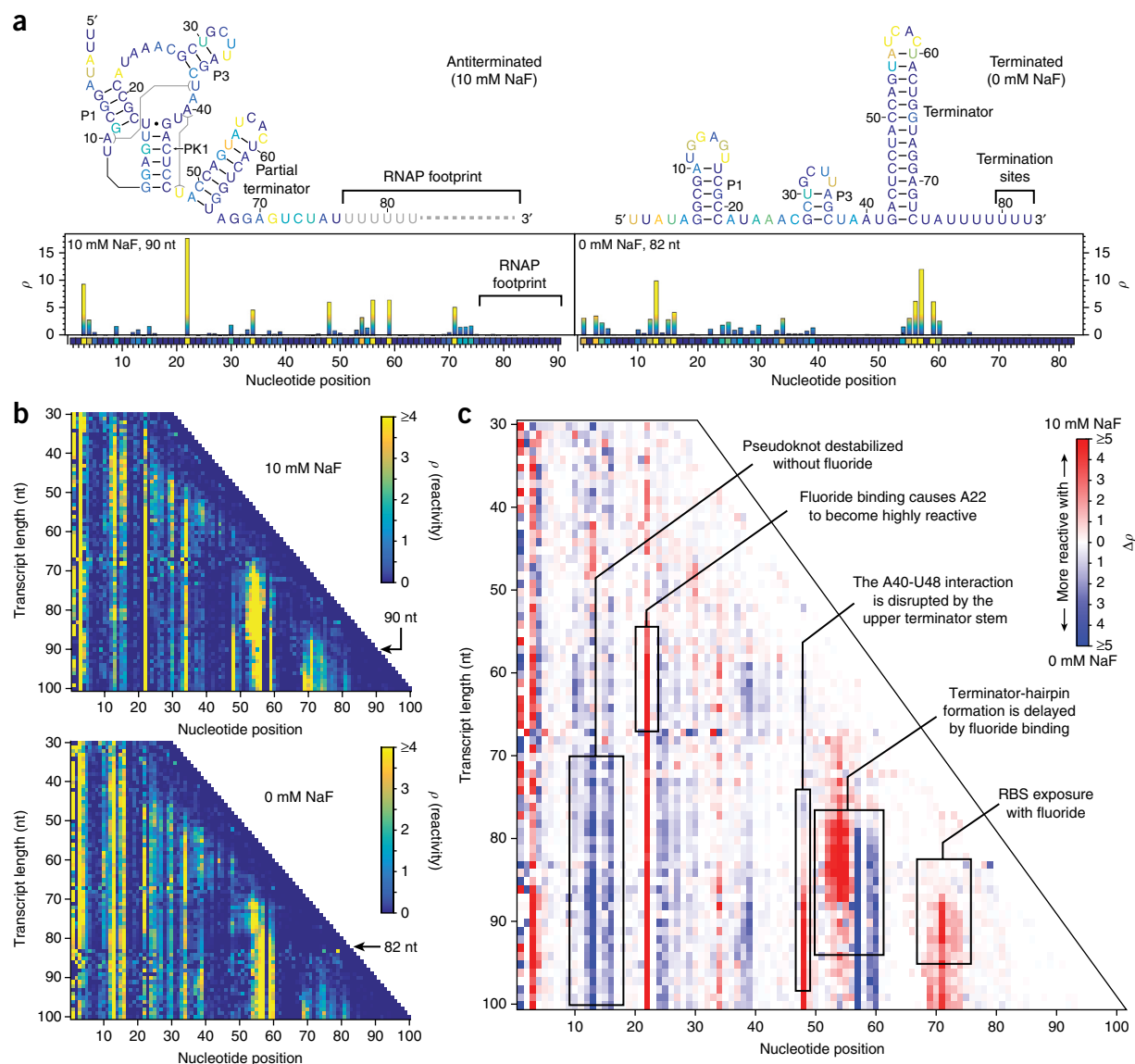
Figure 3 *B. cereus* fluoride-riboswitch cotranscriptional SHAPE-seq data. (**a**) The antiterminated and terminated folds[26] of the fluoride riboswitch (top), colored according to the reactivity values at transcript lengths of 90 nt and 82 nt, respectively (bottom). (**b**) Cotranscriptional SHAPE-seq reactivity matrices for the fluoride riboswitch transcribed with 10 mM (top) or 0 mM NaF (bottom). (**c**) Reactivity differences ($\Delta\rho$) between the matrices in **b**, annotated according to folding events during transcription. The reactivity changes that occur over the course of transcription suggest that the fluoride riboswitch traverses two cotranscriptional folding pathways depending on the presence of ligand. Results in **b** and **c** are derived from one representative of three independent experiments (additional replicates in **Supplementary Fig. 9c,d**). Source data are available online.

Riboswitch mutants alter folding pathways in defined ways

Previous work on the *B. cereus* fluoride riboswitch has examined mutations in the aptamer and terminator stem[26]. Each mutation confers distinct changes in ligand binding and termination capability that correspond to its location in the riboswitch (**Supplementary Fig. 3b–d**). Therefore, we used these mutants to both corroborate our interpretations of the wild-type (WT) fluoride-riboswitch cotranscriptional SHAPE-seq data (**Figs. 3** and **4**) and uncover details of how individual mutations affect the cotranscriptional folding and function of the riboswitch (**Fig. 5** and **Supplementary Figs. 6** and **7**).

We first sought to corroborate our interpretation of structural transitions that lead to aptamer formation by analyzing mutants that disrupt PK1 formation. Mutant M19 (U45A C46U) (**Fig. 5b,c**) disrupts base-pairing in both PK1 and the terminator stem. Mutants M18 (G13A A14U) (**Supplementary Fig. 6a–c**) and M22 (M19 and

M20 mutations) (**Supplementary Fig. 6d–f**) disrupt PK1 folding but maintain base-pairing in the terminator stem. Consequently, M18, M19, and M22 are unable to properly form the fluoride aptamer and were therefore fluoride insensitive, as observed in functional *in vitro* termination assays (**Supplementary Fig. 3d**). From a structural perspective, we observed this defect in aptamer formation as a lack of pronounced differences in cotranscriptional SHAPE-seq data sets for these mutants in the presence or absence of fluoride, and as an absence of transitions reflecting aptamer formation and stabilization. The most notable signature of fluoride insensitivity was consistently high P1-loop reactivities in either the presence or absence of fluoride, thus indicating that PK1 did not form in these mutants. Furthermore, A22 remained weakly reactive compared with the WT riboswitch across all transcript lengths (**Fig. 5b,c** and **Supplementary Fig. 6**), a result further supporting the interpretation that reactivity measurements
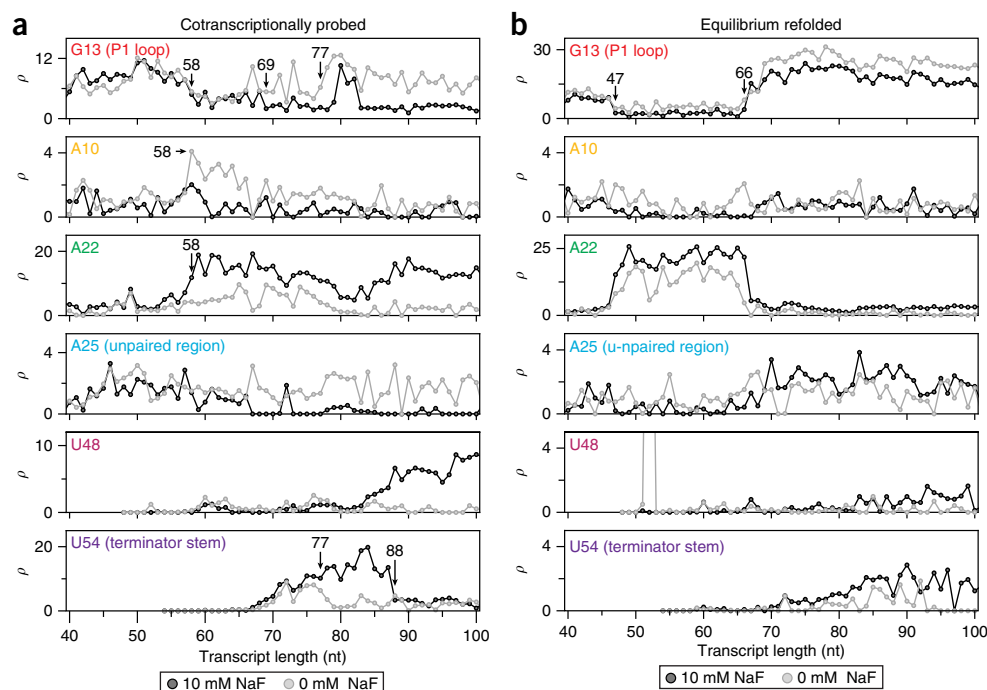
Figure 4 Changes in reactivities over transcript lengths for single nucleotides highlight structural transitions in the cotranscriptional folding pathway of the *B. cereus* fluoride riboswitch. (**a**) Single-nucleotide reactivity trajectories for nucleotides involved in several key structural transitions when transcribed with either 0 mM (gray) or 10 mM NaF (black). The trajectories diverge at lengths at which structural transitions occur. Results shown are extracted from the matrices presented in **Figure 3b**. (**b**) As in **a**, except that the RNAs were transcribed, extracted, denatured, and equilibrium-refolded in transcription buffer with either 0 mM (gray) or 10 mM NaF (black) before SHAPE modification. The lack of divergence between the equilibrium refolding trajectories indicates that cotranscriptional folding is required to obtain alternate ligand-dependent structures. Results shown are extracted from the matrices presented in **Supplementary Figure 8a**. Source data are available online.

at A22 are a strong indicator of aptamer state. Together, these data support the conclusion that aptamer formation is necessary for the ligand-dependent bifurcation of the fluoride-riboswitch folding pathway that leads to the regulation of terminator formation.

We next sought to examine how interactions between the aptamer and intrinsic terminator coordinate structural changes leading to the bifurcation of the folding pathway. Mutant M20 (G69A A70U) contains base substitutions in the 3′ terminator stem that render the intrinsic terminator nonfunctional but does not disrupt the WT aptamer domain (**Supplementary Fig. 3b–d**). Mutant M21 (M18 and M19 mutations) restores PK1 base-pairing but, like M20, does not form a complete terminator stem. As anticipated, the M20 and M21 mutants were nonfunctional, owing to inactivation of the terminator, but underwent all transitions associated with aptamer formation and fluoride binding (**Supplementary Fig. 7a–c**). Interestingly, M21 displayed consistently high P1-loop reactivity in the absence of fluoride and showed delayed formation in the presence of fluoride, thus suggesting that PK1 does not stabilize as readily for this mutant, probably because of the replacement of a G-C base pair in PK1 with A-U. However, the characteristic increase in A22 still occurred, thus suggesting PK1 still formed, but in a smaller fraction of the folded population. Together, these results indicate that M20 and M21 undergo nearly all of the transitions observed in the WT system and are nonfunctional only because the mutations disrupt the final event in termination formation.

To further corroborate our observations of the WT system, we analyzed the M23 mutant (M18, M19 and M20 mutations), which restores base-pairing in both PK1 and the terminator stem (**Fig. 5d,e**). M23 was able to form the fluoride aptamer and was therefore capable of binding fluoride and, in the absence of fluoride, of terminating

transcription at near-WT levels (**Supplementary Fig. 3d**). As expected, we observed similar reactivity patterns within the matrices from M23 and the WT riboswitches during the formation of the aptamer domain; this result indicated that restoring base-pairing was sufficient to reproduce a near-native folding pathway (**Fig. 5d,e**). However, similarly to M21, M23 contains weakened base-pairing in PK1 and does not form a stable aptamer as readily as the WT riboswitch in the absence of fluoride, as seen in the absence of an increase in reactivity at A10 near a transcript length of 58 nt. The analysis of M23 demonstrates that the overall folding pathway and decision-making events of the fluoride riboswitch can be enacted by different RNA sequences.

Cotranscriptional SHAPE-seq accesses kinetically trapped RNA structures

We next sought to assess whether cotranscriptional SHAPE-seq accesses nonequilibrium kinetically trapped folded states of nascent RNAs. To do so, we directly compared our results to an equilibrium analysis whereby RNAs were equilibrium-refolded before probing. To equilibrium-refold all fluoride-riboswitch intermediates, we first generated all of the intermediate transcript lengths, as done for cotranscriptional SHAPE-seq, but extracted, denatured, and equilibrium-refolded the RNAs in transcription buffer before chemical probing.

In our comparison, we found two main differences between the cotranscriptional and equilibrium experiments that indicated that the cotranscriptional experiments probe nonequilibrium folding states (**Fig. 4b** and **Supplementary Fig. 8**). Because cotranscriptional SHAPE-seq experiments probe RNAs that exist as part of a transcription elongation complex, only folding events involving nucleotides that have left, or are within, the RNA-exit channel of RNAP can be observed.
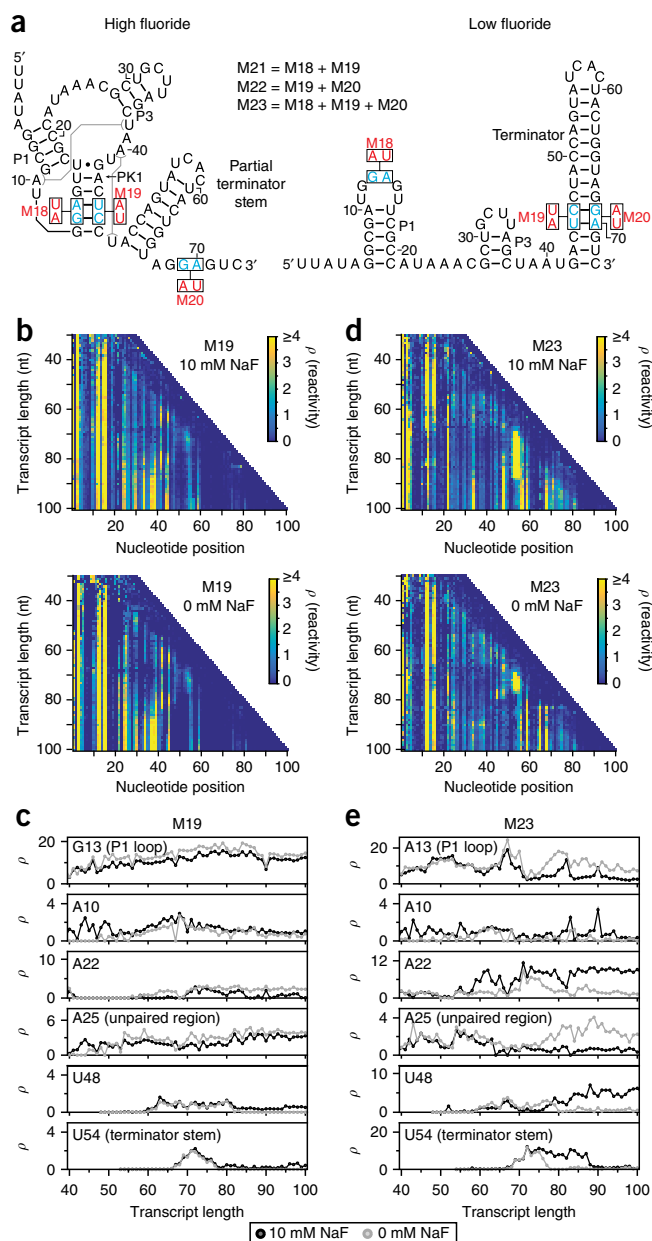
Figure 5 Cotranscriptional SHAPE-seq analysis of *B. cereus* fluoride-riboswitch mutants. (**a**) The locations of mutations M18–M23 from ref. 26 within the antiterminated (high-fluoride) and terminated (low-fluoride) secondary structures for the WT system (as in **Fig. 3a**). (**b**) Reactivity matrices of the M19 mutant (pseudoknot and terminator base-pairing disrupted) transcribed with either 10 mM (top) or 0 mM NaF (bottom). (**c**) Single-nucleotide cotranscriptional reactivity trajectories for the same key nucleotides highlighted in **Figure 4**, for the M19 mutant transcribed with either 0 mM (gray) or 10 mM NaF (black). (**d**) As in **b**, but for mutant M23 (pseudoknot and terminator base-pairing restored through compensatory mutations). (**e**) As in **c**, but for mutant M23. Results were obtained from one experiment. Source data for **b**–**e** are available online.

During equilibrium refolding, however, RNAP is not present to prevent the 3′ end of the transcript from folding with the rest of the RNA, thus causing the structural transitions to occur at shorter transcript lengths than those observed with cotranscriptional SHAPE-seq. Specifically, we observed ligand-independent folding of PK1 (via decreases in P1-loop reactivities) at a transcript length of 47 nt with equilibrium refolding

(**Fig. 4b** and **Supplementary Fig. 8**) as opposed to the length of 58 nt (**Figs. 3** and **4a**) observed in cotranscriptional probing of the arrested elongation complexes. In cotranscriptional SHAPE-seq experiments, a transcript length of 58 nt corresponds to a point at which nucleotide 47 is expected to be leaving the RNA-exit channel. We also observed shifts in structural transitions to earlier lengths during equilibrium refolding for the SRP RNA (**Supplementary Fig. 1a,b**).

The second piece of evidence supporting the capture of nonequilibrium RNA structures is the distinct fluoride-independent behavior of the P1 loop at longer transcript lengths in the equilibrium refolding experiments. Specifically, equilibrium refolding produced a sharp fluoride-independent rise in the reactivities of the P1-loop nucleotides over transcript lengths of 66–69 nt, thus indicating the complete opening of PK1 in either the absence or presence of ligand (**Fig. 4b**). However, in the cotranscriptional SHAPE-seq experiment, PK1 remained stabilized at longer transcript lengths in the presence of fluoride. The observed deviation from equilibrium structures indicated that the RNAs probed in cotranscriptional SHAPE-seq were out of equilibrium.

From these results, we concluded that equilibrium refolding does not permit meaningful analysis of the cotranscriptional folding pathways of the fluoride riboswitch beyond a transcript length of 66 nt, because terminator-hairpin formation is thermodynamically favored regardless of fluoride concentration. However, this observation is crucial because it reveals the relative thermodynamic favorability of the terminator structure over the aptamer. Furthermore, in the terminated structure, G68 pairs with C47, and the latter participates in the last base pair of PK1 within the aptamer. Thus, the loss of a single base pair in PK1 to pairing in the terminator stem is sufficient to preclude aptamer folding in favor of terminator folding. This thermodynamic tipping point corresponds well to the cotranscriptional opening of PK1 at a transcript length of 77 nt in the absence of fluoride. During transcription without fluoride, the instability of PK1 may allow terminator base-pairing to extend into the RNA-exit channel, thereby favoring the terminated structure.

A model of the folding pathway and decision-making process of the *B. cereus* fluoride riboswitch that combines these observations with the cotranscriptional analysis of the WT and mutant systems is shown in **Figure 6**.

DISCUSSION

We developed cotranscriptional SHAPE-seq to facilitate the experimental characterization of cotranscriptional RNA folding at nucleotide resolution. Furthermore, we demonstrated that cotranscriptional SHAPE-seq interrogates the structure of nonequilibrium, kinetically trapped nascent RNAs and consequently generates functionally meaningful structural information. Replicate data sets for the *E. coli* SRP RNA and the *B. cereus crcB* fluoride riboswitch showed that the technique is highly reproducible (**Supplementary Fig. 9**).

When interpreting cotranscriptional SHAPE-seq data, it is important to consider the relative timescales of RNA folding and RNA synthesis as well as the advantages and limitations of chemical RNA-structure probing. Whereas the uninterrupted nucleotide-addition cycle of *E. coli* RNAP occurs at 50–100 nt/s (10–20 ms/nt)[30], simple RNA folding events such as base-pair melting occur on a microsecond-to-millisecond timescale[31], and larger conformation changes occur on the order of seconds to tens of seconds. Although it is desirable, interrogation of RNA folding on the microsecond-to-millisecond timescale is inaccessible with even the most fast-acting SHAPE probe, BzCN, which reacts with a half-life of 250 ms (ref. 19). This inaccessibility is further exacerbated by chemical probes, such as DMS, with reaction times on the orders of minutes[32]. Furthermore, in its current form, the temporal resolution of cotranscriptional SHAPE-seq is restricted by the manual manipulation of samples
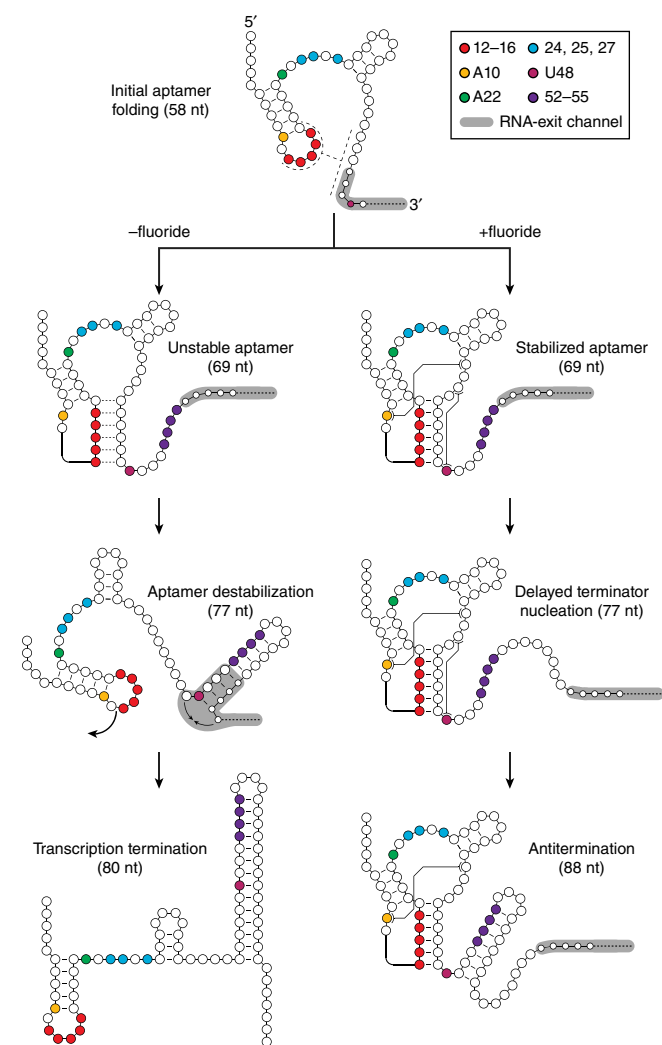
Figure 6 A model for ligand-dependent cotranscriptional folding of the *B. cereus* fluoride riboswitch. The folding pathway for the fluoride riboswitch begins with initial aptamer folding. If fluoride binds (right), the prefolded aptamer then stabilizes through specific interactions, thereby leading to delays in the early folding stages of the intrinsic terminator hairpin, which does not nucleate until RNAP has escaped the poly(U) tract. However, if there is no fluoride binding (left), the top of the terminator hairpin quickly folds, thereby disrupting the pseudoknot and reaching into the RNA-exit channel and allowing the full terminator hairpin to trigger transcription termination. Intermediate structural states are inferred from cotranscriptional SHAPE-seq reactivities, covariation analysis[26], and crystallographic data[27].

and the need to halt transcription elongation complexes before RNA probing. Thus, the power of cotranscriptional SHAPE-seq lies not in temporal sensitivity but in the capability of capturing single-nucleotide-resolution 'snapshots' of kinetically trapped intermediate folds that have not equilibrated within the seconds timescale of the experiment. As such, it complements single-molecule force spectroscopy, which measures RNA structural changes during transcription with a high temporal resolution but a low spatial resolution[14]. Despite its limitations, as described above, the resulting SHAPE-seq reactivity profiles facilitate the identification of distinct nucleotide signatures that indicate the nascent-RNA structural state and the observation of major conformational rearrangements that occur as the RNA sequence is synthesized.

Although cotranscriptional SHAPE-seq occurs on timescales much longer than the rate of transcription, this aspect does not preclude its

use in analyzing the role of transcription dynamics in RNA folding. Transcription pausing, which can occur on the seconds timescale, has been shown to play a critical role in RNA folding through studies of several RNAs, including the *E. coli* SRP RNA[1] and the *btuB* coenzyme B12 riboswitch[33]. By itself, cotranscriptional SHAPE-seq does not directly measure transcription pausing, because transcription roadblocks obscure the observation of pause distributions, and the saturating NTP concentrations used are not conducive to precise measurement of transcription pausing[34]. However, the structural states probed at a specific length of RNA would reflect any folding afforded by a transcription pause upstream of that length. Therefore, the consequences of altering native transcription dynamics on RNA folding would be observable in cotranscriptional SHAPE-seq data sets.

Finally, it is important to consider that the stochastic nature of cotranscriptional RNA folding produces an ensemble of structural conformations. Because SHAPE-seq measurements are made in bulk, reactivity profiles reflect an average over this ensemble. Interestingly, this average can be seen by comparing equilibrium-refolded SHAPE-seq matrices and cotranscriptional SHAPE-seq matrices: the latter generally show blurrier transitions that probably reflect the probing of a subpopulation of molecules that have not yet made the specific transitions (comparison of **Fig. 2b** with **Supplementary Fig. 1a**, and of **Fig. 3b** with **Supplementary Fig. 8a**). Detection of modification sites by mutational profiling (MaP)[35] may be able to access information about the folding pathways of subpopulations if the MaP strategy can robustly detect the modification of RNA by BzCN.

We used cotranscriptional SHAPE-seq to evaluate the RNA folding pathway of a fluoride riboswitch and in this context identified molecular signatures of aptamer folding and ligand binding as well as the precise point at which the nascent RNA mediates a genetic decision (**Fig. 6**). In agreement with force spectroscopy analysis of *pbuE* adenine-riboswitch cotranscriptional folding[14] and biochemical analysis of the folding pathways of the *btuB* coenzyme B12 riboswitch[33], our results indicated that the *crcB* fluoride riboswitch is controlled by the kinetics of its cotranscriptional folding events. Despite this similarity, there are key differences in the cotranscriptional folding pathways of the adenine and fluoride riboswitches, at the levels of both ligand sensing and the stability of the ligand-bound aptamer, which suggest that these riboswitches use different overall strategies to make gene-regulatory decisions. One clear difference is in aptamer folding. Whereas the *pbuE* adenine aptamer has not been observed to fold in the absence of adenine[14], our data indicate that the *crcB* fluoride aptamer adopts an 'unstable' folded state in the absence of fluoride on the timescales of our experiment (**Fig. 4a** and **Supplementary Fig. 5a**). The fate of the fluoride-bound *crcB* aptamer is also distinct from the fate of the adenine-bound *pbuE* aptamer, which rearranges into an aptamer-disrupted structure after antitermination. In contrast, the *crcB* fluoride-riboswitch aptamer is kinetically trapped and remains stable for at least 30 s in the presence of fluoride, even after terminator nucleation (**Figs. 3** and **4a**). The origin of this distinction may be functional in nature: whereas the *pbuE* adenine riboswitch functions at the transcriptional level, the *crcB* fluoride riboswitch may exert dual transcriptional and translational control over gene expression by simultaneously preventing complete terminator folding and exposing a ribosome-binding site that would otherwise be sequestered within the terminator hairpin (**Fig. 5d**). These mechanistic distinctions between folding of fluoride and adenine riboswitches suggest diversity in the mechanisms through which cotranscriptional RNA folding can direct targeted RNA functions, such as ligand sensing and genetic control. A fundamental understanding of the relationship between RNA folding and function may be achieved with a broad characterization of the cotranscriptional folding pathways

of diverse riboswitches and comparative studies of the conservation of RNA folding pathways for specific riboswitch variants, both of which are now afforded by our cotranscriptional SHAPE-seq technique.

The work presented here provides an experimental means through which to answer fundamental questions about how the cotranscriptional nature of RNA folding directs RNA structure and function. We anticipate that the integrated analysis of measurements made with cotranscriptional SHAPE-seq and complementary biophysical, biochemical, and computational techniques should provide a powerful framework with which to understand the roles of cotranscriptional folding in regulating broader cellular processes.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Accession codes. Raw sequencing data that support the findings of this study have been deposited in the Sequence Read Archive under BioProject PRJNA342175. Individual BioSample accession codes are available in **Supplementary Table 1**. SHAPE-seq reactivity spectra generated in this work have been deposited in the RNA Mapping Database under accession codes FLUORSW_BZCN_0001, FLUORSW_BZCN_0002, FLUORSW_BZCN_0003, FLUORSW_BZCN_0004, FLUORSW_BZCN_0005, FLUORSW_BZCN_0006, FLUORSW_BZCN_0007, FLUORSW_BZCN_0008, FLUORSW_BZCN_0009, FLUORSW_BZCN_0010, FLUORSW_BZCN_0011, FLUORSW_BZCN_0012, FLUORSW_BZCN_0013, FLUORSW_BZCN_0014, FLUORSW_BZCN_0015, FLUORSW_BZCN_0016, FLUORSW_BZCN_0017, FLUORSW_BZCN_0018, FLUORSW_BZCN_0019, FLUORSW_BZCN_0020, SRPECLI_BZCN_0001, SRPECLI_BZCN_0002, SRPECLI_BZCN_0003, SRPECLI_BZCN_0004, SRPECLI_DMS_0001, SRPECLI_DMS_0002, and SRPECLI_DMS_0003. Sample details are available in **Supplementary Table 2**.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

1. Wong, T.N., Sosnick, T.R. & Pan, T. Folding of noncoding RNAs during transcription facilitated by pausing-induced nonnative structures. *Proc. Natl. Acad. Sci. USA* **104**, 17995–18000 (2007).
2. Heilman-Miller, S.L. & Woodson, S.A. Effect of transcription on folding of the *Tetrahymena* ribozyme. *RNA* **9**, 722–733 (2003).
3. Kramer, F.R. & Mills, D.R. Secondary structure formation during RNA synthesis. *Nucleic Acids Res.* **9**, 5109–5124 (1981).
4. Pan, T., Artsimovitch, I., Fang, X.W., Landick, R. & Sosnick, T.R. Folding of a large ribozyme during transcription and the effect of the elongation factor NusA. *Proc. Natl. Acad. Sci. USA* **96**, 9545–9550 (1999).
5. Pan, T. & Sosnick, T. RNA folding during transcription. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 161–175 (2006).
6. Russell, R. *et al.* Exploring the folding landscape of a structured RNA. *Proc. Natl. Acad. Sci. USA* **99**, 155–160 (2002).
7. Al-Hashimi, H.M. & Walter, N.G. RNA dynamics: it is about time. *Curr. Opin. Struct. Biol.* **18**, 321–329 (2008).
8. Garst, A.D., Edwards, A.L. & Batey, R.T. Riboswitches: structures and mechanisms. *Cold Spring Harb. Perspect. Biol.* **3**, a003533 (2011).
9. Serganov, A. & Nudler, E. A decade of riboswitches. *Cell* **152**, 17–24 (2013).
10. 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).
11. Rieder, U., Kreutz, C. & Micura, R. Folding of a transcriptionally acting preQ1 riboswitch. *Proc. Natl. Acad. Sci. USA* **107**, 10804–10809 (2010).
12. Mandal, M. & Breaker, R.R. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat. Struct. Mol. Biol.* **11**, 29–35 (2004).
13. Haller, A., Altman, R.B., Soulière, M.F., Blanchard, S.C. & Micura, R. Folding and ligand recognition of the TPP riboswitch aptamer at single-molecule resolution. *Proc. Natl. Acad. Sci. USA* **110**, 4188–4193 (2013).
14. Frieda, K.L. & Block, S.M. Direct observation of cotranscriptional folding in an adenine riboswitch. *Science* **338**, 397–400 (2012).
15. Wickiser, J.K. Kinetics of riboswitch regulation studied by in vitro transcription. *Methods Mol. Biol.* **540**, 53–63 (2009).
16. Wong, T.N. & Pan, T. RNA folding during transcription: protocols and studies. *Methods Enzymol.* **468**, 167–193 (2009).
17. Wright, D.J., King, K. & Modrich, P. The negative charge of Glu-111 is required to activate the cleavage center of EcoRI endonuclease. *J. Biol. Chem.* **264**, 11816–11821 (1989).
18. Pavco, P.A. & Steege, D.A. Elongation by *Escherichia coli* RNA polymerase is blocked *in vitro* by a site-specific DNA binding protein. *J. Biol. Chem.* **265**, 9960–9969 (1990).
19. Mortimer, S.A. & Weeks, K.M. Time-resolved RNA SHAPE chemistry: quantitative RNA structure analysis in one-second snapshots and at single-nucleotide resolution. *Nat. Protoc.* **4**, 1413–1421 (2009).
20. Watters, K.E., Yu, A.M., Strobel, E.J., Settle, A.H. & Lucks, J.B. Characterizing RNA structures *in vitro* and *in vivo* with selective 2′-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Methods* **103**, 34–48 (2016).
21. Aviran, S., Lucks, J.B. & Pachter, L. RNA structure characterization from chemical mapping experiments. *Proc. 49th Annual Allerton Conf. on Communication, Control, and Computing* 1743–1750 (Monticello, Illinois, USA, 2011).
22. Bindewald, E. *et al.* Correlating SHAPE signatures with three-dimensional RNA structures. *RNA* **17**, 1688–1696 (2011).
23. Batey, R.T., Rambo, R.P., Lucast, L., Rha, B. & Doudna, J.A. Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* **287**, 1232–1239 (2000).
24. Komissarova, N. & Kashlev, M. Functional topography of nascent RNA in elongation intermediates of RNA polymerase. *Proc. Natl. Acad. Sci. USA* **95**, 14699–14704 (1998).
25. Ehresmann, C. *et al.* Probing the structure of RNAs in solution. *Nucleic Acids Res.* **15**, 9109–9128 (1987).
26. Baker, J.L. *et al.* Widespread genetic switches and toxicity resistance proteins for fluoride. *Science* **335**, 233–235 (2012).
27. Ren, A., Rajashankar, K.R. & Patel, D.J. Fluoride ion encapsulation by Mg^{2+} ions and phosphates in a fluoride riboswitch. *Nature* **486**, 85–89 (2012).
28. Ali, M., Lipfert, J., Seifert, S., Herschlag, D. & Doniach, S. The ligand-free state of the TPP riboswitch: a partially folded RNA structure. *J. Mol. Biol.* **396**, 153–165 (2010).
29. Ottink, O.M. *et al.* Ligand-induced folding of the guanine-sensing riboswitch is controlled by a combined predetermined induced fit mechanism. *RNA* **13**, 2202–2212 (2007).
30. Roberts, J.W., Shankar, S. & Filter, J.J. RNA polymerase elongation factors. *Annu. Rev. Microbiol.* **62**, 211–233 (2008).
31. Mustoe, A.M., Brooks, C.L. & Al-Hashimi, H.M. Hierarchy of RNA functional dynamics. *Annu. Rev. Biochem.* **83**, 441–466 (2014).
32. Tijerina, P., Mohr, S. & Russell, R. DMS footprinting of structured RNAs and RNA–protein complexes. *Nat. Protoc.* **2**, 2608–2623 (2007).
33. Perdrizet, G.A. II, Artsimovitch, I., Furman, R., Sosnick, T.R. & Pan, T. Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proc. Natl. Acad. Sci. USA* **109**, 3323–3328 (2012).
34. 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).
35. Siegfried, N.A., Busan, S., Rice, G.M., Nelson, J.A. & Weeks, K.M. RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959–965 (2014).

ONLINE METHODS

Plasmids. Plasmids used for DNA-template synthesis contained a chloramphenicol-resistance gene, the p15A origin of replication, and a consensus *E. coli* σ^{70} promoter followed by a sequence encoding the RNA under study. The *E. coli* SRP RNA sequence (GenBank NC_000913.3, bases 476448 to 476561) with the sequence ATC appended at the 5′ end[1] was cloned upstream of the antigenomic hepatitis δ ribozyme. The *B. cereus crcB* fluoride riboswitch (GenBank AE017194.1, bases 4763724 to 4763805) was cloned upstream of a consensus ribosome-binding site and the superfolder GFP (SFGFP) sequence. These non-native downstream sequences were used to allow transcription to proceed far enough that the full length RNA of interest would emerge from RNAP. These sequences did not influence cotranscriptional SHAPE-seq interpretations, because lengths at which non-native RNA had emerged from RNAP were not used to draw conclusions. The fluoride-riboswitch mutants were derived from the plasmid described above.

Proteins. EcoRI E111Q (Gln111) was a generous gift from J. Roberts (Cornell University) and J. Filter (Cornell University).

Template preparation. DNA template libraries (**Supplementary Table 3**) for cotranscriptional SHAPE-seq were prepared by combining individual PCR amplifications of each RNA template length. Each 25-μL PCR reaction included 20.4 μL of H₂O, 2.5 μL of 10× Standard Taq Reaction Buffer (New England BioLabs), 0.5 μL of 10 mM dNTPs, 0.25 μL of 100 μM oligo J (forward primer; **Supplementary Table 4**), 0.1 μL of plasmid DNA template, 0.25 μL of Taq DNA polymerase (New England BioLabs), and 1 μL of 25 μM reverse primer (**Supplementary Tables 4 and 5**). The reverse primer incorporated an EcoRI site. Reaction mixes were run with a standard thermal-cycle program consisting of 30 cycles of amplification at an annealing temperature of 55 °C. After thermal cycling, PCR reactions were pooled, mixed, and split into 500-μL aliquots before addition of 50 μL of 3 M NaOAc, pH 5.5, and 1 mL of 100% EtOH for EtOH precipitation. Precipitated pellets were dried with a SpeedVac and pooled by dissolving all pellets in 30 μL of H₂O. The template was then run on a 1% agarose gel and extracted with a QIAquick Gel Extraction Kit (Qiagen). The concentration of the purified template was determined with a Qubit Fluorometer (Life Technologies), and the molarity of the template was calculated on the basis of the median template length.

Single-length DNA templates (**Supplementary Table 3**) were prepared through five 100-μL PCR reactions including 82.75 μL of H₂O, 10 μL of 10× Standard Taq Reaction Buffer (New England BioLabs), 1.25 μL of 10 mM dNTPs, 2.5 μL of 10 μM of oligo J (forward primer; **Supplementary Table 4**), 2.5 μl of 10 μM oligo K (**Supplementary Table 4**), 0.5 μL of plasmid DNA template, and 0.5 of μL Taq DNA polymerase (New England BioLabs). Reactions were run with the thermal-cycling program described above. After thermal cycling, reactions were pooled before the addition of 50 μL of 3 M NaOAc, pH 5.5, and 1 mL of 100% EtOH for EtOH precipitation. The precipitated pellet was dried with a SpeedVac and dissolved in 30 μL of H₂O. The template was then run on a 1% agarose gel and extracted with a QIAquick Gel Extraction Kit (Qiagen). The concentration of the purified template was determined with a Qubit 2.0 Fluorometer (Life Technologies).

***in vitro* transcription (single length, radiolabeled).** 25-μL reaction mixtures containing 5 nM linear DNA template (described above) and 0.5 U of *E. coli* RNAP holoenzyme (New England BioLabs) were incubated in transcription buffer (20 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 1 mM DTT, and 50 mM KCl), 0.1 mg/mL bovine serum albumin, 200 μM ATP, GTP, and CTP, and 50 μM UTP containing 0.5 μCi/μL [α-³²P]UTP for 10 min at 37 °C to form open complexes. When present, NaF was included at a final concentration of 1 μM, 10 μM, 100 μM, 1 mM or 10 mM, as indicated in **Supplementary Figure 3a**. Single-round transcription reactions were initiated by the addition of MgCl₂ to 5 mM and rifampicin to 10 μg/mL. Transcription was stopped by the addition of 125 μL of stop solution (0.6 M Tris, pH 8.0, 12 mM EDTA, and 0.16 mg/mL tRNA).

RNA from stopped transcription reactions was purified by the addition of 150 μL of phenol/chloroform/isoamyl alcohol (25:24:1 (v/v)), vortexing, centrifugation, and collection of the aqueous phase, which was then ethanol precipitated through the addition of 450 μL of 100% ethanol to each reaction and storage at −20 °C overnight. Precipitated RNA was resuspended in transcription loading dye (1× transcription buffer, 80% formamide, 0.05% bromophenol blue, and xylene cyanol). Reactions were fractionated through electrophoresis on 12% denaturing polyacrylamide gels containing 7.5 M urea (National Diagnostics, UreaGel). Reactive bases were detected with an Amersham Biosciences Typhoon 9400 Variable Mode Imager. Quantification of bands was performed with ImageQuant. For all experiments, individual bands were normalized for incorporation of [α-32P]UTP by dividing the band intensity by the number of Us in the transcript. The percentage readthrough was calculated by dividing the sum of runoff RNAs by the sum of all terminated and runoff products.

***In vitro* transcription (cotranscriptional SHAPE-seq experiment).** 50-μL total reaction mixtures containing 100 nM linear DNA template library (described above) and 4 U of *E. coli* RNAP holoenzyme (New England BioLabs) were incubated in transcription buffer (20 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 1 mM DTT, and 50 mM KCl), 0.2 mg/mL bovine serum albumin, and 500 μM NTPs for 7.5 min at 37 °C to form open complexes. When present, NaF was included at a final concentration of 10 mM. After the first incubation, EcoRI Gln111 dimer was added to a final concentration of 500 nM and incubated at 37 °C for another 7.5 min. Immediately after the second incubation, single-round transcription reactions were initiated with the addition of MgCl₂ to 5 mM and rifampicin to 10 μg/ml. All transcription reactions were allowed to proceed for 30 s. Cotranscriptional experiments were then directly SHAPE-modified (RNA modification and purification described below). Equilibrium refolding experiments were stopped by the addition of 150 μL of TRIzol solution (Life Technologies), purified, and equilibrium-refolded in transcription buffer before SHAPE modification as described below.

***in vitro* transcription (single length, unlabeled).** *in vitro* transcription of single-length unlabeled RNA was performed as described above for cotranscriptional SHAPE-seq, except in a 25-μL total volume with 2 U of *E. coli* RNAP holoenzyme (New England BioLabs) and without Gln111 addition or SHAPE modification. The resulting RNAs were purified as described for cotranscriptional SHAPE-seq and fractionated with a 10% denaturing polyacrylamide gel containing 7.0 M urea. The resulting gel was stained with SYBR Gold (Life Technologies), imaged with a Bio-Rad ChemiDoc MP system and quantified with Image Lab (Bio-Rad). The percentage readthrough was calculated as described above.

RNA modification and purification. For cotranscriptional experiments, the 30-s transcription products were immediately SHAPE-modified by splitting the reaction and mixing half with 2.78 μL of 400 mM benzoyl cyanide (BzCN; Pfaltz & Bower) dissolved in anhydrous dimethyl sulfoxide (DMSO; (+) sample) or anhydrous DMSO only (Sigma Aldrich; (−) sample) for ~2 s before addition of 75 μL of TRIzol solution. DMS modification was performed by splitting the reaction and mixing half with 2.78 μL of 3.5% DMS in ethanol ((+) sample) or 100% ethanol only ((−) sample) and incubating the reactions at 37 °C for 3 min before quenching with 6.67 μL of β-mercaptoethanol, incubation at 37 °C for 1 min and addition of 75 μL of TRIzol solution. Transcription products for equilibrium refolding had 150 μL of TRIzol added after *in vitro* transcription. The products of both reactions were extracted according to the manufacturer's protocol and dissolved in 20 μL total of 1× DNase I buffer (New England BioLabs) containing 1 U of DNase I enzyme. Digestion proceeded at 37 °C for 30 min, after which 30 μL of RNase-free H₂O, then 150 μL of TRIzol, was added. The RNA samples were then extracted again according to the manufacturer's protocol and dissolved in either 10 μL of 10% DMSO in H₂O (cotranscriptional experiments) or 25 μL of RNase-free H₂O (equilibrium refolding experiments). Samples for equilibrium refolding experiments were then heated to 95 °C for 2 min, snap cooled on ice for 1 min, and refolded in 1× folding buffer for 20 min at 37 °C (20 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 1 mM DTT, 50 mM KCl, 0.2 mg/mL bovine serum albumin, and 500 μM NTPs), optionally containing 10 mM fluoride. RNA modification of the equilibrium refolding samples was performed as described above and was followed by the addition of 30 μL of RNase-free H₂O and 150 μL of TRIzol. Samples were extracted a third time according to the manufacturer's instructions. The resulting pellet was dissolved in 10 μL of 10% DMSO in H₂O.

Linker preparation. The phosphorylated linker, oligonucleotide A (**Supplementary Table 4**), was purchased from Integrated DNA Technologies and adenylated with a 5′ DNA Adenylation Kit (New England BioLabs) according to the manufacturer's protocol at a 20× scale, with the reactions divided into

50-µL aliquots. After completion of the reaction, 150 µL of TRIzol was added, and the reactions were extracted according to the manufacturer's instructions, and the products were dissolved in 20 µL of RNase-free H$_2$O. The concentration of purified linker was determined with a Qubit Fluorometer (Life Technologies), and the molarity of the RNA was calculated by using 6782.1 g/mol as the molecular weight. The adenylation reaction was assumed to be 100% efficient. The linker was diluted to a 2 µM stock for subsequent use.

Linker ligation. To the modified and unmodified RNAs in 10% DMSO (RNA modification and purification described above), 0.5 µL of SuperaseIN (Life Technologies), 6 µL of 50% PEG 8000, 2 µL of 10× T4 RNA Ligase Buffer (New England BioLabs), 1 µL of 2 µM 5′-adenylated RNA linker, and 0.5 µL of T4 RNA ligase, truncated KQ (200 U/µL; New England BioLabs) were added to bring the total reaction volume to 20 µL. The reactions were mixed well and incubated overnight (>10 h) at room temperature.

Reverse transcription. The completed linker ligations were adjusted to 150 µL with RNase-free H$_2$O before the addition of 15 µL of 3 M NaOAc, 1 µL of 20 mg/mL glycogen, and 450 µL of EtOH for EtOH precipitation. Precipitated pellets were dissolved in 10 µL of RNase-free H$_2$O. Then, 3 µL of 0.5 µM reverse-transcription primer, oligonucleotide B (**Supplementary Table 4**), was added. The resulting mix was then denatured completely by heating to 95 °C for 2 min, and samples were incubated at 65 °C for 5 min before being placed on ice for ~30 s. Then, 7 µL of SSIII master mix was added, containing 0.5 µL of Superscript III (Life Technologies), 4 µL of 5× First Strand Buffer (Life Technologies), 1 µL of 100 mM DTT, 1 µL of 10 mM dNTPs, and 0.5 µL of RNase-free H$_2$O. The reaction mix was further incubated at 42 °C for 1 min, then at 52 °C for 25 min, and deactivated by heating at 65 °C for 5 min. The RNA was then hydrolyzed by the addition of 1 µL of 4 M NaOH solution and heating at 95 °C for 5 min. The basic solution containing the cDNA was partially neutralized with 2 µL of 1 M HCl and then precipitated with 69 µL of cold EtOH; samples were incubated at 15 min at −80 °C and spun for 15 min at 4 °C at maximal speed to pellet the RNA, and the pellet was washed with 70% EtOH. The washed pellet, free of base, was dissolved in 22.5 µL of nuclease-free H$_2$O.

Adapter ligation. To the cDNA, 3 µL of 10× CircLigase Buffer (Epicentre), 1.5 µL of 50 mM MnCl$_2$, 1.5 µL of 1 mM ATP, 0.5 µL of 100 µM DNA adapter, oligonucleotide C (**Supplementary Table 4**), and 1 µL of CircLigase I (Epicentre) were added. The reaction was incubated at 60 °C for 2 h, then at 80 °C for 10 min to inactivate the ligase. The ligated DNA was EtOH precipitated with 1 µL of 20 mg/mL glycogen as a carrier and dissolved in 20 µL of nuclease-free H$_2$O. Then the cDNA was purified with 36 µL of Agencourt XP beads (Beckman Coulter), according to the manufacturer's instructions and eluted with 20 µL of TE buffer.

Quality analysis. For quality analysis (QA), a separate PCR reaction for each (+) and (−) sample was mixed by combining 13.75 µL of nuclease-free H$_2$O, 5 µL of 5× Phusion Buffer (New England BioLabs), 0.5 µL of 10 mM dNTPs, 1.5 µL of 1 µM labeling primer (oligonucleotides D/E (**Supplementary Table 4**)), 1.5 µL of 1 µM primer PE_F (oligonucleotide F (**Supplementary Table 4**)), 1 µL of 0.1 µM selection primer (oligonucleotides G/H (**Supplementary Table 4**)), 1.5 µL of ssDNA library (+ or −), and 0.25 µL of Phusion DNA polymerase (New England BioLabs). Both fluorescent primers were purchased from Applied Biosystems, and the selection primers were purchased from Integrated DNA Technologies. Asterisks represent phosphorothioate modifications to prevent the 3′→5′ exonuclease activity of Phusion polymerase. Amplification was first

performed for 15 cycles, at an annealing temperature of 65 °C and an extension time of 15 s, excluding the PE_F primer. Then, the PE_F primer was added for an additional ten cycles of amplification. To the complete reactions, 50 µL of nuclease-free H$_2$O was added, and the diluted reaction was ethanol-precipitated. The resulting pellet was dissolved in formamide and analyzed with an ABI 3730xl capillary electrophoresis device.

Library preparation and next-generation sequencing. To construct sequencing libraries, a separate PCR for each (+) and (−) sample was mixed by combining 33.5 µL of nuclease-free H$_2$O, 10 µL of 5× Phusion Buffer (New England BioLabs), 0.5 µL of 10 mM dNTPs, 0.25 µL of 100 µM TruSeq indexing primer (oligonucleotide I (**Supplementary Table 4**)), 0.25 µL of 100 µM primer PE_F, 2 µL of 0.1 µM selection primer (+ or −, as noted above), 3 µL of ssDNA library (+ or −), and 0.5 µL of Phusion DNA polymerase (New England BioLabs). Amplification was performed as indicated above. Completed reactions were chilled at 4 °C for 2 min before the addition of 5 U exonuclease I (New England BioLabs) to remove unextended primer. The reactions were then incubated at 37 °C for 30 min. After incubation, 90 µL of Agencourt XP beads (Beckman Coulter) was added for purification according to the manufacturer's instructions. The complete libraries were eluted with 20 µL of TE buffer and quantified with a Qubit 2.0 Fluorometer (Life Technologies).

To prepare the libraries for sequencing, the average length of each sample was determined according to the results from the quality analysis to calculate the molarity of each (+) or (−) sample separately. Sequencing pools were mixed to be equimolar, such that all of the sequencing libraries were present in the solution at the same level. Sequencing was performed on an Illumina HiSeq 2500 in either 'rapid run' or 'high output' mode, with 2 × 36 bp paired-end reads. To help overcome the low complexity of the linker region during sequencing, 10–20% PhiX DNA was included.

Data analysis. A detailed description of the cotranscriptional SHAPE-seq data analysis pipeline is provided in the **Supplementary Note**.

Code availability. Spats v1.0.1 can be accessed at https://github.com/LucksLab/spats/releases/. Scripts used in data processing are located at https://github.com/LucksLab/Cotrans_SHAPE-Seq_Tools/releases/.

Data availability. Raw sequencing data that support the findings of this study have been deposited in the Sequence Read Archive under BioProject PRJNA342175. Individual BioSample accession codes are available in **Supplementary Table 1**. SHAPE-seq reactivity spectra generated in this work have been deposited in the RNA Mapping Database under accession codes FLUORSW_BZCN_0001, FLUORSW_BZCN_0002, FLUORSW_BZCN_0003, FLUORSW_BZCN_0004, FLUORSW_BZCN_0005, FLUORSW_BZCN_0006, FLUORSW_BZCN_0007, FLUORSW_BZCN_0008, FLUORSW_BZCN_0009, FLUORSW_BZCN_0010, FLUORSW_BZCN_0011, FLUORSW_BZCN_0012, FLUORSW_BZCN_0013, FLUORSW_BZCN_0014, FLUORSW_BZCN_0015, FLUORSW_BZCN_0016, FLUORSW_BZCN_0017, FLUORSW_BZCN_0018, FLUORSW_BZCN_0019, FLUORSW_BZCN_0020, SRPECLI_BZCN_0001, SRPECLI_BZCN_0002, SRPECLI_BZCN_0003, SRPECLI_BZCN_0004, SRPECLI_DMS_0001, SRPECLI_DMS_0002, and SRPECLI_DMS_0003. Sample details are available in **Supplementary Table 2**. Source data for **Figures 2–5** and **Supplementary Figures 1**, **2**, and **4–9** are available with the paper online. All other data that support the findings of this paper are available from the corresponding author upon reasonable request.