

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Structural Acrobatics of Manganese-Sensing RNAs During Transcription

Permalink

<https://escholarship.org/uc/item/0m73s52n>

Author

Stephen, Christine

Publication Date

2025

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Structural Acrobatics of Manganese-Sensing RNAs During Transcription

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Biochemistry and Molecular Biophysics

by

Christine Nicole Stephen

Committee in charge:

Professor Tatiana Mishanina, Chair
Professor Kevin Corbett
Professor Thomas Hermann
Professor Mark Herzik
Professor Dong Wang

2025

Copyright

Christine Nicole Stephen, 2025

All rights reserved.

The Dissertation of Christine Nicole Stephen is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2025

TABLE OF CONTENTS

DISSERTATION APPROVAL PAGE	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	vii
LIST OF SUPPLEMENTAL FIGURES	viii
LIST OF TABLES	xi
LIST OF SUPPLEMENTAL TABLES	xi
LIST OF ABBREVIATIONS	xii
ACKNOWLEDGEMENTS	xv
ABSTRACT OF THE DISSERTATION	xix
Chapter 1 Introduction to Riboswitches	21
1. Synopsis.....	21
2. Riboswitches.....	21
3. Riboswitch co-transcriptional folding.....	23
4. References.....	25
5. Acknowledgements.....	25
Chapter 2 Tools to study riboswitches	27
1. Reconstituting transcription in the test tube.....	27
2. Introduction to RNA chemical probing.....	28
3. Co-transcriptional RNA chemical probing.....	29
4. Gene reporter fusions.....	30
5. References.....	31
6. Acknowledgements.....	32

Chapter 3 Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the *Escherichia coli* pH-responsive riboswitch 33

0. Abstract	36
1. Introduction.....	37
2. Results	41
3. Discussion	57
4. Experimental Procedures	63
4. Figures	82
5. References	89
6. Supporting Information for: Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the <i>Escherichia Coli</i> pH-responsive riboswitch	93
7. Acknowledgements	123

Chapter 4 Structurally distinct manganese-sensing riboswitch aptamers regulate diverse expression platform architectures 124

0. Abstract	125
1. Introduction.....	126
2. Materials and Methods	130
3. Results	142
4. Discussion	167
5. References	172
6. Supplementary Information for: Structurally distinct manganese-sensing riboswitch aptamers regulate diverse expression platform architectures.....	177
7. Acknowledgements	253

Chapter 5 Opportunities for riboswitch inhibition by targeting co-transcriptional RNA folding events.....	254
0. Abstract	256
1. Introduction.....	256
2. Limitations in Designing Synthetic Ligands That Target Riboswitches in Pathogenic Bacteria	260
3. RNA-based Silencing in Natural Contexts.....	264
4. Synthetic RNA-Based Silencing	265
5. Advances in Riboswitch Co-transcriptional Folding for Riboswitch Families Found in Pathogenic Bacteria	267
6. Antisense Oligonucleotides That Act on RNA Co-Transcriptionally: A Case Study on a Group I Intron	285
7. Delivery of ASOs	287
8. Translation of ASO Therapies to the Clinic	288
9. Discussion	289
10. References	292
11. Acknowledgements	304

LIST OF FIGURES

Figure 1.1 Overview of riboswitch-controlled gene expression.	22
Figure 1.2 RNA forms its biologically relevant folds during transcription.	23
Figure 1.3 Riboswitches sense ligand and rearrange their expression platform structure co-transcriptionally.	24
Figure 2.1 Reconstitution of promoter-initiated transcription in the test tube.	27
Figure 2.2 Overview of <i>in vitro</i> pausing assays with RNA polymerase (RNAP) assembled on a short, synthetic nucleic acid scaffold.	28
Figure 2.3 Chemical probing to map RNA secondary structure.	29
Figure 2.4 Design of reporter fusions used in this work.	30
Figure 3.1 Alternative folding outcomes of the pH-responsive riboswitch.	82
Figure 3.2 RNAP transcription through the PRE at neutral vs. alkaline pH.	83
Figure 3.3 Locating the first critical pauses for PRE folding.	84
Figure 3.4 Measurements of elemental pause escape rate at the sequential pauses under neutral vs. alkaline conditions.	85
Figure 3.5 Measurement of pause kinetics at the third, backtracked sequential pause during active transcription elongation in presence and absence of Gre factors.	86
Figure 3.6 Measurement of pause kinetics at Pause 2 during active transcription elongation.	88
Figure 4.1 Introduction to the <i>E. coli alx</i> and <i>mntP</i> Mn-sensing translational riboswitches.	129
Figure 4.2 Aptamer architectures and riboswitch-controlled translational response <i>in vivo</i>	144
Figure 4.3 Co-transcriptional DMS probing of the <i>mntP</i> riboswitch aptamer at alkaline pH.	150
Figure 4.4 Co-transcriptional DMS probing of the <i>alx</i> riboswitch aptamer at alkaline pH.	153
Figure 4.5 Decoupling the effects of alkaline pH and Mn on <i>alx</i> riboswitching <i>in vivo</i> and pH-dependent reactivity changes in the divergent <i>alx</i> and <i>mntP</i> L3 sequences.	157
Figure 4.6 Co-transcriptional DMS probing of the <i>mntP</i> riboswitch during expression platform folding.	160
Figure 4.7 Co-transcriptional DMS probing of the <i>alx</i> riboswitch during expression platform folding.	164

Figure 4.8 Changes in the <i>alx</i> aptamer during expression platform folding.	166
Figure 4.9 Model for <i>alx</i> and <i>mntP</i> riboregulation in <i>E. coli</i>	170
Figure 5.1 Overview of riboswitch-controlled gene expression and riboswitch inhibition strategies.	258
Figure 5.2 Overview of ASO-directed co-transcriptional silencing of target riboswitch-controlled genes, using the TPP riboswitch as an example.	266
Figure 5.3 Folding pathway for <i>E. coli thiC</i> riboswitch in the presence and absence of TPP ligand and its proposed targeting with ASO.	274
Figure 5.4 <i>E. coli tbpA</i> riboswitch folding intermediates experimentally validated by co-transcriptional SHAPE-seq.	277
Figure 5.5 Folding outcomes for <i>B. subtilis yxjA</i> riboswitch in the presence and absence of guanine ligand and its proposed targeting with ASO.	281
Figure 5.6 ASO targeting of <i>C. albicans</i> 26S rRNA group I intron	286

LIST OF SUPPLEMENTAL FIGURES

Supplemental Figure 3.1 Secondary structure models of the <i>E. coli</i> pH-responsive riboswitch at pH 7.2 and 8.5.	95
Supplemental Figure 3.2 Quantitation of synchronized single-round <i>in vitro</i> transcription from the <i>alx</i> promoter and transcription in the presence of pause-modulating factors.	97
Supplemental Figure 3.3 Pause 1 mapping at limiting NTP concentrations and NET-seq data analysis.	100
Supplemental Figure 3.4 Sequential pause escape measurements with ECs positioned directly at each pause and Gre-catalyzed RNA cleavage at the third pause.	101
Supplemental Figure 3.5 Analysis of Pause 1 kinetics during active transcription elongation.	103
Supplemental Figure 3.6 Pause 2 mapping at limiting GTP concentration and NET-seq data analysis.	105
Supplemental Figure 3.7 Analysis of potential RNA structures in the vicinity of Pause 1 and test of <i>his</i> pause escape kinetics at different pH.	106
Supplemental Figure 3.8 Analysis of previously reported mutations demonstrated to affect Pause 1 longevity or location.	107
Supplemental Figure 4.1 Structure probing of equilibrium-refolded <i>alx</i> aptamer RNA.	178
Supplemental Figure 4.2 TECprobe-VL construct designs for the <i>alx</i> and <i>mntP</i> riboswitches.	180

Supplemental Figure 4.3 Co-transcriptional DMS probing of the <i>mntP</i> riboswitch aptamer at alkaline pH +Mn.	181
Supplemental Figure 4.4 Synchronized, single-round <i>in vitro</i> transcription initiated from the <i>mntP</i> promoter and RNA structure models for pause-adjacent aptamer folding intermediates.	182
Supplemental Figure 4.5 Mapping <i>mntP</i> pause locations during aptamer and early expression platform folding.	183
Supplemental Figure 4.6 Equilibrium DMS probing of <i>mntP</i> aptamer folding intermediates at alkaline pH $\pm$ Mn.	184
Supplemental Figure 4.7 Co-transcriptional DMS probing of the <i>alx</i> riboswitch aptamer at alkaline pH +Mn.	186
Supplemental Figure 4.8 Equilibrium DMS probing of <i>alx</i> aptamer folding intermediates at alkaline pH $\pm$ Mn.	187
Supplemental Figure 4.9 Co-transcriptional DMS probing of the <i>alx</i> and <i>mntP</i> aptamers at neutral pH $\pm$ Mn and pH-dependent reactivity changes in L3.	188
Supplemental Figure 4.10 Equilibrium DMS probing of <i>alx</i> and <i>mntP</i> aptamer folding intermediates at neutral pH $\pm$ Mn and pH-dependent reactivity changes in L3.	190
Supplemental Figure 4.11 pH-dependent reactivity changes in <i>alx</i> and <i>mntP</i> L1 in co-transcriptionally folded and equilibrium refolded aptamer intermediates.	192
Supplemental Figure 4.12 Co-transcriptional DMS probing of <i>mntP</i> riboswitch during expression platform synthesis at neutral and alkaline pH.	193
Supplemental Figure 4.13 Equilibrium DMS probing on <i>mntP</i> folding intermediates during expression platform synthesis.	195
Supplemental Figure 4.14 Structure probing of equilibrium-refolded <i>mntP</i> full-length riboswitch RNA.	197
Supplemental Figure 4.15 Co-transcriptional DMS probing of the <i>alx</i> riboswitch during expression platform synthesis at alkaline pH.	199
Supplemental Figure 4.16 Co-transcriptional DMS probing of the <i>alx</i> riboswitch during expression platform synthesis at neutral pH and pH-dependent reactivity differences in the metal-sensing core.	201
Supplemental Figure 4.17 Equilibrium DMS probing of the <i>alx</i> riboswitch during expression platform synthesis at neutral vs. alkaline pH and Mn-dependent reactivity differences in the expression platform.	202
Supplemental Figure 4.18 DMS reactivities of <i>mntP</i> RBS nt in co-transcriptionally folded and equilibrium refolded intermediates.	203

Supplemental Figure 4.19 Dropoff in sequencing depth following second <i>alx</i> intrinsic termination site and co-transcriptional SHAPE probing of full-length <i>alx</i> to capture RBS reactivities.....	205
Supplemental Figure 4.20 Structure probing of equilibrium-refolded <i>alx</i> full-length riboswitch RNA.....	206
Supplemental Figure 4.21 Aligned reads and alignment rates for co-transcriptional DMS probing of the <i>alx</i> riboswitch.....	207
Supplemental Figure 4.22 Reactivity matrices for individual and concatenated replicates for co-transcriptional DMS probing of the <i>alx</i> riboswitch at pH 7.2....	208
Supplemental Figure 4.23 Reactivity matrices for individual and concatenated replicates for co-transcriptional DMS probing of the <i>alx</i> riboswitch at pH 7.2 +Mn.	209
Supplemental Figure 4.24 Reactivity matrices for individual and concatenated replicates for co-transcriptional DMS probing of the <i>alx</i> riboswitch at pH 8.5....	210
Supplemental Figure 4.25 Reactivity matrices for individual and concatenated replicates for co-transcriptional DMS probing of the <i>alx</i> riboswitch at pH 8.5 +Mn.	211
Supplemental Figure 4.26 Aligned reads and alignment rates for DMS probing of <i>alx</i> equilibrium refolded intermediates.....	212
Supplemental Figure 4.27 Reactivity matrices for individual and concatenated replicates for equilibrium DMS probing of the <i>alx</i> riboswitch at pH 7.2.....	213
Supplemental Figure 4.28 Reactivity matrices for individual and concatenated replicates for equilibrium DMS probing of the <i>alx</i> riboswitch at pH 7.2 +Mn.	214
Supplemental Figure 4.29 Reactivity matrices for individual and concatenated replicates for equilibrium DMS probing of the <i>alx</i> riboswitch at pH 8.5.....	215
Supplemental Figure 4.30 Reactivity matrices for individual and concatenated replicates for equilibrium DMS probing of the <i>alx</i> riboswitch at pH 8.5 +Mn.	216
Supplemental Figure 4.31 Aligned reads and alignment rates for co-transcriptional DMS probing of the <i>mntP</i> riboswitch.....	217
Supplemental Figure 4.32 Aligned reads and alignment rates for equilibrium DMS probing of the <i>mntP</i> riboswitch.....	218

LIST OF TABLES

Table 5.1 Riboswitch classes in pathogenic bacteria whose co-transcriptional folding has been experimentally characterized.....	270
--	------------

LIST OF SUPPLEMENTAL TABLES

Supplemental Table 3.1 PRE nucleotide sequence, plasmids and oligonucleotides used in this study.	108
Supplemental Table 3.2 Summary of the sequential pause kinetic parameters from scaffold-based experiments positioning ECs at each pause (Fig. 3.4).	111
Supplemental Table 3.3 Summary of the sequential pause kinetic parameters from scaffold-based experiments during active transcription elongation (Fig. 3.5)....	112
Supplemental Table 3.4 Summary of kinetic parameters for pauses prior to termination site D from scaffold-based experiments during active transcription elongation (Fig. 3.6).....	113
Supplemental Table 4.1 Key reagents used in this study.....	240
Supplemental Table 4.2 Oligonucleotides used in this study.....	242
Supplemental Table 4.3 DNA templates used in this study.....	244
Supplemental Table 4.4 <i>E. coli</i> K-12 strains used in this study.....	245
Supplemental Table 4.5 Plasmids used in this study.	245
Supplemental Table 4.6 Sequences of DNA templates used in this study.	249
Supplemental Table 4.7 Biosample information for all co-transcriptional chemical probing experiments performed in this work.	250
Supplemental Table 4.8 Biosample information for all single-length, equilibrium refolded chemical probing experiments performed in this work.....	251

LIST OF ABBREVIATIONS

ASO	antisense oligonucleotide
BzCN	benzoyl cyanide
Cadmium	Cd
c-di-GMP	cyclic di-guanosine monophosphate
cDNA	complementary DNA
cryo-EM	cryogenic electron microscopy
DMS	dimethyl sulfate
DNA	deoxyribonucleic acid
dNMP	deoxynucleotide monophosphate
dNTP	deoxynucleotide triphosphate
dsRNA	double-stranded RNA
EC	elongation complex
ePEC	elemental paused elongation complex
FMN	flavin mononucleotide
Glc6NP	glucosamine-6-phosphate
IC ₅₀	half-maximal inhibitory concentration
K _D	dissociation constant
kDa	kilodalton
MaP	mutational profiling
MD	molecular dynamics
Mg	magnesium

Mn	manganese
mRNA	messenger RNA
NMR	nuclear magnetic resonance
nt	nucleotide
NT-DNA	non-template DNA
NTP	nucleotide triphosphate
ORF	open reading frame
PCR	polymerase chain reaction
PEC	paused elongation complex
PRE	pH-responsive element
pre-miRNA	pre-microRNA
preQ ₁	prequeuosine 1
RBS	ribosome binding site or ribosome binding sequence
RNA	ribonucleic acid
RNAi	RNA interference
RNAP	RNA polymerase
rRNA	ribosomal RNA
RT	reverse transcription
RT-PCR	reverse transcription polymerase chain reaction
SAM	S-adenosyl methionine
SHAPE	selective 2'-hydroxyl acylation analyzed by primer extension
smFRET	single-molecule fluorescence resonance energy transfer

smTIRF	single-molecule total internal reflection fluorescence
sRNA	small RNA
ssRNA	single-stranded RNA
T-DNA	template DNA
TECprobe-VL	<u>v</u> ariable <u>l</u> ength <u>t</u> ranscription <u>e</u> longation <u>c</u> omplex RNA chemical <u>p</u> ro <u>b</u> ing
TPP	thiamine pyrophosphate
UTR	untranslated region
ZTP	5'-aminoimidazole-4-carboxamide-1-beta-D-ribofuranosyl-5'-triphosphate
2D	two-dimensional
3D	three-dimensional
3WJ	three-way junction
4WJ	four-way junction

ACKNOWLEDGEMENTS

I would first and foremost like to thank my close friends and family members, whose support throughout my PhD has been invaluable. First, I thank my parents, Janis and James Stephen, who have always supported my ambitions and told me that I can accomplish whatever I set my mind to. Their humble, hardworking nature is truly admirable and are qualities I strive to emulate every day. Next, I express my sincerest gratitude for my best friends – Jennifer Guan, Brina Patel, Nimrita Singh, Yesenia Garcia, Olivia Schmidt, and Bastiaan van Ravenhorst. To each of you, your friendship has enriched my life in more ways than I can count. Our numerous phone calls and meetups throughout my PhD were filled with joy and laughter, which nourished my spirit and kept me going when I was stressed about my lab projects. I also thank my current roommate and friend Michaela Horger. Michaela, it has been a blast living with you this final year of my PhD and your resilience, creativity, and fun-loving nature has been a source of inspiration to me.

I would next like to thank numerous members of the Mishanina lab. First, I thank my thesis advisor, Dr. Tatiana Mishanina for her incredible mentorship these past six years. I thank Tatiana for her unwavering support and belief in me over the course of my thesis work. Tatiana, working with you on this project is the greatest accomplishment of my life so far, and the enthusiasm, love, and creativity you bring to work every day continually inspires me. Some of the best memories of my PhD are from our weekly meetings, where we discussed ideas and new research directions and you helped nurture and build my confidence over time.

Next, I would like to thank Sean Reardon, my best friend in the Mishanina lab. Sean, you are truly one of a kind, and I can't image having a lab mate more fun than you. I am so grateful my office seat and lab bench were right next to you, because this is where we forged a friendship over our many common interests, including meditation, books, music, and of course, science.

Your support over the course of my PhD means a lot to me and has helped to motivate me when I was stressed, which was often. I would also like to thank Danae Palmer, who has been an incredible partner in crime on this project. Danae, I can't wait to see where you take this project in the future. Additionally, I thank Ravish Sharma, who has been an incredible asset to the riboswitch projects in the lab. Ravish, your calm, kind, and thoughtful nature are inspiring, and it was a pleasure sitting next to you these past few years. Lastly, I thank all my other lab mates in the Mishanina lab. You are all remarkable individuals who have together created a positive, and warm environment that benefits all of our trainees.

Chapter 1, in part, is a reprint of the material as it appears in Opportunities for riboswitch inhibition by targeting co-transcriptional RNA folding events. Stephen, Christine; Palmer, Danae; Mishanina, Tatiana, International Journal of Molecular Sciences, 2024. The dissertation author was the primary co-author of this paper.

Chapter 2, in part, is a reprint of the material as it appears in Opportunities for riboswitch inhibition by targeting co-transcriptional RNA folding events. Stephen, Christine; Palmer, Danae; Mishanina, Tatiana, International Journal of Molecular Sciences, 2024. The dissertation author was the primary co-author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the *E. coli* pH-responsive riboswitch. Stephen, Christine; Mishanina, Tatiana, Journal of Biological Chemistry, 2022. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is a reprint of the material in peer review as it appears in Structurally distinct manganese-sensing riboswitch aptamers regulate diverse expression platform architectures. Stephen, Christine; Palmer, Danae; Bautista, Clarisa; Mishanina, Tatiana, in peer review at Nucleic Acids Research, 2025. The dissertation author was the primary investigator and author of this material.

Chapter 5, in full, is a reprint of the material as it appears in Opportunities for riboswitch inhibition by targeting co-transcriptional RNA folding events. Stephen, Christine; Palmer, Danea; Mishanina, Tatiana, International Journal of Molecular Sciences, 2024. The dissertation author was the primary co-author of this paper.

VITA

- 2018 Bachelor Science in Biochemistry, University of California Santa Barbara
- 2020 Master of Science in Chemistry, University of California San Diego
- 2025 Doctor of Philosophy in Biochemistry and Molecular Biophysics, University of California San Diego

PUBLICATIONS

Stephen, C., Palmer, D., Mishanina, T.V. Structurally distinct manganese-sensing riboswitch aptamers regulate diverse expression platform architectures. *In review at Nucleic Acids Res.* (2025).

Stephen, C., Palmer, D., Mishanina, T.V. Opportunities for riboswitch inhibition by targeting co-transcriptional RNA folding events. *International Journal of Molecular Sciences*. **25**, 10495 (2024).

Tingstedt, J.L., **Stephen, C.**, Risinger, C., Blixt, O., Gunalan, V., Johansen, I.S., Fomsgaard, A., Polacek, C., Lassaunière, R. Differential recognition of influenza A virus H1N1 neuraminidase by DNA vaccine-induced antibodies in pigs and ferrets. *Frontiers in Immunology*. **14** (2023).

Stephen, C. and Mishanina, T.V. Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the *Escherichia coli* pH-responsive riboswitch. *Journal of Biological Chemistry*. **298**, 102302 (2022).

ABSTRACT OF THE DISSERTATION

Structural Acrobatics of Manganese-Sensing RNAs During Transcription

by

Christine Nicole Stephen

Doctor of Philosophy in Biochemistry and Molecular Biophysics

University of California San Diego, 2025

Professor Tatiana Mishanina, Chair

Riboswitches are 5'-untranslated regions of a messenger RNA that control the fate of their respective transcript by changing their conformation in response to a ligand-binding event, effectively turning the gene “on” or “off”. This dissertation investigates two riboswitches in *E. coli* that sense manganese (Mn), which upon Mn binding, rearrange their folded structures into a

conformation that is accessible for ribosome binding and subsequent translation of the downstream genes, *alx* and *mntP*. While prior study uncovered the structure and dynamics of Mn-sensing riboswitch aptamers, it remains unknown how Mn binding rearranges the full-length riboswitch into a translationally active conformation. Further, prior work missed the aspect of active transcription that is critical for riboswitch folding and ligand-sensing. To address these gaps in knowledge, we employed a suite of biochemical techniques to elucidate how the *E. coli* Mn-sensing riboswitches, *alx* and *mntP*, fold and respond to Mn binding co-transcriptionally. First, we have applied *in vitro* transcription methods to study the kinetics of RNA polymerase (RNAP) transcription on the *alx* and *mntP* riboswitch sequences. Specifically, we characterized transcriptional pauses encountered by RNAP during synthesis of the *alx* and *mntP* riboswitches, which may coordinate key RNA folding transitions. Next, we explored how the *alx* and *mntP* RNAs fold co-transcriptionally using RNA chemical probing methods. These data uncovered that the *alx* and *mntP* riboswitch aptamer folding intermediates can sense Mn, challenging the notion that a fully folded riboswitch aptamer is required for ligand sensing. Further, we provide the first insights into how the conserved *alx* and *mntP* aptamer domains “communicate” with their expression platforms, which diverge in sequence and in structure. Lastly, we show that specific differences in *alx* and *mntP* aptamer sequence may explain how pH differentially tunes Mn affinity for the riboswitch aptamers. To underscore the importance of riboswitch co-transcriptional folding, we highlight how RNA folding intermediates may be targeted by antisense oligonucleotides (ASOs) – an exciting new strategy for riboswitch inhibitor design.

Chapter 1 Introduction to Riboswitches

1. Synopsis

This dissertation investigates how ligand-sensing RNAs called riboswitches fold and bind to ligand during active transcription by RNA polymerase (RNAP). Chapter 1 gives the reader background on riboswitches and how they fold co-transcriptionally and builds the foundation for the studies of the *E. coli* manganese (Mn)-sensing riboswitches, *alx* and *mntP*, detailed in Chapters 3-4.

Chapter 2 is an overview of the experimental techniques used to study riboswitches both in the test tube (e.g., *in vitro* transcription assays and chemical probing of RNA structure) as well as in cells (e.g., gene reporter fusions). Further, Chapter 2 summarizes the state-of-the-art RNA chemical probing techniques that were used to map the co-transcriptional RNA folding transitions detailed in Chapter 4.

2. Riboswitches

Bacteria employ non-coding RNA elements called riboswitches to attune their gene expression programs to shifting nutrient availability¹. Riboswitches are structured, 5'-untranslated regions of a messenger RNA (mRNA) that change their structure upon binding to a nutrient ligand, which communicates the fate of the mRNA to the transcription or translation machinery². Generally, riboswitches are found in bacteria and archaea, with the exception being the thiamine pyrophosphate-sensing riboswitch family that has eukaryotic representatives¹. Riboswitches are composed of two domains: an aptamer domain that binds a ligand, typically with high affinity, and an expression platform that undergoes structural changes to affect gene expression² (Fig. 1.1). There are at least 55 classes of characterized riboswitches that are grouped by what ligand they bind, typically with a highly conserved aptamer domain². Expression platforms are designated as “off” or “on” depending on whether ligand binding rearranges the RNA structure to downregulate

(“off”) or upregulate (“on”) gene expression levels. “Off” riboswitch expression platforms form either an intrinsic terminator hairpin that prompts transcription termination (Fig. 1.1A) or a hairpin stem that sequesters the ribosome binding site, leading to inhibition of translation (Fig. 1.1B). Conversely, for “on” riboswitch expression platforms, ligand binding leads to the formation of an anti-terminator stem or release of the ribosome binding site, promoting downstream gene expression.

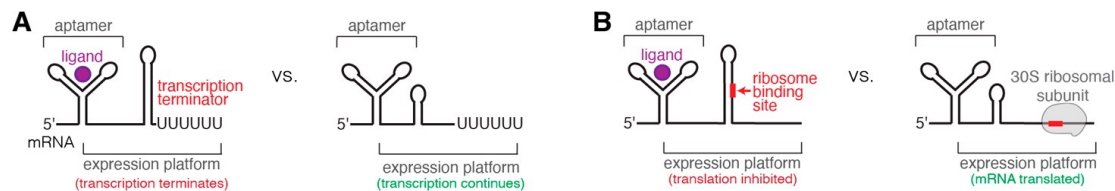


Figure 1.1 Overview of riboswitch-controlled gene expression.

A. Riboswitch responding to ligand by terminating transcription, turning gene expression “off”. **B.** Riboswitch responding to ligand by inhibiting translation, turning gene expression “off”.

3. Riboswitch co-transcriptional folding

RNA folding is a co-transcriptional process by nature, with ligand-binding and folding of RNA secondary structure elements occurring on a faster time scale than transcription³ (Fig. 1.2). Experimental evidence pointing towards the *in vivo* relevance of co-transcriptional RNA folding emerged from early studies on the *Tetrahymena* group I intron, where it was observed that the autocatalyzed splicing of this intron occurs in a few seconds⁴ and thus must take place within the timeframe of transcription. In stark contrast, when the same *Tetrahymena* group I intron RNA is synthesized *in vitro*, denatured and refolded, the majority of the RNA becomes “trapped” in intermediate folds that rearrange into the native state on a slow timescale (~1-100 min)⁵. Therefore, developing a model of how an RNA acquires its functionally relevant folds requires incorporating the transcription machinery into experimental methods.

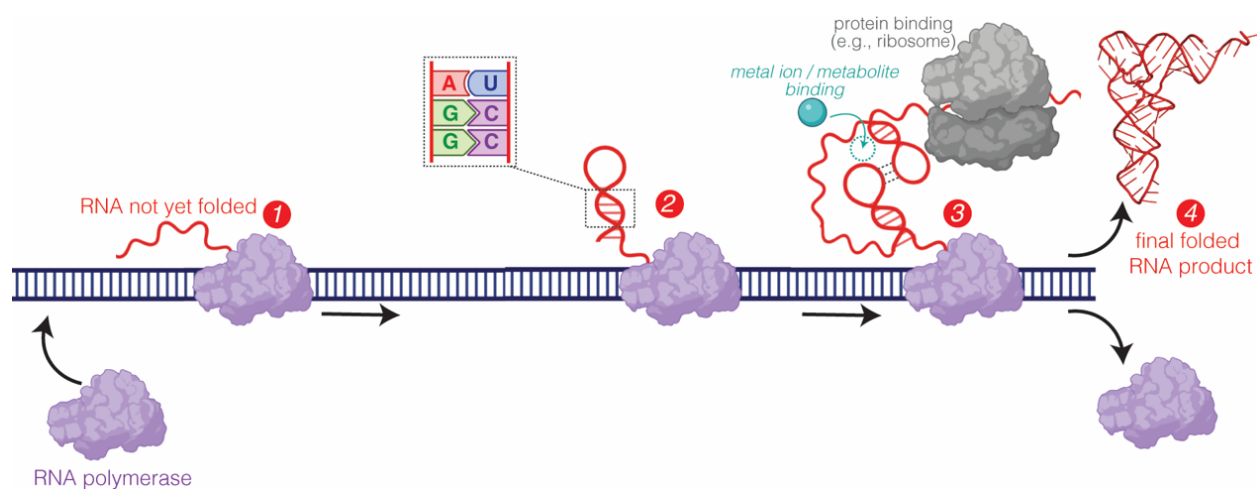


Figure 1.2 RNA forms its biologically relevant folds during transcription.

Early on in transcription by RNA polymerase (RNAP), an RNA molecule is predominantly single-stranded (1). As more of the RNA sequence emerges from RNAP, the RNA can form base-pairs and nucleate secondary structural motifs such as hairpin stems (2). Further along in the RNA synthesis, RNA can nucleate higher-order tertiary interactions required for interaction with binding partners such as a ligand, or the translational machinery (3). Lastly, once transcription terminates, the fully folded RNA is released from RNAP (4).

It is now well-established that many riboswitches rearrange their structure during active transcription^{6–8} (Fig. 1.3). Further, many riboswitch RNAs traverse through intermediate folds that are critical for execution of their gene regulatory decisions⁹, underscoring the complex nature of riboswitch-based gene regulation. The attention of the riboswitch biology community is shifting away from isolated aptamer-ligand interactions and towards the interplay between RNA folding dynamics during transcription, ligand affinity, and resulting gene expression outcomes^{9–12}. The specific riboswitches studied in this dissertation are members of the manganese (Mn)-sensing riboswitch family, which is the largest metal-sensing riboswitch family characterized to date². The specific Mn-sensing riboswitches investigated in this work precede the *alx* and *mntP* genes in *E. coli*. Additional background on the *alx* riboswitch is presented in Chapters 3.1 and 4.1 and additional background on the *mntP* riboswitch is presented in Chapter 4.1 of this dissertation.

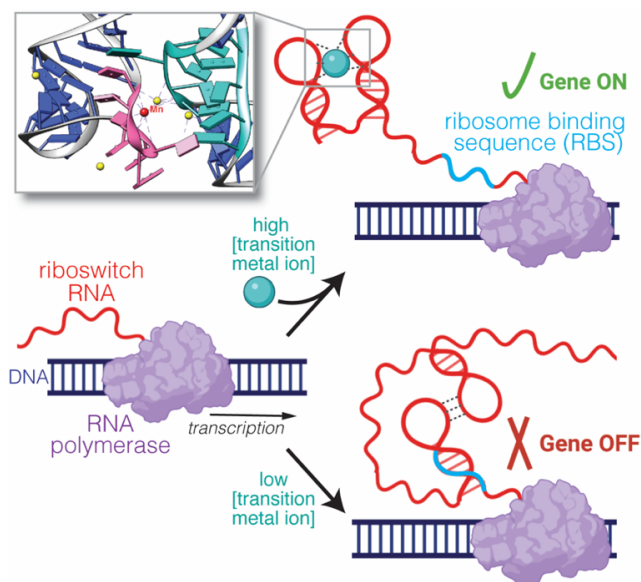


Figure 1.3 Riboswitches sense ligand and rearrange their expression platform structure co-transcriptionally.

Scheme highlighting co-transcriptional binding of a transition metal ion by a riboswitch, ultimately controlling whether the ribosome binding sequence (RBS) is available for translation of the downstream mRNA (messenger RNA).

4. References

1. Serganov, A. & Nudler, E. A Decade of Riboswitches. *Cell* **152**, 17–24 (2013).
2. Kavita, K. & Breaker, R. R. Discovering riboswitches: the past and the future. *Trends Biochem. Sci.* **48**, 119–141 (2023).
3. Al-Hashimi, H. M. & Walter, N. G. RNA dynamics: it is about time. *Curr. Opin. Struct. Biol.* **18**, 321–329 (2008).
4. Brehm, S. L. & Cech, T. R. The fate of an intervening sequence RNA: Excision and cyclization of the *Tetrahymena* ribosomal RNA intervening sequence *in vivo*. *Biochemistry* **22**, 2390–2397 (1983).
5. Woodson, S. A. Folding mechanisms of group I ribozymes: role of stability and contact order. *Biochem. Soc. Trans.* **30**, 1166–1169 (2002).
6. 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).
7. Chauvier, A., Picard-Jean, F., Berger-Dancuse, J.-C., Bastet, L., Naghdi, M.R., Dube, A., Turcotte, P., Perreault, J. & Lafontaine, D.A. Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation. *Nat. Commun.* **8**, 13892 (2017).
8. Strobel, E. J., Cheng, L., Berman, K. E., Carlson, P. D. & Lucks, J. B. A ligand-gated strand displacement mechanism for ZTP riboswitch transcription control. *Nat. Chem. Biol.* **15**, 1067–1076 (2019).
9. Szyjka, C. E. & Strobel, E. J. Observation of coordinated RNA folding events by systematic cotranscriptional RNA structure probing. *Nat. Commun.* **14**, 7839 (2023).
10. Bushhouse, D. Z. & Lucks, J. B. Tuning strand displacement kinetics enables programmable ZTP riboswitch dynamic range *in vivo*. *Nucleic Acids Res.* **51**, 2891–2903 (2023).
11. Bushhouse, D. Z., Fu, J. & Lucks, J. B. RNA folding kinetics control riboswitch sensitivity *in vivo*. *bioRxiv*. doi:10.1101/2024.03.29.587317 (2024).
12. Lennon, S. R. & Batey, R. T. Regulation of Gene Expression Through Effector-dependent Conformational Switching by Cobalamin Riboswitches. *J. Mol. Biol.* **434**, 167585 (2022).

5. Acknowledgements

Chapter 1, in part, is a reprint of the material as it appears in Opportunities for riboswitch inhibition by targeting co-transcriptional RNA folding events. Stephen, Christine; Palmer, Danea;

Mishanina, Tatiana, International Journal of Molecular Sciences, 2024. The dissertation author was the primary co-author of this paper.

Chapter 2 Tools to study riboswitches

1. Reconstituting transcription in the test tube

Since riboswitches fold co-transcriptionally, a main experimental goal of this work was to study transcription by *E. coli* RNA polymerase (RNAP) on the *alx* and *mntP* riboswitch sequences. To achieve this goal, transcription was reconstituted in the test tube by combining the purified components required for transcription: a DNA template, purified *E. coli* RNAP, nucleotide triphosphates (NTPs), and one nucleotide with a radiolabel that is incorporated into the RNA by RNAP (Fig. 2.1). Transcription reactions are quenched at varying time points, and the radiolabeled RNA is visualized by gel analysis to assess the distribution of RNA lengths over time (Fig. 2.1).

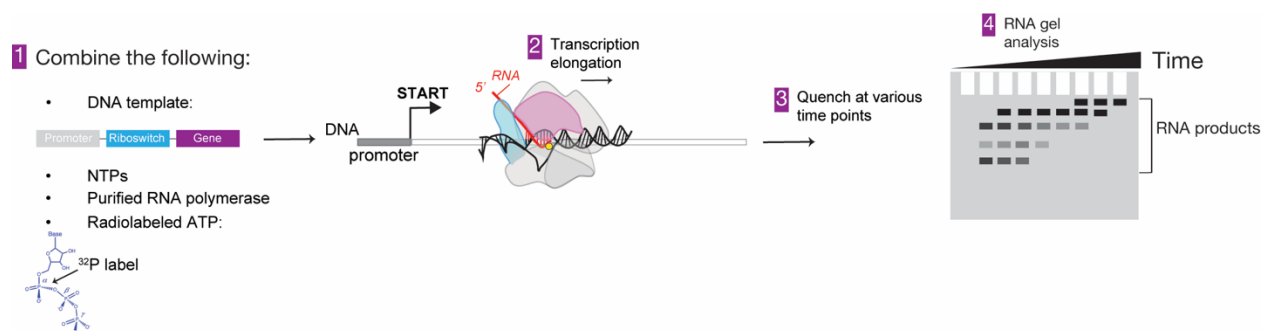


Figure 2.1 Reconstitution of promoter-initiated transcription in the test tube.

A key feature of transcription by RNAPs is that RNA extension is interrupted by pauses – a transient halt in nucleotide incorporation during transcription elongation¹. Pausing during transcription serves key gene regulatory functions across all domains of life¹ and is established to provide a temporal window for riboswitch structural rearrangements^{2–4}. Motivated by the importance of transcriptional pausing to riboswitch folding, this work provides a detailed experimental analysis of transcriptional pausing on the *alx* and *mntP* riboswitch sequences. A useful experiment to study the rate of RNAP escape from a pause is to employ nucleic acid

scaffolds, wherein RNAP is positioned on a short, synthetic RNA-DNA scaffold that recreates the transcription bubble, eliminating the requirement for initiation from a promoter (Fig. 2.2). Once RNAP is positioned onto the scaffold, transcription is initiated and nucleotide incorporation is assayed as a function of time.

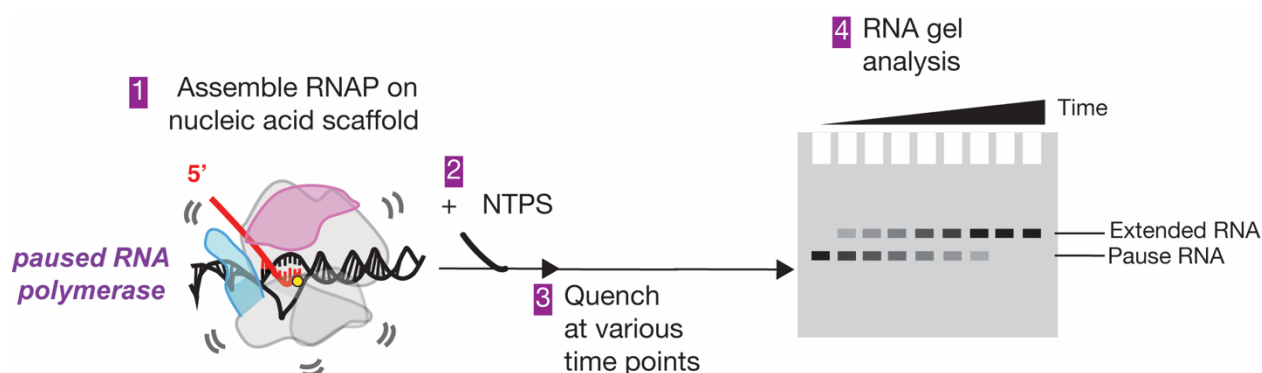


Figure 2.2 Overview of *in vitro* pausing assays with RNA polymerase (RNAP) assembled on a short, synthetic nucleic acid scaffold.

2. Introduction to RNA chemical probing

RNA chemical probing is a widespread technique used to develop secondary structure models of an RNA. The first step of an RNA chemical probing experiment is covalent and irreversible modification of RNA by a reagent that preferentially modifies flexible nucleotides that are not base-paired in the RNA structure⁵. Commonly used chemical probing reagents are either SHAPE (Selective 2'-Hydroxyl Acylation Analyzed by Primers Extension)⁶ reagents or dimethyl sulfate (DMS)⁷. Following modification, the RNA is reverse-transcribed into complementary DNA (cDNA) under mutational profiling (MaP) conditions, resulting in mutations incorporated opposite the modification site⁵. Next, cDNA libraries are prepared and sequenced, permitting calculation of reactivity profiles that ultimately inform RNA secondary structure modeling⁵ (Fig. 2.3A).

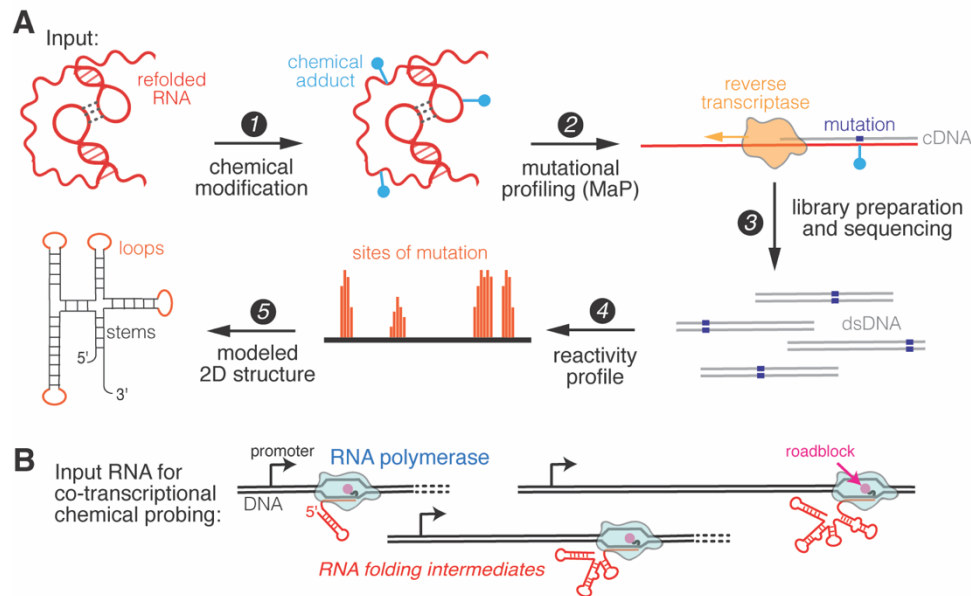


Figure 2.3 Chemical probing to map RNA secondary structure.

A. Workflow of an RNA chemical probing experiment and **B.** RNA polymerase roadblocking strategy for co-transcriptional chemical probing.

3. Co-transcriptional RNA chemical probing

To better recapitulate how riboswitches fold and respond to a ligand in a native context, more laboratories are incorporating the transcriptional machinery into their experimental methods. The chemical probing workflow outlined in the previous section has been updated to incorporate RNAP (Fig. 2.3B), permitting interrogation of riboswitch RNA structures in the context of transcription complexes^{8–10}. The most up-to-date method for co-transcriptional RNA chemical probing is called variable length Transcription Elongation Complex RNA structure probing (TECprobe-VL)¹⁰, and is the method used to study the *alx* and *mntP* riboswitches in Chapter 4. In brief, the TECprobe-VL method involves the following steps: (1) Preparation of a DNA template encoding the RNA sequence of interest with biotinylated dNMPs randomly distributed throughout the template. Streptavidin binds at the randomly distributed biotinylation sites, creating roadblocks to RNAP movement across every position on the DNA template. (2) Reconstitution of transcription *in vitro*, resulting in formation of stalled transcription elongation complexes that are exposing the

nascent, structured RNA transcript. (3) Chemical probing of the displayed RNA is performed with either a SHAPE reagent or a base-specific reagent, which is followed by RNA processing steps to read out where the modification sites are and inform structure models.

4. Gene reporter fusions

LacZ-based reporter gene fusions are used in this work to validate RNA structural predictions with a functional gene expression readout *in vivo*¹¹. Reporter gene fusions are modular and allow for a straightforward test of variables that may alter expression of a gene of interest, such as the promoter sequence, riboswitch variants, length of the protein-coding sequence, and environmental conditions. A transcriptional reporter contains a β -galactosidase (β -gal) gene (*lacZ*) with its own RBS (ribosome binding sequence) (Fig. 2.4A), thus *lacZ* mRNA translation and consequently observed β -gal activity are uncoupled from the riboswitch-controlled translation initiation. The β -gal gene in a translational reporter does not have its own RBS and instead directly follows the first several codons of the gene of interest (Fig.2.4B). Therefore, translation of the β -gal mRNA and observed β -gal activity depend on accessibility of the riboswitch-controlled RBS.

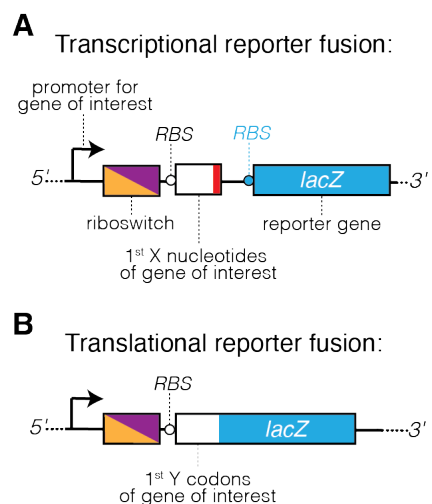


Figure 2.4 Design of reporter fusions used in this work.

A. Sample construct design for *lacZ* transcriptional reporter fusion. **B.** Sample construct design for *lacZ* translational reporter fusion.

5. References

1. Kang, J. Y., Mishanina, T. V., Landick, R. & Darst, S. A. Mechanisms of Transcriptional Pausing in Bacteria. *J. Mol. Biol.* **431**, 4007–4029 (2019).
2. Perdrizet, G. A., Artsimovitch, I., Furman, R., Sosnick, T. R. & Pan, T. Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proc. Natl. Acad. Sci.* **109**, 3323–3328 (2012).
3. Steinert, H., Sochor, F., Wacker, A., Buck, J., Helmling, C., Hiller, F., Keyhani, S., Noeske, J., Grimm, S., Rudolph, M.M., Keller, H., Mooney, R.A., Landick, R., Suess, B., Furtig, B., Wohnert, J. & Schwalbe, H. Pausing guides RNA folding to populate transiently stable RNA structures for riboswitch-based transcription regulation. *eLife* **6**, e21297 (2017).
4. Widom, J. R., Nedialkov, Y.A., Rai, V., Hayes, R.L., Brooks, C.L., Artsimovitch, I. & Walter, N.G. Ligand Modulates Cross-Coupling between Riboswitch Folding and Transcriptional Pausing. *Mol. Cell* **72**, 541-552.e6 (2018).
5. Siegfried, N. A., Busan, S., Rice, G. M., Nelson, J. A. E. & Weeks, K. M. RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959–965 (2014).
6. Merino, E. J., Wilkinson, K. A., Coughlan, J. L. & Weeks, K. M. RNA Structure Analysis at Single Nucleotide Resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE). *J. Am. Chem. Soc.* **127**, 4223–4231 (2005).
7. Mustoe, A. M., Lama, N. N., Irving, P. S., Olson, S. W. & Weeks, K. M. RNA base-pairing complexity in living cells visualized by correlated chemical probing. *Proc. Natl. Acad. Sci.* **116**, 24574–24582 (2019).
8. Strobel, E. J., Watters, K. E., Nedialkov, Y., Artsimovitch, I. & Lucks, J. B. Distributed biotin–streptavidin transcription roadblocks for mapping cotranscriptional RNA folding. *Nucleic Acids Res.* **45**, e109–e109 (2017).
9. Strobel, E. J., Cheng, L., Berman, K. E., Carlson, P. D. & Lucks, J. B. A ligand-gated strand displacement mechanism for ZTP riboswitch transcription control. *Nat. Chem. Biol.* **15**, 1067–1076 (2019).
10. Szyjka, C. E. & Strobel, E. J. Observation of coordinated RNA folding events by systematic cotranscriptional RNA structure probing. *Nat. Commun.* **14**, 7839 (2023).
11. Miller, JH. *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia Coli and Related Bacteria*. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1992).

6. Acknowledgements

Chapter 2, in part, is a reprint of the material as it appears in Opportunities for riboswitch inhibition by targeting co-transcriptional RNA folding events. Stephen, Christine; Palmer, Danea; Mishanina, Tatiana, International Journal of Molecular Sciences, 2024. The dissertation author was the primary co-author of this paper.

Chapter 3 Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the *Escherichia coli* pH-responsive riboswitch

Journal of Biological Chemistry

Title:

Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the
Escherichia Coli pH-responsive riboswitch

Authors:

Christine Stephen¹ and Tatiana V. Mishanina^{1,*}

¹Department of Chemistry and Biochemistry, University of California San Diego, 9500 Gilman
Dr, La Jolla, CA 92093

***Corresponding author information:**

Phone: (858) 246-5883

Email: tmishanina@ucsd.edu

Keywords:

Riboswitch

Bacterial transcription

Transcription

Pausing

RNA

RNA polymerase

Gene regulation

Abbreviations:

PRE, pH-responsive element

sRNA, small RNA

RNAP, RNA polymerase

RBS, ribosome binding site

EC, elongation complex

PEC, paused elongation complex

ePEC, elemental paused elongation complex

NTP, nucleotide triphosphate

DMS, dimethyl sulfate

nt, nucleotide

0. Abstract

Riboswitches are 5'-untranslated regions of mRNA that change their conformation in response to ligand binding, allowing post-transcriptional gene regulation. This ligand-based model of riboswitch function has been expanded with the discovery of a "pH-responsive element" (PRE) riboswitch in *E. coli*. At neutral pH, the PRE folds into a translationally inactive structure with an occluded ribosome binding sequence, whereas at alkaline pH the PRE adopts a translationally active structure. This unique riboswitch does not rely on ligand binding in a traditional sense to modulate its alternative folding outcomes. Rather, pH controls riboswitch folding by two possible modes that are yet to be distinguished; pH either regulates the transcription rate of RNA polymerase (RNAP) or acts on the RNA itself. Previous work suggested that RNAP pausing is prolonged by alkaline pH at two sites, stimulating PRE folding into the active structure. To date, there has been no rigorous exploration into how pH influences RNAP pausing kinetics during PRE synthesis. To provide that understanding and distinguish between pH acting on RNAP versus RNA, we investigated RNAP pausing kinetics at key sites for PRE folding under different pH conditions. We find that pH influences RNAP pausing, but not in the manner proposed previously. Rather, alkaline pH either decreases or has no effect on RNAP pause longevity, suggesting the modulation of RNAP pausing is not the sole mechanism by which pH affects PRE folding. These findings invite the possibility that the RNA itself actively participates in the sensing of pH.

1. Introduction

Organisms implement a diverse array of strategies to tune their gene expression programs in response to changing environmental conditions. One such tuning strategy employed by bacteria is the use of riboswitches – structured, 5'-untranslated regions of a messenger RNA that dictate whether the downstream gene is transcribed and/or translated into protein. Riboswitches typically regulate gene expression by altering their structure upon binding to a nutrient ligand (e.g., metal ions, amino acids, vitamins)¹. A pH-responsive riboswitch discovered in *E. coli*, however, does not require a ligand binding event in a traditional sense to rearrange its structure and instead rearranges its structure in response to the environmental H⁺ concentration. This riboswitch is a “pH-responsive element” (PRE): at neutral pH, the PRE folds into a translationally inactive structure with an occluded ribosome binding sequence (RBS), whereas at alkaline pH, the PRE adopts a translationally active structure with an exposed RBS². This conformational isomerization regulates the expression of the downstream gene, which encodes a putative metal transporter that is highly upregulated under extreme alkaline conditions, thus called “*alx*” (Figs. 3.1 and S3.1)³.

The *alx* gene was first identified in a genetic screen seeking *E. coli* genes that are upregulated in response to alkaline conditions³. Later, a screen for small RNAs (sRNAs) with potential regulatory function in *E. coli* uncovered a sRNA termed SraF that is immediately upstream of the *alx* gene (also known as *ygjT* gene)⁴. Upon further investigation, SraF was definitively characterized as a pH-responsive riboswitch controlling translation of the *alx* mRNA through a combination of *lacZ* reporter fusion assays and RNA chemical probing experiments performed both *in vitro* and *in vivo*². The pH-responsive riboswitch is generally conserved amongst bacteria in the *Enterobacteriales* order², a large and diverse group of gram-negative bacteria that have been found in soil, water, and in association with living organisms⁵ – all of

which are environments where bacteria can experience an external alkaline pH⁶. It appears that the *alx* gene and its control by the PRE has arisen to help counter the stress imposed on bacteria upon a shift into an external alkaline environment that can be encountered in their ecological niches. While non-extremophilic bacteria deploy a wide range of strategies to maintain the cytoplasmic pH in a neutral range (generally 7.4-7.8), previous experiments measuring the cytoplasmic pH of log-phase *E. coli* after a rapid shift to an alkaline medium demonstrated that the cytosol alkalinizes before recovering the desired neutral pH⁷. Thus, the PRE (and RNA polymerase transcribing it) can experience an alkaline environment *in vivo* to prompt its structural rearrangement and increase translation of the Alx protein. The Alx protein is a member of the TerC family of membrane transporter proteins and has been implicated in Mn²⁺ transport⁸; however, the connection between Mn²⁺ transport activity and alkaline pH awaits elucidation. Additionally, the molecular mechanisms underpinning the pH-dependent riboregulation of the *alx* gene remain unknown.

Previous studies suggest that PRE folding is kinetically controlled. Specifically, the model for PRE folding proposed that RNAP polymerase (RNAP) pausing – a temporary inhibition of nucleotide (nt) addition during transcript elongation – is enhanced by alkaline pH at two specific sites during PRE synthesis, which promotes folding of the PRE into the active structure (Fig. 3.1B)^{2,9}. Two key pieces of evidence supported this model. First, unfolding of RNA synthesized at neutral pH, followed by re-folding at alkaline pH did not produce the translationally active conformer, which indicates that pH exerts control over PRE folding co-transcriptionally². Second, RNA made by T7 RNAP (which in general pauses minimally during transcription)¹⁰ assumed the same inactive structure at either neutral or alkaline pH². While these observations support the importance of strategically placed transcriptional pauses to the RNA folding outcome, they do not distinguish between the two possible mechanisms of pH action: (i) pH altering paused RNAP

structure to prolong its residence time at the key pause sites versus (ii) pH acting on the RNA to favor its co-transcriptional folding into the active conformer.

Pausing by RNAP plays crucial regulatory roles in both prokaryotes and eukaryotes by creating windows of time for regulatory events. In addition to allowing transcriptional regulators to (un)bind to RNAP, pausing in bacteria allows synchronization of transcription and translation, promotes the proper folding of nascent RNA transcripts, and is an initial step in transcription termination¹¹. Previous studies of *E. coli* RNAP have elucidated key structural changes in the enzyme when it reaches DNA sequences that induce pausing¹². These structural changes rearrange RNAP's active site into a conformation that cannot catalyze nt addition. Despite these advances, it remains unclear how RNAP pausing might be prolonged as it transcribes the *same* sequence of DNA at alkaline pH compared to neutral pH in the context of the PRE. To address this gap in knowledge, we performed the first rigorous exploration into how pH influences RNAP pausing during PRE synthesis. An improved understanding of the kinetic control mechanism underlying pH-dependent riboswitching would both enrich our basic knowledge of kinetically controlled riboswitches and facilitate its use as an engineerable switch for synthetic gene circuits, which have various biotechnological applications, ranging from biosynthesis of antibacterial drugs to biofuel production¹³.

Our experiments revealed that the first and most critical pause site for RNA folding, termed "Pause 1" in the originally published work, is actually a composite of three sequential pauses. Furthermore, measurements of RNAP pausing kinetics at these pauses showed that pH impacts pause longevity but not in the previously proposed manner. Rather, alkaline pH either accelerates or has no effect on RNAP pause escape. Additionally, we explored RNAP pausing kinetics at "Pause 2", the second critical decision point for the RNA folding. At this site, we find that alkaline pH accelerates the pause escape rate. Taken together, our findings suggest that modulation of RNAP pausing alone is insufficient to explain the alternative PRE folding outcomes under different

pH. We posit that the RNA is involved in sensing and rearranging its structure in response to alkaline conditions.

2. Results

Alkaline pH accelerates *E. coli* RNAP transcription through the pH-responsive riboswitch

To monitor the transcription progress of *E. coli* RNAP under neutral vs. alkaline pH on the full PRE sequence, we performed synchronized single-round *in vitro* transcription assays initiated at the native promoter that controls PRE synthesis *in vivo* (see Fig. S3.2A and S3.2B for detailed DNA template sequence and experimental scheme). The transcription time courses demonstrated that *E. coli* RNAP transcribes through the PRE sequence faster under alkaline pH conditions relative to neutral pH (Fig. 3.2B and 3.2C). This observation is consistent with the chemical mechanism of nt addition, where in the first step a base in the RNAP active site deprotonates the 3' hydroxyl of the terminal nt on the RNA transcript (Fig. 3.2A). Therefore, as pH alkalinizes, the fraction of the catalytic base in the deprotonated form increases, explaining the observed increase in the transcription rate. This observation is also consistent with previous experimental analysis of *E. coli* RNAP transcription kinetics under different pH conditions¹⁴. Although alkaline pH clearly enhances the average transcription rate (Fig. 3.2C), these measurements do not directly address the question of pause longevity, which may be obscured by other events in the transcription elongation assay. The prevailing PRE folding model proposes that RNAP pause escape is *slowed* at alkaline pH versus neutral pH to support folding of specific structural motifs in the active PRE conformer.

We observed pausing on the PRE sequence at similar positions to those reported in the original work (Fig. 3.2B)². To begin we analyzed the pause site termed “Pause 1”, which was identified as the first critical pause location for PRE folding. Specifically, it was proposed that prolonged pausing at this site at alkaline pH provides a time window for folding of “Hairpin S”, a key structural element in the PRE active conformer (Fig. 3.1B; see Fig. S3.1 for PRE secondary structure models at different pH and Table S3.1 for PRE sequence)². First, we find that the RNA

band identified as “Pause 1” in previous work is actually a composite of three sequential pause RNA species (Fig. 3.3B). We analyze the reasons for the discrepancy in the following section, and for consistency with prior work, we will continue to refer to this composite of three pauses collectively as “Pause 1”. We estimated the average dwell time of RNAP at Pause 1 by treating the diffuse band as a single RNA species, following established methodologies to determine pause kinetic parameters^{15,16}. Since RNAP arrives at Pause 1 more rapidly at pH 8.5 vs 7.2, we fit kinetic parameters to data points where continued arrival of RNAP at the pause was negligible (10 s – 20 s for pH 7.2 and 8 s – 12 s for pH 8.5). This analysis yielded average pause dwell times of 13 s at pH 7.2 and 9 s at pH 8.5 (Fig. S3.2C), suggesting that, in contrast to the prevailing model, RNAP resides at Pause 1 for longer under neutral conditions. However, this promoter-initiated data did not permit individual kinetic treatment of each sequential pause comprising Pause 1; therefore, it was unclear whether each pause lifetime was longer under neutral conditions, or if a single pause RNA species in the diffuse band was biasing the results. To truly address whether alkalinity prolongs RNAP pausing at the first, and most important site for PRE folding into the active conformer, we employed scaffold-based transcription assays, described in the subsequent sections.

The prevailing PRE folding model proposes that an additional pause in the folding pathway, “Pause 2”, is prolonged by alkaline pH to support folding of Hairpin C/D in the active conformer (Figs. 3.1B and S3.1)². We observed multiple pause RNA species in the expected region for Pause 2 prior to the termination site D (Fig. 3.2B, see the next paragraph for description of termination sites on PRE). Similar to our analysis of Pause 1, we fit kinetic parameters to data points where continued arrival of RNAP at the pause was negligible (14 s – 20 s for pH 7.2 and 8 s – 12 s for pH 8.5). This analysis yielded average pause dwell times of 11 s at pH 7.2 and 7 s at pH 8.5 (Fig. S3.2C), suggesting that RNAP also resides at Pause 2 for longer under neutral conditions. This data, however, presents the same limitations addressed in our analysis of Pause

1, prompting us to explore the kinetics of Pause 2 further with scaffold-based transcription assays that are also described in the subsequent sections.

Besides the Pause 1 and 2 sites, the other prominent RNA species observed in this experiment (~108 nt and ~190 nt) correspond to intrinsic termination sites (Fig. 3.2B). Intrinsic termination occurs when an RNA hairpin forms in the RNAP exit channel while RNAP is positioned on an unstable RNA-DNA hybrid, prompting dissociation of the elongation complex (EC)¹⁷. Hairpins B and D in the PRE are characteristic intrinsic terminator hairpins with hairpin structures followed by a U-tract (Fig. S3.1); for consistency with prior work, we will refer to these termination sites as termination site B and termination site D, named for their respective hairpin structures. A slightly larger fraction of ECs terminated at site B at neutral pH (Fig. S3.2C), with ~46% of ECs terminating here at pH 7.2 and ~43% of ECs at pH 8.5. A slightly larger fraction of ECs terminated at site D at neutral pH (Fig. S3.2C), with ~13% of ECs terminating here at pH 7.2 and ~10% of ECs at pH 8.5. The RNA species corresponding to termination sites B and D persist in the “chase” lane (Fig. 3.2B), which represents a 5-min reaction at saturating nucleotide triphosphates (NTPs, 500 μ M). Thus, these locations are definitively termination sites and not long-lived pauses.

Modulation of RNAP pausing by transcription factors during synthesis of the pH-responsive riboswitch

To gain further insights into the nature of the pauses, we performed transcription elongation assays in the presence of *E. coli* transcription factors known to modulate pausing. We first tested for backtracked pauses by including the transcription factor GreA during *in vitro* promoter-initiated transcription assays at different pH. A backtracked pause is a long-lived pause that occurs when RNAP reverse-translocates on the RNA-DNA hybrid and extrudes the 3'-end of the RNA through its secondary channel, thereby precluding an incoming NTP from entering the active site¹¹. GreA relieves RNAP from a backtrack by stimulating cleavage of the extruded 3'-

RNA piece, permitting RNAP to continue transcription elongation^{18,19}. GreA exerts a clear effect at the Pause 1 site under both neutral and alkaline conditions by consolidating the multiple RNA bands observed at Pause 1 into a single band, which persists at later time points relative to the no-factor condition (Fig. S3.2D). This suggests that a backtracking event occurs at Pause 1, which we confirmed with scaffold-based transcription assays outlined in later sections.

We next included the transcription factor NusA into *in vitro* transcription elongation assays to expose potential hairpin-stabilized pauses. Hairpin-stabilized pauses occur when an RNA hairpin in the RNAP exit channel extends pause longevity¹¹ and are enhanced in the presence of NusA through stabilizing contacts between NusA and paused RNAP²⁰. Previous work supported the existence of hairpin-stabilized pauses at both Pause 1 and Pause 2²; however, reproducibility and gel quantification of these results were not reported. We first performed promoter-initiated transcription elongation assays in the presence of 500 nM NusA, matching the concentration used in previously reported experiments². In our experiments, we observed multiple putative hairpin-stabilized pauses prior to termination site B that are enhanced by NusA as well as increased termination at site B (Fig. S3.2E). We also noticed a significant decrease in transcription elongation past the termination site B at 500 nM NusA such that Pauses 1 and 2 were not detected (Fig. S3.2E), in stark contrast to previous work². This is likely caused by NusA stabilization of the intrinsic terminator hairpin at termination site B, increasing the fraction of ECs that dissociate here.

In an attempt to observe transcription and the pauses past termination site B, we performed transcription elongation assays with NusA present at lower concentrations (50 nM, 100 nM, and 250 nM). The effect of NusA at 100 nM and 250 nM mimics that seen with the 500 nM condition, with multiple pauses enhanced by NusA and increased termination at termination site B, but not many transcripts produced beyond that point (Fig. S3.2E). At 50 nM NusA, we observe sufficient elongation past termination site B to visualize Pauses 1 and 2; however, we cannot

visualize enhancement of the pauses prior to termination site B (Fig. S3.2E), suggesting that the 50 nM concentration of NusA is insufficient to observe pause enhancement in this experiment.

Further gel analysis of transcription elongation time courses with NusA suggests that three hairpin-stabilized pauses reside at positions C36, C58, and C73/A74 on the PRE (Figs. S3.2F and S3.1). At positions C73/A74, Hairpin A is appropriately spaced from the RNAP active site such that it could reside in the RNA exit channel and prompt a hairpin-stabilized pause (Fig. S3.1; see Fig. S3.7A for an example of appropriate spacing in the context of a well-studied *his* pause). At positions C36 and C58, there are no hairpin structures present in the fully folded RNA to rationalize the observed enhancement of these pauses in the presence of NusA. Nevertheless, it is possible the PRE traverses through labile folding intermediates during its synthesis similarly to other non-coding RNAs^{21–23} and riboswitches in particular^{24–26}, thus the putative hairpins present in the RNAP exit channel at positions C36 and C58 cannot be visualized in the fully folded RNA structure. Prior work on the pH-responsive riboswitch reported that NusA promotes formation of the active PRE conformer under neutral conditions, mimicking the effect of alkaline pH². Since we cannot find evidence for hairpin-stabilized pauses in the critical region for pH-responsive riboswitching, it is unclear how NusA produced this effect.

Three sequential pauses constitute the first critical decision point for pH-responsive riboswitch folding

To assess RNAP pausing on the PRE under different pH conditions, we performed a different class of transcription assay that bypasses the need for initiating transcription from a promoter and zooms in on the locations of the pauses, thus simplifying kinetic analysis. We first positioned RNAP on a short, synthetic nucleic acid scaffold corresponding to the reported location of Pause 1^{2,9}, which was reportedly prolonged by alkaline pH to promote the folding of Hairpin S in the PRE active conformer (Figs. 3.1B and S3.1). To test whether the pause escape rate at this

site is slowed by alkaline pH, we measured the escape rate of *E. coli* RNAP from this proposed pause under neutral versus alkaline pH conditions. Under both pH conditions, rapid RNA extension was observed (Fig. S3.3A), which is behavior indicative of an active EC rather than a paused elongation complex (PEC). Therefore, this experiment demonstrated that the location of Pause 1 was originally mapped incorrectly, likely due to limited resolution of the gel analysis of long RNAs produced from promoter-initiated transcription.

To elucidate the sequence motifs responsible for the pauses, we performed a scaffold-based transcription assay that positioned the ECs upstream of the sequence window where Pause 1 lies (Fig. 3.3A). An RNA sequencing ladder was generated for precise assignment of the pause site. These experiments located Pause 1 ~4 nt downstream of the originally published site and uncovered that the paused RNA species termed “Pause 1” in the originally published work is a composite of three sequential pauses (Fig. 3.3B). In the original RNA gel analysis, which employed a 6% denaturing PAGE, these paused RNAs migrated as a single diffuse band. We also observed this diffuse band on our 6% gels (Fig. S3.2B). To ensure that these sequential pauses are the biologically relevant ones, and not an artifact from using a minimalistic transcription scaffold, we clearly captured the sequential pauses in promoter-initiated transcription experiments (Fig. 3.3D). Specifically, each of the sequential pause RNA species was observed by limiting the incoming NTP at each of the respective pause sites and analyzing RNA products on a 10% (vs. originally used 6%) denaturing PAGE, which allowed nt resolution in the region of the gel where the pause RNA species migrate. We also performed the scaffold-based pause mapping experiment with limiting NTP conditions and observed an increase in pause duration at each sequential pause site (Fig. S3.3B).

We compared the sequential pause sequences to the consensus elemental pause sequence (Fig. 3.3C), which was previously identified using high-throughput sequencing of nascent RNAs from bacteria (NET-seq)²⁷. The pauses at position 142 and 143 on the PRE

sequence lack the -1 CG base pair (bp) and +1 GC bp present in the consensus sequence. However, the pause at 142 shares the -10 GC bp with the consensus sequence, and the pause at 143 shares both the -10 and -11 GC bp with the consensus sequence. Lastly, the pause at position 144 on the PRE sequence shares the +1 GC bp and -11 GC bp with the consensus sequence. Previous kinetic analyses of *E. coli* RNAP on the consensus elemental pause sequence have demonstrated that the pause signal is multipartite, with individual sequence elements having an additive effect on overall pause strength²⁸. Therefore, all elements of the consensus pause sequence are not necessary to observe a pause, as seen with the sequential pauses we identified on the PRE. Furthermore, to assess whether the newly mapped pauses at positions 142-144 on the PRE sequence occur *in vivo*, we analyzed the previously reported NET-seq data for the entire *E. coli* genome²⁷. From this data, three peaks are observed at the genomic locations corresponding to positions 142-144 on the PRE (Fig. S3.3C), supporting their existence *in vivo*.

***E. coli* RNAP pause longevity is either decreased or unaffected by alkaline pH at the first major pause sites in the pH-responsive riboswitch folding pathway**

To test whether RNAP pausing is prolonged by alkaline pH at the sequential pauses, we reconstituted ECs at each of the sequential pause sites and measured the “elemental” pause lifetimes at neutral and alkaline pH. When RNAP first encounters a pause signal in the transcribed nucleic acids, it rearranges into an elemental paused elongation complex (ePEC), entering into a transient, catalytically inactive state (Fig. 3.3C)²⁸. Depending on the sequence, the ePEC can either return to an active elongation conformation or isomerize into a longer-lived pause state such as a backtracked or a hairpin-stabilized pause described in the earlier section. In these experiments, we focused solely on elemental pause longevity (without upstream RNA structure) and its response to pH because the elemental pause is an obligate precursor for longer-lived

pauses that may mediate PRE folding. This setup permitted direct evaluation of the alkaline pH effect on the paused RNAP and thus the proposed PRE folding model. An evaluation of the proposed folding model could not be performed based on promoter-initiated transcription experiments alone for the following reasons: (i) the elemental pause longevities could not be measured since the full-length nascent RNA is present and (ii) promoter-initiated experiments did not permit rigorous kinetic treatment of each individual pause in the sequential series due to poor resolution of pause bands on the gel and large number of other RNA species present. Kinetic and structural analyses of PECs reconstituted on synthetic nucleic acid scaffolds have enriched our fundamental understanding of transcriptional pausing^{12,28,29}, prompting our application of scaffold-based transcription assays to study the PRE pauses here.

In our experiments, ECs were assembled on synthetic nucleic acid scaffolds corresponding to each sequential pause that contain a fully complementary transcription bubble. The 5' radiolabeled RNA primer used for each scaffold was 17-nt in length with an 11-nt region of complementarity (to form an RNA-template DNA hybrid) and a 6-nt non-complementary 5'-tail mimicking nascent RNA in the RNA exit channel of RNAP. Additional complementarity was not introduced in the upstream RNA segment to prevent ECs from assembling at an undesired location upstream of the pauses. Following PEC assembly, we measured the fraction of pause RNA over time at pH 7.2 and 8.5 when the radiolabeled PECs were extended with the NTP required for pause escape. To obtain kinetic parameters, a semilogarithmic plot was generated for the pause RNA fraction over time, and a single-exponential decay fitting was performed on the plot (Fig. S3.4A). At the first pause, pause decay was faster at alkaline pH, with an average RNAP dwell time at the pause of 54 s at pH 8.5 vs 174 s at pH 7.2 (Fig. 3.4A and Table S3.2). At the second pause in the sequential series, the average pause dwell time was 33 s at pH 7.2 vs 34 s at pH 8.5, demonstrating a similar RNAP residence time at this pause at both pH (Fig. 3.4B). At the third pause, pause decay was faster at alkaline pH, with an average pause dwell time of

16 s at pH 8.5 and 32 s at pH 7.2 (Fig. 3.4C). However, the calculated dwell times for the third pause do not accurately depict the pause escape rate since a large fraction of the ECs are inactivated, evidenced by the significant remaining fraction of PECs after incubation with a saturating concentration of the incoming NTP. We demonstrated that ECs remaining at the third pause are backtracked (see Supplemental Discussion), with a backtracking event prompted by the presence of a destabilizing AU-rich DNA-RNA hybrid – a sequence motif known to promote backtracking^{11,30}. Besides the backtracking event at the third pause, in general we observed RNA products resulting from intrinsic cleavage of ECs positioned at each pause, which are characterized and analyzed in detail in the Supplemental Discussion.

Taken together, the sequential pause lifetimes observed in our scaffold experiments are significantly longer than those calculated from promoter-initiated transcription experiments, which can be attributed to multiple factors. The shortened downstream DNA in scaffold experiments compared to promoter-initiated experiments likely increased pause lifetimes, an observation consistent with prior analyses of *E. coli* RNAP pausing³¹. This, for example, is consistent with reported kinetic analyses of the *his* pause – a well-characterized hairpin-stabilized pause in the leader region of the histidine biosynthetic operon – in a promoter vs scaffold-based transcription experiments^{12,15}. The absence of the nascent riboswitch RNA in scaffold experiments also presumably affected the details of observed pausing kinetics. In particular, a significant fraction of RNAP cannot escape the backtrack at position 144 in our scaffold experiments (Fig. 3.4C); however, when the full-length nascent RNA is present in promoter-initiated experiments the pause RNA species at 142-144 largely disappear after 30 s (Fig. 3.2B). This agrees with previous reports that highlight co-transcriptionally folded RNA imposing a kinetic barrier to RNAP backtracking^{32,33}. Although there are clear differences in pause longevity between the scaffold and promoter-initiated experiments, the general effect of pH on pause kinetics was consistent: the average time

RNAP spends at any of the individual sequential pauses, or at Pause 1 collectively, is not increased by alkaline pH.

While the above scaffold experiments provided a useful initial characterization of the sequential pauses, many open questions remained regarding how pH influences RNAP pausing. For example, these scaffold experiments only permitted evaluation of pause lifetimes and not pause efficiencies – a parameter that describes the fraction of RNAP molecules captured by a particular pause sequence³⁴. Additionally, these experiments lacked a complete characterization of the backtracking event at the third pause. Experiments positioning ECs directly at the sequential pauses outlined here *suggest* that PECs enter a backtracked registers at the third pause; however, this data could not provide a comprehensive depiction of both the backtracked register and the pause escape kinetics when ECs encounter this pause during active transcription elongation. To address these limitations and provide a precise dissection of how pH modulates both pause escape and pause efficiency at the sequential pauses, we performed additional scaffold-based transcription assays described in the next section.

Gre factor-stimulated RNA cleavage increases traffic through the sequential pauses

To further investigate both the backtracking at the third pause and the sequential pauses in general, we performed additional scaffold-based transcription assays where RNAP arrived at each pause during active transcription elongation vs positioned directly at the pause, as in the preceding section. These experiments permitted evaluation of both pause lifetimes and pause efficiencies at different pH in the presence and absence of Gre factors (Table S3.3). Briefly, we reconstituted ECs on a synthetic DNA-RNA scaffold corresponding to position 136 on the PRE (G10 ECs), and the G10 ECs were radiolabeled via incorporation of [α -³²P] ATP, extending the RNA to form A11 ECs (Fig. 3.5A). Transcription elongation was then initiated upon addition of all 4 NTPs to the A11 ECs, with the incoming NTP at the pause being analyzed present at a limiting

concentration. Limiting the incoming NTP at each pause increased the pause duration and permitted manual quenching. Similar to the analysis performed in the previous section, a semilogarithmic plot was generated for the pause RNA fraction over time and a single-exponential decay fitting was performed on the plot.

Pause 144 analysis. We first performed transcription elongation assays at a limiting concentration of GTP (20 μ M) to increase the duration of the third pause and evaluate how pH affects both the pause lifetime and pause efficiency in the presence and absence of Gre factors (Figs. 3.5B and 3.5C). The fraction of U18 RNA corresponding to the third pause was plotted over time to perform kinetic analysis (Figs. 3.5B and 3.5C). In the absence of Gre factors, pause escape was faster at alkaline pH, with an average pause dwell time of 130 s at pH 7.2 compared to 108 s at pH 8.5 (Figs. 3.5B and 3.5C). 22% of ECs entered the pause at pH 7.2 vs. 11% at pH 8.5 (Figs. 3.5B and 3.5C), suggesting that the third pause captured a greater fraction of ECs at neutral pH vs. alkaline pH. These observations are consistent with the experiments positioning ECs directly at the third pause to measure the escape rate, where we observed an incomplete decay for the pause RNA fraction over time and a greater population of PECs that escaped the pause at alkaline pH vs. neutral pH (Fig. 3.4C).

Upon addition of GreA, pause escape was slower at alkaline pH, with an average pause dwell time of 40 s at pH 7.2 and 70 s at pH 8.5 (Figs. 3.5B and 3.5C). When GreB was present, pause escape was faster at alkaline pH, with an average pause dwell time of 38 s at pH 7.2 and 27 s at pH 8.5 (Figs. 3.5B and 3.5C). On the whole, addition of either GreA or GreB decreased the average pause dwell time compared to the no-factor condition under both neutral and alkaline conditions, supporting that the Gre factors stimulated RNA cleavage to promote rescue of backtracked PECs. It is unclear why the pause lifetime increased at pH 8.5 in the presence of GreA yet decreased at pH 8.5 in the presence of GreB, since it is expected that alkaline pH would increase the rate of RNA cleavage via hydrolysis in general¹⁹. While we did not observe a

consistent trend between pH and pause lifetimes at the third pause, the effect of pH on pause efficiency was similar in the presence and absence of Gre factors. With GreA present in the transcription elongation assay, 12% of ECs entered the pause at neutral pH compared to 5% at alkaline pH (Figs. 3.5B and 3.5C). When GreB was present, 18% of ECs entered the pause at neutral pH compared to 11% at alkaline pH (Figs. 3.5B and 3.5C). We consistently observed that the pause efficiency was greater at neutral pH vs. alkaline pH in the presence and absence of Gre factors, suggesting that a smaller population of ECs entered the third, backtracked pause under alkaline pH during active transcription elongation. Additionally, we observed a clear increase in the U16 RNA fraction upon addition of Gre factors (Fig. S3.5A) and a concomitant decrease in the U18 fraction, supporting that the U16 RNA results from Gre factor-stimulated cleavage of backtracked PECs. Thus, the cleavage event at the third pause repositions ECs at the first pause in the sequential series, increasing the flux of ECs through sequential pauses (see Supplemental Discussion for further analysis of the translocation register at the third pause during active elongation). Altogether, our data supports that ECs at the third pause are backtracked and this backtracking event occurs when ECs encounter the third pause during active transcription elongation, which holds potential relevance for PRE RNA folding.

Pause 142 analysis. To visualize the first pause and probe the effect of Gre factors on its longevity, we performed transcription elongation assays at a limiting concentration of ATP (10 μ M) (Fig. S3.5B). The fraction of U16 RNA corresponding to the first pause was plotted over time to perform kinetic analysis (Fig. S3.5B). In the absence of Gre factors, pause escape was slower at neutral pH, with an average pause dwell time of 45 s at pH 7.2 compared to 20 s at pH 8.5 (Fig. S3.5B). This trend is consistent with results from the transcription assay that positioned ECs directly at the first pause (Fig. 3.4A). 55% of ECs entered the pause at pH 7.2 vs. 27% at pH 8.5 (Fig. S3.5B), demonstrating that the first pause captured more ECs at neutral pH vs. alkaline pH.

When either GreA or GreB was introduced in the limiting ATP condition, there was a marked increase in the U16 RNA fraction, resulting from a backtracking and Gre-catalyzed 2-nt or 3-nt cleavage at the third pause that repositioned the ECs at the first pause (Fig. S3.5B). The U16 RNA fraction became more prominent in this experiment because ATP, the incoming NTP at the first pause, was present at a limiting concentration. In the presence of GreA, pause escape was slower at neutral pH, with an average pause dwell time of 85 s at pH 7.2 compared to 29 s at pH 8.5 (Fig. S3.5B). In the presence of GreB, pause escape was also slower at neutral pH, with an average pause dwell time of 178 s at pH 7.2 compared to 53 s at pH 8.5 (Fig. S3.5B). Under both pH conditions, the average time spent by ECs at the first pause increased in the presence of Gre factors due to repopulation of ECs at the first pause following the cleavage event at the third, backtracked pause. Specifically, in the presence of GreA, 63% of ECs entered the first pause at pH 7.2 compared to 40 % at pH 8.5 (Fig. S3.5B), exhibiting that the first pause captured more ECs at neutral pH vs. alkaline pH. In the presence of GreB, 73% of ECs entered the first pause at pH 7.2 compared to 76% at pH 8.5 (Fig. S3.5B), demonstrating that the first pause captured a slightly greater fraction of ECs at alkaline pH vs. neutral pH. Altogether, our analysis of the first sequential pause corroborates the kinetic results obtained at the third sequential pause: ECs at the third pause are backtracked and Gre factor-stimulated cleavage increases RNAP traffic through the sequential pauses. Furthermore, we consistently observed that pause escape is accelerated by alkaline pH when the first pause is encountered by ECs during active transcription elongation.

Pause 143 analysis. Lastly, we performed transcription elongation assays at a limiting concentration of UTP (10 μ M) to increase the duration of the second pause and evaluate the pause lifetime and efficiency at different pH. The fraction of A17 RNA corresponding to the second pause was plotted over time to perform kinetic analysis (Fig. S3.5C). In the absence of Gre factors, pause escape was faster at alkaline pH, with an average pause dwell time of 34 s at pH

7.2 and 30 s at pH 8.5 (Fig. S3.5C). 18% of ECs entered the pause at pH 7.2 compared to 17% at pH 8.5 (Fig. S3.5C), demonstrating that the second pause captured a similar fraction of ECs at both pH.

When either GreA or GreB was introduced in the limiting UTP condition, there was a marked increase in both the U16 and A17 RNA fractions (Fig. S3.5C). The increase in the U16 and A17 RNA fractions was due to an increased flux of ECs through the first and second pauses following the backtracking and Gre-catalyzed 2-3-nt RNA cleavage at the third pause. The A17 RNA corresponding to the second pause became more prominent in this experiment since UTP, the incoming NTP at the second pause, was limiting. In general, the A17 RNA fraction did not increase to the same degree as the U16 RNA fraction upon addition of Gre factors. It appears that the second pause is weaker than the first pause and consequently captures fewer ECs that re-enter elongation following backtracking and cleavage at the third pause. Interestingly, we observed that the pause efficiencies were greater at alkaline pH vs. neutral pH in the presence of Gre factors. Specifically, upon addition of GreA 32% of ECs entered the pause at pH 7.2 vs. 36% of ECs at pH 8.5 (Fig. S3.5C). In the presence of GreB, 26% of ECs entered the pause at pH 7.2 vs. 47% of ECs at pH 8.5 (Fig. S3.5C). For the measured pause escape rates, the effect of pH was consistent with what was observed in the no-factor condition, with faster pause escape at alkaline pH. In the presence of GreA, pause escape was faster at alkaline pH, with an average pause dwell time of 26 s at pH 7.2 and 20 s at pH 8.5 (Fig. S3.5C). In the presence of GreB, pause escape was also faster at alkaline pH, with an average pause dwell time of 70 s at pH 7.2 and 34 s at pH 8.5 (Fig. S3.5C). Overall, our observations in these experiments remained consistent with the results outlined in prior sections, further supporting an increase in flux of ECs through the sequential pauses following Gre factor-assisted cleavage at the third pause. Additionally, similar to our observations at the first pause, we consistently observed that pause

escape is faster at alkaline pH when the second pause is encountered by ECs during active transcription elongation.

***E. coli* RNAP pause longevity is decreased by alkaline pH at the second major pause site in the pH-responsive riboswitch folding pathway**

We next investigated how pH affects elemental pause escape kinetics at Pause 2, the second critical decision point for PRE folding. To locate Pause 2 and measure the pause lifetimes at different pH, we performed scaffold-based transcription assays in a similar fashion to our analysis of Pause 1. Specifically, we reconstituted ECs on a synthetic DNA-RNA scaffold corresponding to position 173 on the PRE (G10 ECs) – upstream of the originally mapped Pause 2 location at the PRE nt 182 – and the G10 ECs were radiolabeled via incorporation of [α - 32 P] ATP, extending the RNA to form A11 ECs (Fig. 3.6A). Transcription elongation was initiated upon addition of all 4 NTPs to the A11 ECs, with the incoming NTP corresponding to each pause present at a limiting concentration to increase pause duration and permit manual quenching. To obtain kinetic parameters, a semilogarithmic plot was generated for the pause RNA fraction over time and a single-exponential decay fitting was performed on the plot. In the original PRE folding model, it was proposed that Pause 2, in addition to Pause 1, is prolonged by alkaline pH to promote folding of Hairpin C/D in the PRE active conformer (Figs. 3.1B and S3.1). Similar to our investigation of Pause 1, we focused solely on elemental pause longevity and its response to pH since the elemental pause is a necessary precursor for longer-lived pauses that may mediate PRE folding.

We first performed this experiment limiting the concentration of GTP to test for Pause 2, which was originally reported to reside at position 182 on the PRE sequence² (see Fig. S3.1 and Table S3.1 for PRE sequence). However, we did not find evidence for this pause location in our experiments (Fig. S3.6A), suggesting that Pause 2 was mapped incorrectly. The reported location

of Pause 2 was likely incorrect due to limited resolution of the gel analysis of long RNAs produced from promoter-initiated transcription, similarly to Pause 1 mapping. Nevertheless, multiple pauses were enhanced in the UTP limiting condition at both pH, indicating that these are most likely the relevant pauses prior to the termination site D (Fig. 3.6B). Consistent with this observation, in promoter-initiated transcription experiments we observed increased pause longevity for the pauses prior to termination site D in the UTP limiting condition (Fig. 3.6C). We mapped these pauses to positions 177, 183, and 184 on the PRE sequence (Fig. 3.6B), placing them in the vicinity of the originally mapped Pause 2 location. We compared these pause sequences to the consensus elemental pause sequence (Fig. S3.6B), which was identified in a prior NET-seq study²⁷. Overall, the pauses do not contain many sequence features in common with those most frequently observed *in vivo* in *E. coli*; however, the pause consensus results from genome-wide averaging and it is not necessarily expected that all pauses will share sequence similarity with the consensus sequence. Additionally, to assess whether the newly mapped pauses at positions 177, 183, and 184 occur *in vivo*, we analyzed the previously reported NET-seq data for the entire *E. coli* genome²⁷. From this data, we observed sequencing reads above baseline for position 177 but not for positions 183 and 184 (Fig. S3.6C). It is unclear why there is a discrepancy between our *in vitro* pause mapping data and the NET-seq results. Further experimentation will be required to confirm the locations of the pauses prior to termination site D *in vivo*.

With the pauses prior to termination site D mapped, we performed additional transcription elongation assays at a limiting concentration of UTP (10 μ M) to increase the duration of each pause and evaluate the pause lifetimes and efficiencies at different pH (Table S3.4). The fraction of U14, G20, and U21 RNAs were plotted over time to perform kinetic analysis of the pauses at positions 177, 183, and 184, respectively (Fig. 3.6D). For the elemental pause at position 177, pause escape was faster at alkaline pH, with an average pause dwell time of 50 s at pH 7.2 and 34 s at pH 8.5 (Fig. 3.6D). At position 183, pause escape was also faster at alkaline pH, with an

average pause dwell time of 252 s at pH 7.2 and 80 s at pH 8.5 (Fig. 3.6D). Lastly, at position 184, the average pause dwell time was 110 s at pH 7.2 and 37 s at pH 8.5 (Fig. 3.6D), demonstrating faster pause escape at alkaline pH. Altogether, pause escape was faster at alkaline pH for each pause prior to the termination site D, which is consistent with our analysis from promoter-initiated transcription experiments where we also observed faster pause escape at alkaline pH (Fig. S3.2C).

3. Discussion

In this work we investigated the proposed folding mechanism for the *E. coli* pH-responsive riboswitch. Previous study clearly supports that PRE folding is kinetically controlled via strategically located RNAP pause sites². This prior study reported promoter-initiated transcription experiments in which multiple pause sites were visualized at both neutral and alkaline pH². From this data, gel quantification of the pause RNA species was performed using the metric “half maximal intensity” and two pause species in particular were found to have greater half maximal intensity values for later time points at alkaline pH². This analysis led to the conclusion that alkaline pH enhances RNAP pausing at two sites to support PRE folding into the translationally active conformer. The method of gel quantification and reproducibility, however, were not specified and the overall effect of alkalinity on the transcription rate was not considered. Thus, these previously reported observations did not discriminate between two potential mechanisms of pH action: pH rearranging paused RNAP structure to prolong its stall time at key pause sites vs. pH acting directly on the RNA co-transcriptionally to promote its folding into the active conformer. Here we performed a rigorous analysis of *E. coli* RNAP pausing kinetics under different pH conditions at the biologically relevant pauses for PRE folding and posit that the RNA plays a role in sensing the cytosolic pH.

We initially investigated the first major pause site in the PRE folding pathway, which was suggested to facilitate the folding of Hairpin S, a key structural element in the active conformer thought to ultimately determine whether the ribosome binding site in the PRE is accessible (Fig. S3.1). We uncovered that the pause downstream of Hairpin S was mapped incorrectly; instead, our pause mapping experiments demonstrated a series of three sequential pauses ~4 nt downstream of the originally published site. This updated location is still consistent with Pause 1 being a critical driver of PRE RNA folding, where point mutations within the pause sequence that changed either the duration or location of Pause 1 altered *alx* gene expression⁹. Specifically, the A132G mutation was found to significantly increase the duration of Pause 1 relative to the wild-type PRE sequence, whereas the C135U and A137G substitutions shifted the location of Pause 1 downstream from its original position⁹. With the improved sequence location of Pause 1, we can shed new insights into the observed effects of the mutations on pausing. The A132G mutation places a G in the -11 position at the first sequential pause, increasing its sequence similarity to the pause consensus and explaining the observed increase in Pause 1 duration (Fig. S3.8). Since C135U and A137G were found to shift the location of Pause 1 downstream of the original site, it is possible the mutations increased the propensity of RNAP to pause at a location downstream of the wild-type sequence. For example, the A137G mutation creates a sequence motif with a high degree of similarity to the pause consensus shifted a few nt downstream of the sequential pauses (Fig. S3.8).

pH effect on the key pauses in the PRE folding pathway. Once mapped more precisely, we placed RNAP at the Pause 1 sequential pauses and measured the elemental pause escape rate at neutral vs. alkaline pH, thus testing whether alkaline pH acts directly on paused RNAP to drive the ultimate RNA folding outcome. We found that the elemental pause escape rate at these pauses is not increased by alkaline pH, demonstrating that alkaline pH does not alter paused RNAP structure directly to facilitate PRE folding into the active conformer. Further

experiments to measure pause escape kinetics at the sequential pauses produced the following insights: (i) the third pause in the sequential series is backtracked and likely adopts the 1- to 2-nt backtracked registers when the pause is encountered during active transcription elongation; (ii) Gre factor-stimulated RNA cleavage at the third pause increases traffic through the sequential pauses, and (iii) in general, pause escape was faster at alkaline pH vs. neutral pH across all pauses in the sequential series when encountered by ECs during active transcription elongation. Additionally, in promoter-initiated transcription assays where the full-length PRE RNA is present, we observed that addition of GreA consolidated the multiple RNA bands constituting Pause 1 and enhanced the duration of earlier pauses in the sequential series – confirming the relevance of our results from scaffold-based transcription experiments.

We also investigated the pH dependence of pause escape at Pause 2, which was proposed to be prolonged by alkaline pH to promote the folding of Hairpin C/D in the active conformer (Figs. 3.1B and S3.1). Similar to Pause 1, we found that pause escape at Pause 2 is accelerated by alkaline pH. Altogether, our results demonstrate that pH exerts an effect on paused RNAP independent of the nascent riboswitch RNA but not in the originally proposed manner. Instead, our data support that ECs spend more time on average at Pauses 1 and 2 at *neutral pH*. Therefore, a new framework is required to explain the changes in PRE RNA structure that occur at different pH, as we propose in a later section.

Potential role of nascent RNA structure modulating pause longevities during PRE synthesis. While previous work suggested nascent RNA hairpin modulating Pause 1 longevity², the newly mapped locations for Pause 1 call this assertion into question. The most well-characterized pause class involving nascent RNA structure is the hairpin-stabilized pause, where the RNA hairpin is positioned ~11 nt away from the RNA 3' end at the active site (Fig. S3.7A). The three sequential pauses we mapped downstream of Hairpin S are 17-19 nt away from the RNA 3' end, positioning Hairpin S well outside of the RNAP exit channel, even upon backtracking

by 1-2 nt at the third pause; therefore, the sequential pauses are likely unaffected by the proposed Hairpin S in the nascent RNA (Fig. S3.7B). Additionally, Hairpin S is not the only RNA structure forming in the vicinity of Pause 1 since Hairpin C, which forms in the inactive conformer, also nucleates in proximity of Pause 1 (Fig. S3.1). In contrast to Hairpin S, which is spaced too far from the RNAP exit channel to promote a hairpin-stabilized pause, Hairpin C can only begin nucleation after ECs have bypassed Pause 1 (Fig. S3.7C).

Although it appears unlikely that either Hairpin S or Hairpin C modulate Pause 1 longevity, an interplay between the nascent PRE RNA and pausing cannot be ruled out. Our test for hairpin-stabilized pauses in the region critical for pH-responsive riboswitching yielded inconclusive results. Specifically, NusA significantly enhanced termination at site B and decreased the flux of ECs through Pauses 1 and 2, preventing a definitive conclusion as to whether hairpin-stabilized pauses exist at these locations. At this point, we can only consider the influence of RNA structural motifs present in the fully folded RNA on the pauses, which further limits our analysis. It is possible that the PRE traverses through transient structures that have not yet been characterized. For example, we observed multiple pauses enhanced by NusA prior to the termination site B that could not be rationalized in the context of the full-length RNA secondary structure models. Thus, a necessary aim of future work will be to map RNA structural intermediates with newly developed approaches such as co-transcriptional SHAPE-seq³⁵. It is also possible that a more complex RNA structural motif modulates pause longevity in the vicinity of the sequential pauses, which could not be captured in our experiments. We currently lack detailed three-dimensional structural information for the PRE RNA which further limits investigation into a potential interplay between the nascent RNA structure and pausing. Recently, a novel pause stabilized by an RNA pseudoknot was identified for the preQ₁ riboswitch, highlighting a cross-coupling between riboswitch folding and RNAP pausing³⁶. By analogy to preQ₁ riboswitch, it is possible that a complex RNA structural motif modulates pause longevity at the sequential pauses within the PRE.

Additional experiments that report on both local and global RNA structure will be required to investigate how the pH-responsive riboswitch folds co-transcriptionally in general and the specific RNA structural motifs proximal to RNAP at the sequential pauses.

To assess the effect of alkaline pH on a classical pause stabilized by nascent RNA structure, we investigated the pH-dependence of RNAP pause escape kinetics at the *his* pause – a well-characterized hairpin-stabilized pause that has served as an important model system to advance our understanding of hairpin-stabilized pauses¹¹. We performed scaffold-based experiments to characterize RNAP pause escape kinetics at the *his* pause under neutral and alkaline conditions both in the presence and absence of NusA (Fig. S3.7D). We observed that pause escape was faster at pH 8.5 compared to pH 7.2 both in the presence and absence of NusA (Fig. S3.7E), consistent with our results testing the pH-dependence of the elemental pauses relevant to PRE folding. Therefore, if a nascent RNA hairpin that modulates Pause 1 longevity was missed in the scope of our experiments, it is unlikely that the pause duration is enhanced by alkaline pH either in the presence or absence of NusA.

An alternative model for pH-dependent PRE folding. Since we observed that ECs transcribe through the PRE sequence faster at alkaline pH, one possibility is that a faster transcription rate alone is required for formation of the translationally active PRE conformer. However, this explanation is inconsistent with the prior observation that *in vitro* synthesis of the PRE by T7 RNAP at alkaline pH does not shift the RNA ensemble to the active conformer². Thus, the faster transcription rate at alkaline pH is insufficient to explain the RNA structural switch. Since T7 RNAP pauses less than *E. coli* RNAP¹⁰, this observation supports that pauses during PRE synthesis are strategically located to guide the RNA folding. Pauses in other riboswitch folding pathways have been found to guide proper folding decisions and can provide a time window for ligand binding^{26,36,37}, so it is possible that the pauses in the PRE folding pathway fit this paradigm and provide a time window for alkaline pH to act on the RNA. Also, it is intriguing that three

sequential pauses were uncovered in our experiments. Sequential pauses have been observed in leader regions that regulate *B. subtilis* gene expression^{38,39}, and were also found throughout the *B. subtilis* genome in an RNET-seq study⁴⁰. It is currently unclear what functional role sequential pauses play in the context of the PRE. Altogether, while we conclude that both Pause 1 and Pause 2 do not display the kinetics proposed in prior work, they are likely important to the ultimate RNA folding outcome at different pH.

An alternative model for PRE folding that unifies our observations with prior work is a model with elements of *both* kinetic control and a direct effect of pH on the RNA. We propose that pausing provides a temporal window for pH to exert an effect on the RNA structure and guide the biologically functional outcomes. Since the PRE alters its co-transcriptionally folded structure upon a shift from neutral to alkaline pH, it is possible that a nucleobase deprotonation event triggers the structural switch at alkaline pH. Typically, unmodified nucleosides in RNA are expected to be neutrally charged at biological pH, since the pK_a values closest to neutrality in RNA nucleobases are 3.5 for N1 of adenosine, 4.2 for N3 of cytidine, and 9.2 for both N1 of guanosine and N3 of uridine⁴¹. However, the structural environment in a folded RNA can shift these pK_a values significantly towards neutrality^{42–46}. There are many examples where upwardly shifted pK_a s of A or C residues in an RNA result in protonation under physiological conditions, which in turn modulates the RNA structural ensemble. For example, protonation of a C in a C⁺•G-C base triple ($pK_a = 6.8$) in the beet western yellow virus mRNA was found to stabilize an RNA pseudoknot critical for translational frameshifting⁴². Protonation of an A with a pK_a of 6.23 in the murine leukemia virus mRNA tunes the conformational equilibria to control the level of translational frameshifting⁴³. Examples of functionally relevant protonation events span beyond viruses and have been characterized for the U6 spliceosomal RNA⁴⁴, a bacterial group II intron⁴⁵, and a microRNA⁴⁶. The most commonly protonated base pairs are A⁺-C and C⁺-C⁴⁷, and while the PRE structural models lack such base pairs as they currently stand, many of the RNA

examples outlined above highlight the potential for more complex effects of base protonation/deprotonation events on RNA structural motifs beyond simple base pairing. Furthermore, functionally relevant protonation events for both the bacterial group II intron and microRNA noted above involve a protonated A⁺-G base pair, further expanding the scope of possibilities for pH-sensing in the PRE. Finally, a quantum chemical study that surveyed available RNA crystal structures identified eight base pairs that can exist in both protonated and non-protonated geometries and could act as pH-dependent switches⁴⁸. This study also identified a wide range of higher order RNA structural interactions where protonation likely plays a stabilizing role⁴⁸. Thus, it is possible that a deprotonation event destabilizes a particular tertiary structural element of PRE at alkaline pH, altering the energetics underlying the RNA folding and shifting the RNA ensemble towards the translationally active conformer. We currently lack detailed three-dimensional structural information for the PRE, which will be an essential aim of future work to identify specific putative protonation sites that may mediate a conformational switch. More broadly, a complete mechanistic picture for the pH-responsive riboswitch will provide important insights for the development of pH-tuned genetic control elements.

4. Experimental Procedures

Reagents and materials

All plasmids and oligonucleotides used in this study are listed in Table S3.1. Synthetic construct encoding wild-type PRE (nt 1-207) under control of its native *a/x* promoter was purchased in pUC57-Kan vector from Bio Basic. All RNA and DNA oligos for transcription experiments were obtained from Integrated DNA Technologies (IDT). RNAs were purified before use by denaturing polyacrylamide gel electrophoresis (PAGE) through 15% gels (19:1 acrylamide:bis-acrylamide) with 8 M urea and TBE buffer (89 mM Tris base pH 8.3, 89 mM boric acid, 2 mM EDTA). Full-length RNAs were visualized with UV-shadowing, excised, and incubated

overnight in low salt oligo purification buffer (10 mM Tris-Cl pH 7.5, 1 mM EDTA, 100 mM NaCl). The gel-extracted RNAs were applied to DEAE Sepharose resin (Cytiva), eluted with a high salt oligo purification buffer (10 mM Tris-Cl pH 7.5, 1 mM EDTA, 2 M NaCl), and ethanol precipitated. Linear double-stranded DNA templates for promoter-initiated *in vitro* transcription were prepared by PCR with high-fidelity Q5 DNA polymerase (New England Biolabs), treated with *DpnI* (New England Biolabs) and purified by phenol/chloroform extraction followed by ethanol precipitation prior to use. Primers for the PCR were from IDT. Initiating dinucleotide for promoter-initiated transcription (GpC) and 3'-deoxynucleotides for RNA sequencing ladder were from TriLink BioSciences. NTPs for transcription experiments were purchased as HPLC-purified from Cytiva. ³²P-labeled NTPs were from PerkinElmer.

Preparation of *E. coli* RNAP and σ^{70}

E. coli RNAP was prepared as follows. BL21(DE3) cells were transformed with pRM756 (obtained from the Landick lab, University of Wisconsin-Madison), an overexpression plasmid for wild-type *E. coli* RNAP with a PreScission Protease-cleavable His₁₀ tag on the C-terminus of the β' subunit. Cells were grown in 3x1L Luria Broth on a 37 °C shaker (200 rpm) to an OD₆₀₀ of 0.25. The temperature was dropped to 16 °C and cultures continued to grow with shaking (200 rpm) to an OD₆₀₀ of 0.4. Cultures were induced with 1 mM Isopropyl β -D-1 thiogalactopyranoside (IPTG) and grown overnight with shaking (200 rpm) at 16 °C. Cells were pelleted by centrifugation at 4 °C (4000 rpm). The cell pellet was resuspended in 30 mL of lysis buffer (50 mM Tris-Cl, pH 7.9, 5% glycerol, 233 mM NaCl, 2 mM EDTA, 10 mM 2-mercaptoethanol, 10 mM DTT, protease inhibitor cocktail (Sigma Aldrich)) per 2 L of culture and sonicated for a total of 30 minutes (20 s on/ 20 s off, amplitude 60%) at 4 °C. All subsequent purification steps were carried out at 4 °C. The lysate was spun at 11,000g for 15 min. Cleared supernatant was subjected to polyethyleneimine (PEI) precipitation by slow addition of PEI to 0.6% v/v with gentle stirring. After

stirring for 10 min, the precipitate was recovered by centrifugation at 11,000g for 15 min. The pellet was resuspended in 1x TGEDZ (10 mM Tris-Cl, pH 7.9, 5% glycerol, 0.1 mM EDTA, 5 μ M ZnCl₂) + 0.5 M NaCl to salt-in unwanted proteins and the suspension was centrifuged at 11,000g for 15 min. The pellet was resuspended in 1x TGEDZ + 1M NaCl to salt-in RNAP and the suspension was centrifuged at 11,000g for 15 min. The supernatant was removed and subjected to ammonium sulfate precipitation. Ammonium sulfate was added to a final concentration of 0.37 g/L, and the solution was left stirring overnight. The solution was centrifuged at 27,000g for 15 min. The pellet was resuspended in HisTrap Buffer A (20 mM Tris-Cl pH 7.9, 500 mM NaCl, 5 mM imidazole, 5 mM 2-mercaptoethanol), filtered with a 0.2 μ M filter to remove particulates and applied to a GE HisTrap HP 5 mL column. RNAP was eluted over 10 column volumes with a gradient of 5 mM – 250 mM imidazole. Fractions containing RNAP were identified via SDS-PAGE and pooled. The NaCl concentration in the pooled fractions was reduced to 200 mM with addition of 1x TGEDZ. Pooled fractions were then loaded onto a GE HiTrap Heparin HP 5 mL column to compete nucleic acids off of RNAP. RNAP was eluted via a step elution by increasing the NaCl concentration from 200 mM to 500 mM. Fractions containing RNAP were identified via SDS-PAGE, pooled, and concentrated to a final volume of 2 mL with an Amicon Ultra Centrifugal Filter (MWCO 100 kDa). The concentrate was loaded into a dialysis cassette (ThermoFisher) and dialyzed into storage buffer (10 mM Tris-Cl, pH 7.9, 25% glycerol, 100 mM NaCl, 100 μ M EDTA, 1 mM MgCl₂, 20 μ M ZnCl₂, 10 mM DTT) with stirring overnight. Small aliquots of purified RNAP were transferred into microcentrifuge tubes and stored at -80 °C until use.

Wild-type *E. coli* σ^{70} was prepared as follows. T7 Express cells were transformed with pMIS4 (originally obtained from Burgess lab, University of Wisconsin-Madison), an overexpression plasmid for wild-type *E. coli* σ^{70} with an N-terminal His₆ tag. Cultures were grown on a shaker (200 rpm) at 37 °C to an OD₆₀₀ of 0.5 and induced by addition of IPTG to 0.4 mM.

The temperature was dropped to 16 °C and cells were grown with shaking (200 rpm) overnight. Cells were harvested by centrifugation at 3,000 g for 15 min at 4 °C. All subsequent purification steps were carried out at 4 °C. The cell pellet was resuspended in lysis buffer (50 mM Tris-Cl, pH 7.9, 5 mM imidazole, 5% glycerol, 233 mM NaCl, 2 mM EDTA, 10 mM 2-mercaptoethanol, protease inhibitor cocktail (Sigma Aldrich)). The cells were lysed via sonication. The lysate was centrifuged at 11,000g for 15 min. The cleared supernatant was filtered and applied to a GE HisTrap HP 5 mL column. σ^{70} was eluted over 6 column volumes via a two-step elution: the first step was performed over 3 column volumes at 300 mM imidazole and the second step was performed over 3 column volumes at 500 mM imidazole. The fractions containing σ^{70} were identified via SDS-PAGE and pooled. The pooled fractions were run on a HiLoad 16/60 Superdex 200 preparative grade size exclusion column (Cytiva). The fractions containing σ^{70} were identified via SDS-PAGE and pooled. The pooled fractions were dialyzed overnight into a storage buffer (50 mM Tris pH 7.9, 250 mM NaCl, 30% glycerol). Small aliquots of purified σ^{70} were transferred into microcentrifuge tubes and stored at -80 °C.

Preparation of NusA, GreA, and GreB

NusA was prepared as follows. BL21(DE3) cells were transformed with pRA46, an overexpression plasmid for wild-type NusA with a His₆ tag on its N-terminus. Cultures were grown on a shaker (200 rpm) at 37 °C to an OD₆₀₀ of 0.5 and induced by addition of IPTG to 1 mM. The temperature was dropped to 16 °C and cells were grown with shaking (200 rpm) overnight. Cells were harvested by centrifugation at 5000 rpm for 15 min at 4 °C. All subsequent purification steps were carried out at 4 °C. The cell pellet was resuspended in lysis buffer (50 mM Tris-Cl, pH 8.0, 10 mM imidazole, 500 mM NaCl, 2 mM 2-mercaptoethanol, 0.1 mM PMSF, 1 mM benzamidine, 2 μ g/mL DNase I, protease inhibitor cocktail (Sigma Aldrich)). The cells were lysed via sonication. The lysate was centrifuged at 13,000 g for 30 min. The cleared supernatant was applied to a GE

HisTrap HP 5 mL column. NusA was eluted via a step elution over 6 column volumes with 500 mM imidazole. The fractions containing NusA were identified via SDS-PAGE and pooled. The pooled fractions were run on a HiLoad 16/60 Superdex 200 preparative grade size exclusion column (Cytiva). The fractions containing NusA were identified via SDS-PAGE and pooled. The pooled fractions were dialyzed overnight into a storage buffer (20 mM Tris, pH 8.0, 200 mM NaCl, 1 mM EDTA, 1 mM DTT, 30% glycerol, 0.1 % Triton-X-100). Small aliquots of purified NusA were transferred into microcentrifuge tubes, flash frozen, and stored at -80 °C.

GreB was prepared as follows. BL21(DE3) cells were transformed with pIA577 (obtained from the Landick lab, University of Wisconsin-Madison), an overexpression plasmid for wild-type GreB with a His₆ tag on its C-terminus. Cultures were grown on a shaker (200 rpm) at 37 °C to an OD₆₀₀ of 0.5 and induced by addition of IPTG to 1 mM. The temperature was dropped to 16 °C and cells were grown with shaking (200 rpm) overnight. Cells were harvested by centrifugation at 5,000 rpm for 15 min at 4 °C. All subsequent purification steps were carried out at 4 °C. The cell pellet was resuspended in lysis buffer (50 mM NaPO₄ pH 8.0, 5 mM imidazole, 10% glycerol, 250 mM NaCl, protease inhibitor cocktail (Sigma Aldrich)). The cells were lysed via sonication. The lysate was centrifuged at 11,000 g for 15 min. The cleared supernatant was filtered and applied to a GE HisTrap HP 5 mL column. GreB was eluted over 7 column volumes with a gradient of 25-300 mM imidazole. The fractions containing GreB were identified via SDS-PAGE and pooled. The pooled fractions were dialyzed overnight into a storage buffer (10 mM Tris-Cl pH 8.0, 50% glycerol, 0.1 mM EDTA, 250 mM NaCl, 2 mM DTT). Small aliquots of purified GreB were transferred into microcentrifuge tubes and stored at -80 °C.

GreA was prepared as follows. BL21(DE3) cells were transformed with pIA578 (obtained from the Landick lab, University of Wisconsin-Madison), an overexpression plasmid for wild-type GreA with a His₆ tag on its C-terminus. Cultures were grown on a shaker (200 rpm) at 37 °C to an OD₆₀₀ of 0.5 and induced by addition of IPTG to 1 mM. The temperature was dropped to 16 °C

and cells were grown with shaking (200 rpm) overnight. Cells were harvested by centrifugation at 5,000 rpm for 15 min at 4 °C. All subsequent purification steps were carried out at 4 °C. The cell pellet was resuspended in lysis buffer (100 mM Tris-Cl pH 8.0, 100 mM NaCl, 10 mM EDTA, 40 μ M PMSF, 3 mg/mL lysozyme). The cells were lysed via sonication. The lysate was centrifuged at 11,000g for 15 min. The cleared supernatant was filtered and applied to a GE HisTrap HP 5 mL column. GreA was eluted over 7 column volumes with a gradient of 25-300 mM imidazole. The fractions containing GreA were identified via SDS-PAGE and pooled. The pooled fractions were dialyzed overnight into a storage buffer (10 mM Tris-Cl pH 8.0, 50% glycerol, 0.1 mM EDTA, 250 mM NaCl, 2 mM DTT). Small aliquots of purified GreA were transferred into microcentrifuge tubes and stored at -80 °C.

Synchronized single-round *in vitro* transcription from a promoter

To monitor transcription on the PRE sequence when initiated from the *alx* promoter, we performed synchronized, single-round *in vitro* transcription assays. The DNA template for the *in vitro* transcription reactions was prepared via PCR using pMIS301 as a plasmid template. The transcription template contains the *alx* promoter, full PRE sequence, and a portion of the downstream gene (nt 1-59) termed *alx'*. Both the oligos and plasmid used for DNA template preparation are outlined in Table S3.1.

To prepare the RNAP holoenzyme, *E. coli* RNAP (2 μ M) and σ^{70} (4 μ M) were combined in 1X transcription buffer (20 mM Tris-Cl pH 7.2 or 8.5, 150 mM KCl, 10 mM MgCl₂, 2.77 mM DTT) and incubated for 30 min at 37 °C. Following HoloRNAP assembly, *in vitro* transcription reactions were initiated and synchronized, taking advantage of the natural lack of cytosines between positions 4-15 of the PRE. The initiation mix for *in vitro* transcription thus lacked CTP and contained 0.1 μ M of the HoloRNAP, 0.08 μ M of the DNA template, 150 μ M GpC dinucleotide, 25 μ M UTP, 25 μ M GTP, 10 μ M ATP, [α -³²P]ATP, and 1X transcription buffer at the appropriate pH. After allowing initiation to proceed for 15 min at 37 °C, rifampicin (Fisher) was added to a final

concentration of 10 $\mu\text{g/mL}$, to inhibit re-initiation. Completion of this step yielded ECs stalled at the A15 position on the PRE.

Next, transcription elongation reactions were performed by combining 7.5 μL of ECs with 2.5 μL of an NTP mix containing 1 mM each ATP, UTP, CTP, and GTP (250 μM final, each NTP). All elongation reactions were performed at 37 °C. The final concentration of all NTPs in the transcription reactions was chosen to match the conditions used by Nechooshtan *et. al.*² Transcription factors were introduced concurrently with NTPs to test their effect on transcription elongation and pausing. Specifically, to test the effect of NusA on transcription elongation, NTP mixes were prepared following the recipe outlined above and NusA was included at concentrations of 200 nM, 400 nM, 1 μM , and 2 μM such that its final concentrations were 50 nM, 100 nM, 250 nM, and 500 nM in the experiment. To test the effect of GreA on transcription elongation, an NTP mix was prepared following the recipe outlined above and GreA was included at a concentration of 2 μM such that its final concentration was 500 nM in the experiment. After beginning elongation, the reaction was quenched at desired time points with an equal volume of 2X stop buffer (90 mM Tris-borate buffer, 8 M urea, 50 mM EDTA, and 0.02% of both xylene cyanol and bromophenol blue). The time points collected were: 4 s, 6 s, 8 s, 10 s, 12 s, 14 s, 20 s, 30 s, 45 s, 1 min, 1.5 min, and 2 min. Additionally, chase reactions were performed at both pH 7.2 and 8.5 with each NTP present at a final concentration of 500 μM . Radiolabeled RNAs in the quenched samples were denatured for 2 min at 95 °C and separated using denaturing PAGE (10% 19:1 acrylamide:bis-acrylamide, 8 M urea). Samples were run alongside a radiolabeled ladder (*MspI*-digested pBR322 plasmid). The gel was exposed to a PhosphorImager screen, and the screen was scanned using a Typhoon PhosphorImager.

Pause mapping using synchronized single-round *in vitro* transcription from a promoter

To better visualize the sequential pauses on the PRE when initiating transcription from the *a/x* promoter, a similar protocol was used to the one described above with a few modifications. The 1X transcription buffer contained 20 mM Tris-Cl pH 8.5², 150 mM KCl, 5 mM MgCl₂, and 5 mM DTT. The NTP concentrations used for the transcription reactions were lowered to slow pause escape and thus better visualize pause RNA species on denaturing PAGE gels. First, a set of transcription elongation reactions was performed with ATP, UTP, CTP, and GTP at a final concentration of 100 μ M. The reactions were quenched at 5 s, 10 s, 20 s, and 30 s. Additional transcription elongation reactions were executed with one NTP limiting, to extend the lifetime of each sequential pause. Specifically, three sets of reactions were performed with either ATP, UTP, or GTP limiting at a concentration of 20 μ M, with the other three NTPs at a concentration of 100 μ M. The reactions were quenched at 10 s, 20 s, 30 s, and 40 s. In parallel with the transcription elongation reactions, RNA sequencing ladders were generated to map the pauses prior to termination site B. The ECs were combined with ATP, UTP, CTP and GTP to 100 μ M and one 3'-deoxyNTP (TriLink BioTechnologies) to 500 μ M. The sequencing reactions proceeded for 20 min at 37 °C and were quenched with an equal volume of 2X stop buffer. All sequencing reactions were performed at pH 8.5. Radiolabeled RNA in the quenched samples were heat-denatured for 2 min at 95 °C and separated using denaturing PAGE (10% 19:1 acrylamide:bis-acrylamide, 8 M urea). Samples were run alongside a radiolabeled pBR322 ladder. A 10% denaturing PAGE gel was chosen to best resolve the sequential pause RNA species.

Pause mapping using scaffold-based transcription

To precisely map the pauses downstream of Hairpin S in the PRE, a transcription elongation assay using a nucleic acid scaffold was performed. The nucleic acid scaffold initiated transcription upstream of the sequence window where pauses were expected to occur. PAGE-

purified 19-mer RNA was 5'-³²P labelled. The labeling reaction contained 20 μ M RNA, T4 Polynucleotide Kinase (PNK, New England Biolabs), [γ -³²P]ATP, 100 μ M ATP, and 1X PNK buffer. PNK was then heat-inactivated for 3 min at 95 °C, and the 5'-³²P labelled RNA was purified with an Illustra Microspin G50 column (Cytiva) to remove unincorporated ATP. The nucleic acid scaffold was annealed in 1X transcription buffer (20 mM Tris-Cl pH 8.2, 150 mM KCl), with 5 μ M 5'-³²P labelled RNA and 10 μ M template DNA (T-DNA).

To reconstitute ECs, RNAP was combined with the nucleic acid scaffold in 1X transcription buffer and incubated for 15 min at 37 °C, then the non-template DNA (NT-DNA) was added and incubation continued for an additional 10 min at 37 °C. The ratio of RNA:T-DNA:NT-DNA:RNAP was 1:2:5:3, corresponding to concentrations of 0.5 μ M, 1 μ M, 2.5 μ M, and 1.5 μ M, respectively. Next, the ECs were diluted 2-fold in 1X transcription buffer and applied to G50 resin, to remove unbound scaffold. Heparin (Fisher) was added to the ECs at a final concentration of 0.1 mg/mL to bind any free RNAP. ECs were incubated with heparin for 3 min at 37 °C before starting transcription elongation reactions. Reactions were initiated by adding ATP, UTP, CTP and GTP to a final concentration of 100 μ M, each. To increase pause duration, additional reactions were performed with either ATP, UTP, or GTP limiting at a concentration of 10 μ M, with the other three NTPs at a concentration of 100 μ M. MgCl₂ was added to a final concentration of 5 mM alongside NTP introduction. Samples were removed at time points ranging from 10 s to 80 s and quenched with an equal volume of 2X stop buffer. All remaining active ECs were chased off of the DNA scaffold by incubation with NTPs to a final concentration of 1 mM for 5 min at 37 °C.

In parallel with the transcription elongation reactions, RNA sequencing ladders were generated for precise mapping of pause locations. The ECs were prepared as described above. The ECs were combined with ATP, UTP, CTP, and GTP to 100 μ M and one 3'-deoxyNTP (TriLink BioTechnologies) to 1 mM. The sequencing reactions proceeded for 20 min at 37 °C and were

quenched with an equal volume of 2X stop buffer. All sequencing reactions were performed at pH 8.2.

Radiolabeled RNAs from both the quenched time points and sequencing reactions were heat-denatured for 2 min at 95 °C and separated using denaturing PAGE (15% 19:1 acrylamide:bis-acrylamide, 8 M urea). The gel was exposed to a PhosphorImager screen, and the screen was scanned using a Typhoon PhosphorImager.

Sequential pause escape kinetics positioning RNAP directly at the pauses

To measure the pause escape rate at each of the sequential pauses, PAGE-purified 17-mer RNA was 5'-³²P labelled using the same protocol implemented for the pause mapping assays (see above). The nucleic acid scaffolds for each pause in the sequential series (positions 142-144 on the PRE sequence) were annealed in a buffer (20 mM Tris pH 7.2 and 150 mM KCl), with 5 μ M 5'-³²P labelled RNA and 10 μ M template DNA (T-DNA). The template and non-template DNAs used for these experiments included additional downstream DNA sequence compared to the DNA oligos used for pause mapping experiments, to ensure that lost RNAP-DNA contacts due to shortening downstream DNA during transcription would not introduce artifacts into the measurements of pause lifetime.

To reconstitute ECs at the first pause, RNAP was combined with the nucleic acid scaffold in 1X transcription buffer (20 mM Tris-Cl pH 7.2 or 8.5, 150 mM KCl, 5 mM MgCl₂, and 5 mM DTT) and incubated for 15 min at 37 °C, then the non-template DNA (NT-DNA) was added and incubation continued for 10 min at 37 °C. The ratio of RNA:T-DNA:NT-DNA:RNAP was 1:2:3:3, corresponding to concentrations of 0.5 μ M, 1 μ M, 1.5 μ M, and 1.5 μ M, respectively. To reconstitute ECs at the second and third pauses, the same recipe as above was used but with Mg omitted, to prevent intrinsic cleavage at the RNA 3' end prior to initiating transcription. Next, the ECs were diluted 2-fold in 1X transcription buffer and applied to an Illustra Microspin G50

column (Cytiva), to remove unbound scaffold. Heparin was added to the ECs at a final concentration of 0.1 mg/mL to sequester free RNAP. ECs were incubated with heparin for 3 min at 37 °C before starting the transcription reactions.

For the first pause in the sequential series, ATP was added to the PECs at a final concentration of 10 μ M. Samples were removed at various time points and quenched with an equal volume of 2X stop buffer. All remaining active PECs were chased by incubation with ATP to a final concentration of 500 μ M for 5 min at 37 °C.

For the second pause in the sequential series, UTP was added to the PECs at a final concentration of 10 μ M. $MgCl_2$ was added to a final concentration of 5 mM alongside NTP introduction. Samples were removed at various time points and quenched with an equal volume of 2X stop buffer. All remaining active PECs were chased by incubation with UTP to a final concentration of 500 μ M for 5 min at 37 °C.

For the third pause in the sequential series, GTP was added to the PECs at a final concentration of 20 μ M. $MgCl_2$ was added to a final concentration of 5 mM alongside NTP introduction. Samples were removed at various time points and quenched with an equal volume of 2X stop buffer. All remaining active PECs were chased to product by incubation with GTP to a final concentration of 500 μ M for 5 min at 37 °C.

Radiolabeled RNAs from the quenched time points were heat-denatured for 2 min at 95 °C and separated using denaturing PAGE (15% 19:1 acrylamide:bis-acrylamide, 8 M urea was used for first and third pause in the sequential series, 20% 19:1 acrylamide:bis-acrylamide, 8 M urea was used for the second pause in the sequential series). The gel was exposed to a PhosphorImager screen, and the screen was scanned using a Typhoon PhosphorImager. The gel images were quantified using ImageQuant software (GE Healthcare). For measurement of pause lifetimes, the average fraction of RNA at the pause was quantified for each time point. All

measurements of the average pause RNA fraction report the average and standard deviation of time points performed in triplicate.

In addition to the sequential pause escape rate measurements, the Pause 1 location reported by Nechooshtan *et. al.* (nt 138 on the PRE sequence)² was tested with a scaffold-based transcription assay. The general protocol was the same as that used for the sequential pauses described above. The experiments were conducted at pH 7.2 and 8.2 to match the conditions used in the original analysis of the PRE pauses². The nucleic acid scaffold used did not preserve the -11 position of the putative pause. In the wild-type PRE sequence, there is a C at the -11 position in the NT-DNA; the NT-DNA in our experiments had a T at this position instead. However, based on published work probing the contribution of the -11 position to overall pause strength, this substitution would only weaken the pause signal, not abolish it altogether as observed in our experiment (Fig. S3.3A)²⁸.

Testing for a backtracked register with ECs positioned at pause 144

To test for a backtracked register at the third pause, ECs were reconstituted directly at the pause following the same protocol outlined above to measure the GMP incorporation rate at different pH. ECs were reconstituted at both pH 7.2 and 8.5 in the presence or absence of 5 mM MgCl₂. Following reconstitution, ECs for each pH and Mg condition were incubated for 10 min at 37 °C in the presence or absence of 1 μ M GreA or GreB. The reactions were quenched with an equal volume of 2X stop buffer. Radiolabeled RNAs from the quenched samples were heat-denatured for 2 min at 95 °C and separated using denaturing PAGE (15% 19:1 acrylamide:bis-acrylamide, 8 M urea). The gel was exposed to a PhosphorImager screen, and the screen was scanned using a Typhoon PhosphorImager. The gel image was quantified using ImageQuant software (GE Healthcare).

Sequential pause kinetics initiating transcription upstream of the pauses

To further investigate the pause escape kinetics at the sequential pauses, a scaffold-based transcription elongation assay was designed to initiate transcription elongation upstream of the sequential pauses. A 10-mer RNA was used to position RNAP upstream of the sequential pauses, with the 3' G in the RNA corresponding to position 136 on the PRE sequence. The nucleic acid scaffold was annealed in a buffer (20 mM Tris pH 7.2 and 150 mM KCl), with 5 μ M RNA and 10 μ M template DNA (T-DNA). To reconstitute the G10 ECs, RNAP was combined with the nucleic acid scaffold in 1X transcription buffer (20 mM Tris pH 7.2 or 8.5, 150 mM KCl, and 5 mM DTT) and incubated for 15 min at 37 °C, then the non-template DNA (NT-DNA) was added and incubation continued for 10 min at 37 °C. The ratio of RNA:T-DNA:NT-DNA:RNAP was 1:2:5:3 corresponding to concentrations of 0.5 μ M, 1 μ M, 2.5 μ M, 1.5 μ M, respectively. Next, the ECs were diluted 5-fold and heparin was added to the ECs at a final concentration of 0.1 mg/mL to sequester free RNAP. ECs were incubated with heparin for 3 min at 37 °C. To radiolabel the ECs, [α -³²P]ATP was added to the ECs for 1 min at 37 °C, which extended the RNAs by 1 nt to form A11 ECs. To extend all ECs to A11, ECs were incubated with ATP at a final concentration of 2 μ M for 3 min at 37 °C.

Multiple sets of transcription elongation reactions were performed to increase the duration of each pause in the sequential series and test the effects of GreA and GreB on pause escape at different pH. For all reactions, MgCl₂ was introduced at a concentration of 5 mM alongside the incoming NTPs. To assess the effect of pH and Gre factors on pause escape kinetics at each sequential pause, either ATP, UTP, or GTP was added at limiting concentrations of 10 μ M, 10 μ M, and 20 μ M, respectively. All other NTPs were added at a concentration of 100 μ M. GreA or GreB was added to a final concentration of 1 μ M alongside NTP introduction. Samples were removed at various time points and quenched with an equal volume of 2X stop buffer. All

remaining active PECs were chased to product by incubation with NTPs to a final concentration of 500 μ M for 5 min at 37 °C.

Radiolabeled RNAs from the quenched time points were heat-denatured for 2 min at 95 °C and separated using denaturing PAGE (15% 19:1 acrylamide:bis-acrylamide, 8 M urea). The gel was exposed to a PhosphorImager screen, and the screen was scanned using a Typhoon PhosphorImager. The gel images were quantified using ImageQuant software (GE Healthcare). For measurement of pause lifetimes, the average fraction of RNA at the pause was quantified for each time point. All measurements of the average pause RNA fraction report the average and standard deviation of time points performed in triplicate.

Pause 2 mapping and kinetics

To both map and investigate the pause escape kinetics at Pause 2, a scaffold-based transcription elongation assay was designed to initiate transcription elongation upstream of the termination site D. A 10-mer RNA was used to position RNAP upstream of the sequence region where Pause 2 occurs, with the 3' G in the RNA corresponding to position 173 on the PRE sequence. The nucleic acid scaffold was annealed in a buffer (20 mM Tris pH 7.2 and 150 mM KCl), with 5 μ M RNA and 10 μ M template DNA (T-DNA). To reconstitute the G10 ECs, RNAP was combined with the nucleic acid scaffold in 1X transcription buffer (20 mM Tris pH 7.2 or 8.5, 150 mM KCl, and 5 mM DTT) and incubated for 15 min at 37 °C, then the non-template DNA (NT-DNA) was added and incubation continued for 10 min at 37 °C. The ratio of RNA:T-DNA:NT-DNA:RNAP was 1:2:5:3 corresponding to concentrations of 0.5 μ M, 1 μ M, 2.5 μ M, 1.5 μ M, respectively. Next, the ECs were diluted 5-fold and heparin was added to the ECs at a final concentration of 0.1 mg/mL to sequester free RNAP. ECs were incubated with heparin for 3 min at 37 °C. To radiolabel the ECs, [α -³²P]ATP was added for 1 min at 37 °C, which extended the

RNAs by 1 nt to form A11 ECs. To extend all ECs to A11, ECs were incubated with ATP at a final concentration of 2 μ M for 3 min at 37 °C.

Reactions were initiated by adding ATP, UTP, CTP, and GTP to a final concentration of 100 μ M, each. To increase pause duration, additional reactions were performed with UTP or GTP at a limiting concentration of 10 μ M, with the other three NTPs at a concentration of 100 μ M. MgCl₂ was added to a final concentration of 5 mM alongside NTP introduction. Samples were removed at various time points and quenched with an equal volume of 2X stop buffer. All remaining active PECs were chased to product by incubation with NTPs to a final concentration of 1 mM for 5 min at 37 °C.

In parallel with the transcription elongation reactions, RNA sequencing ladders were generated for precise mapping of pause locations. The ECs were prepared as described above. The ECs were combined with ATP, UTP, CTP, and GTP to 100 μ M and one 3'-deoxyNTP to 1 mM. The sequencing reactions proceeded for 20 min at 37 °C and were quenched with an equal volume of 2X stop buffer. All sequencing reactions were performed at pH 8.5.

Radiolabeled RNAs from both the quenched time points and sequencing reactions were heat-denatured for 2 min at 95 °C and separated using denaturing PAGE (15% 19:1 acrylamide:bis-acrylamide, 8 M urea). The gel was exposed to a PhosphorImager screen, and the screen was scanned using a Typhoon PhosphorImager. The gel images were quantified using ImageQuant software (GE Healthcare). For measurement of pause lifetimes, the average fraction of RNA at the pause was quantified for each time point. All measurements of the average pause RNA fraction report the average and standard deviation of time points performed in triplicate.

***His* pause kinetics at different pH**

To investigate the pH-dependence of pause escape at the *his* pause, PAGE-purified 27-mer RNA was 5'-³²P labelled using the same protocol implemented for the pause mapping assays

(see above). The RNA positions ECs 2 nt upstream of the *his* pause site. The nucleic acid scaffold was annealed in a buffer (20 mM Tris pH 7.2 and 150 mM KCl), with 5 μ M 5'-³²P labelled RNA and 10 μ M template DNA (T-DNA). To reconstitute the G27 ECs, RNAP was combined with the nucleic acid scaffold in 1X transcription buffer (20 mM Tris-Cl pH 7.2 or 8.5, 150 mM KCl, 5 mM MgCl₂, and 5 mM DTT) and incubated for 15 min at 37 °C, then the non-template DNA (NT-DNA) was added and incubation continued for 10 min at 37 °C. The ratio of RNA:T-DNA:NT-DNA:RNAP was 1:2:5:3, corresponding to concentrations of 0.5 μ M, 1 μ M, 2.5 μ M, and 1.5 μ M, respectively. The ECs were applied to an Illustra Microspin G50 column (Cytiva), to remove unbound scaffold. Heparin was added to the ECs at a final concentration of 0.1 mg/mL to sequester free RNAP. ECs were incubated with heparin for 3 min at 37 °C. Next, the G27 ECs were diluted 2-fold and walked from G27 to C28 by addition of 10 μ M CTP. The ECs were incubated with CTP for 1 min at 37 °C. The C28 ECs were diluted 2-fold and split, with half of the ECs undergoing incubation with NusA. For a subset of the C28 ECs, NusA was added to a final concentration of 4 μ M (for a final concentration of 2 μ M in the transcription reactions). The C28 ECs were incubated with NusA for 5 min at 37 °C. To initiate the pause escape measurements UTP and GTP were added to the C28 ECs. UTP was added to a concentration of 100 μ M and GTP was added at a limiting concentration of 10 μ M to increase pause duration. Samples were removed at various time points and quenched with an equal volume of 2X stop buffer. All remaining active PECs were chased to product by incubation with UTP and GTP to a final concentration of 500 μ M for 5 min at 37 °C.

Radiolabeled RNAs from the quenched time points were heat-denatured for 2 min at 95 °C and separated using denaturing PAGE (15% 19:1 acrylamide:bis-acrylamide, 8 M urea). The gel was exposed to a PhosphorImager screen, and the screen was scanned using a Typhoon PhosphorImager. The gel images were quantified using ImageQuant software (GE Healthcare). For measurement of pause lifetimes, the average fraction of RNA at the pause was quantified for

each time point. Three replicates were performed for all time courses except for the pH 8.5 - NusA condition, which only had two replicates.

Quantification of average transcription rate on the PRE

The gels were quantified with ImageJ software, and densitometry plots for both the pBR322 ladder standards and time course samples were generated. A calibration curve was created by plotting the relative migration distance vs. the log of transcript length for the pBR322 ladder. A linear regression was performed on this plot, generating a function to estimate transcript lengths for the time course samples based on migration distance. For each time point, the average transcript length was determined by summing the product of each transcript length and its signal intensity and dividing the sum by total intensity. Elongation rates were determined by dividing the average transcript length by the reaction time. Rates for three independent samples were determined at both pH 7.2 and 8.5.

Quantification of pause kinetics and termination efficiency

For promoter-initiated experiments, pause kinetics were calculated as follows. The gel images were quantified using ImageQuant software (Cytiva). The band intensity for the pause RNA was divided by the total RNA, with a correction applied for major termination events prior to the pause as previously described¹⁶. A semilogarithmic plot was generated for the average pause RNA fraction over time and a linear regression was performed on a range of time points, starting where continued arrival of RNAP at the pause was negligible and ending before the semilogarithmic plot deviated from linearity (Pause 1: 10 s – 20 s points for pH 7.2 and 8 s – 12 s for pH 8.5, Pause 2: 14 s – 20 s for pH 7.2 and 8 s – 12 s for pH 8.5). Pause kinetic parameters were obtained from this regression analysis, following a published protocol for determination of pause kinetic parameters¹⁵. All measurements of the average pause RNA fraction report the average and standard deviation of time points performed in triplicate.

To quantify the degree of termination at termination sites B and D in promoter-initiated transcription experiments, the band intensity for the termination RNA was divided by the total RNA for each time individual time point. The fraction of terminated RNA was plotted over time to visualize the percentage of ECs that dissociate at termination sites B and D at both pH 7.2 and pH 8.5

Pausing kinetics for scaffold transcription experiments were calculated as follows. The gel images were quantified using ImageQuant software (Cytiva). A semilogarithmic plot was generated for the pause RNA fraction over time (t), and a single-exponential decay fitting was performed on the plot. Pause kinetic parameters were obtained from this analysis, according to the following equation, for individual time courses (n=3):

$$\text{Fraction of RNA at the pause} = (\text{Pause efficiency}) \cdot e^{-kt}$$

Specifically, pause lifetime (or dwell time, t) was determined as a reciprocal of the rate constant (k) for pause RNA disappearance. The reported dwell time and pause efficiency (f) for each pause is the average of individual parameters determined from fits of three individual time courses, and their standard deviation.

The pause kinetics for pause 3 in the experiment outlined in Fig. 3.4C were not obtained with a single-exponential decay fitting due to a poor fit. Instead, the average pause dwell time was calculated using the alternative method outlined above that was applied to analysis of promoter-initiated transcription experiments.

Supporting information

This article contains Supporting Information (Figs. S3.1-S3.8, Tables S3.1-S3.4).

Acknowledgements

We thank the Landick lab at the University of Wisconsin-Madison for generously supplying us with expression plasmids for wild-type *E. coli* RNAP, σ^{70} , NusA, GreA, and GreB. Additionally, we thank the former Mishanina lab members, Xing Quan, for purifying σ^{70} , and Tommy Le for assisting with the purification of GreA and GreB. We thank Dr. Ravish Sharma for purifying NusA, An Hsieh, Dr. Gairika Ghosh, and Rushabh Bhakta for assisting with wild-type *E. coli* RNAP purifications, and members of our research group for their valuable feedback during the course of this study. Finally, we are grateful to Drs. Robert Landick, Daniel Roston, and Eric Strobel for critical reading of the manuscript and helpful suggestions for improvement.

All cartoons depicting the pH-responsive riboswitch RNA were prepared using BioRender. PRE secondary structure models were generated using forna⁴⁹. This work was supported by NIH/NIGMS (ESI) R35 GM142785-01, Burroughs Wellcome Fund CASI (1016945.01), UCSD institutional support and Yinan Wang Memorial Chancellor's Endowed Junior Faculty Fellowship to T.V.M., and NIH Molecular Biophysics Training Grant (T32 GM008326) to C.S.

Data availability

Raw data for all experiments described in this manuscript is available upon request to the corresponding author.

4. Figures

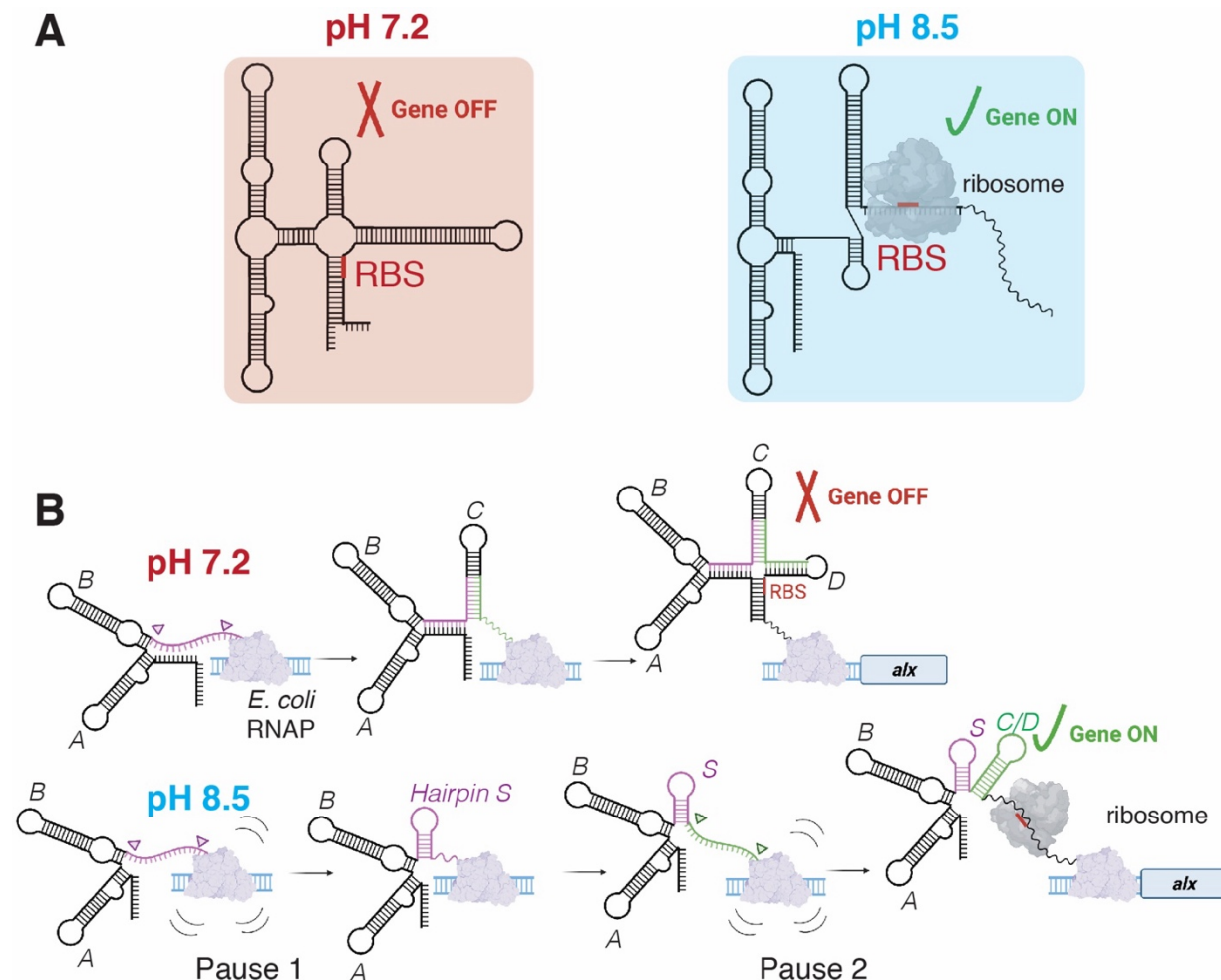


Figure 3.1 Alternative folding outcomes of the pH-responsive riboswitch.

A. Overview of the pH-responsive riboswitch folding outcomes under different pH conditions. When the PRE is synthesized under conditions of alkaline pH, a translationally active structure is formed with the ribosome binding site (RBS) available for ribosome binding and subsequent translation, turning on the downstream gene (*alx*). When the PRE is synthesized under conditions of neutral pH, a translationally inactive structure is formed with the RBS occluded, turning off the *alx* gene. The PRE structure models presented here are based on previously published structure predictions, which were generated via chemical probing with dimethyl sulfate (DMS) followed by primer extension². **B.** Proposed folding pathways for the pH-responsive riboswitch under both alkaline and neutral pH. The folding pathways shown illustrate the PRE folding model proposed by Nechooshtan *et al.*². Under alkaline conditions, it was proposed that RNAP pause longevity is enhanced at two key sites to promote folding of specific structural elements in the active conformer (Hairpin S and Hairpin C/D, shown in purple and green, respectively). The colored arrows denote the RNA portions that alternatively base pair to form Hairpin S and Hairpin C/D in the active conformer.

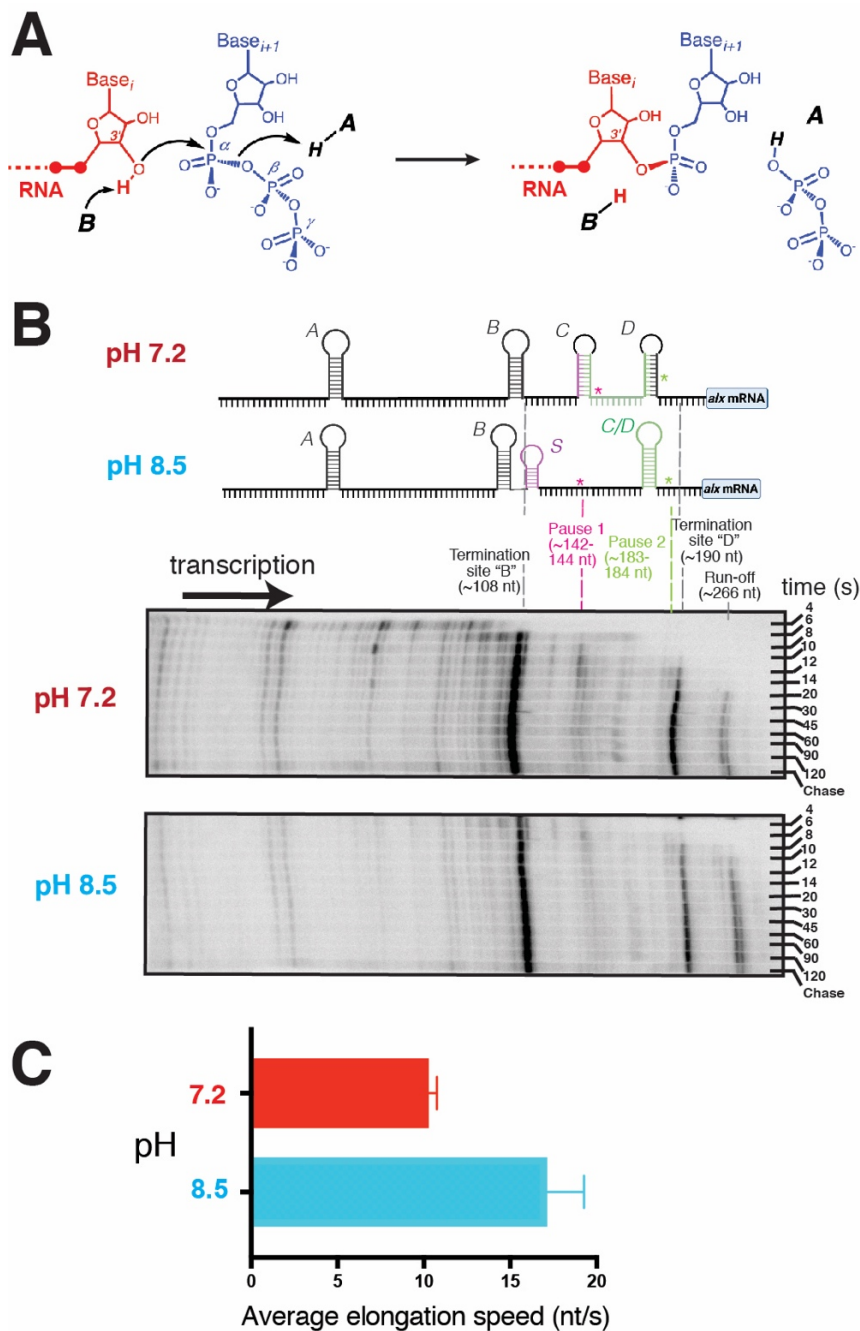


Figure 3.2 RNAP transcription through the PRE at neutral vs. alkaline pH.

A. Chemical mechanism of nucleotide addition by RNAP. The flow of electrons in the reaction is marked with arrows. Proton abstraction and donation are performed by the putative base (B) and acid (A), respectively. **B.** Synchronized single-round *in vitro* transcription from the *alx* promoter upstream of the PRE. Transcription progress was monitored at pH 7.2 and 8.5. RNA transcripts were radiolabeled via incorporation of [α - 32 P]ATP and visualized on a 10% denaturing polyacrylamide gel. Prominent RNA species are labelled with approximate transcript lengths. **C.** Average transcription rate of *E. coli* RNAP on the PRE at pH 7.2 vs. pH 8.5. Error shown is s.d. of a triplicate measurement.

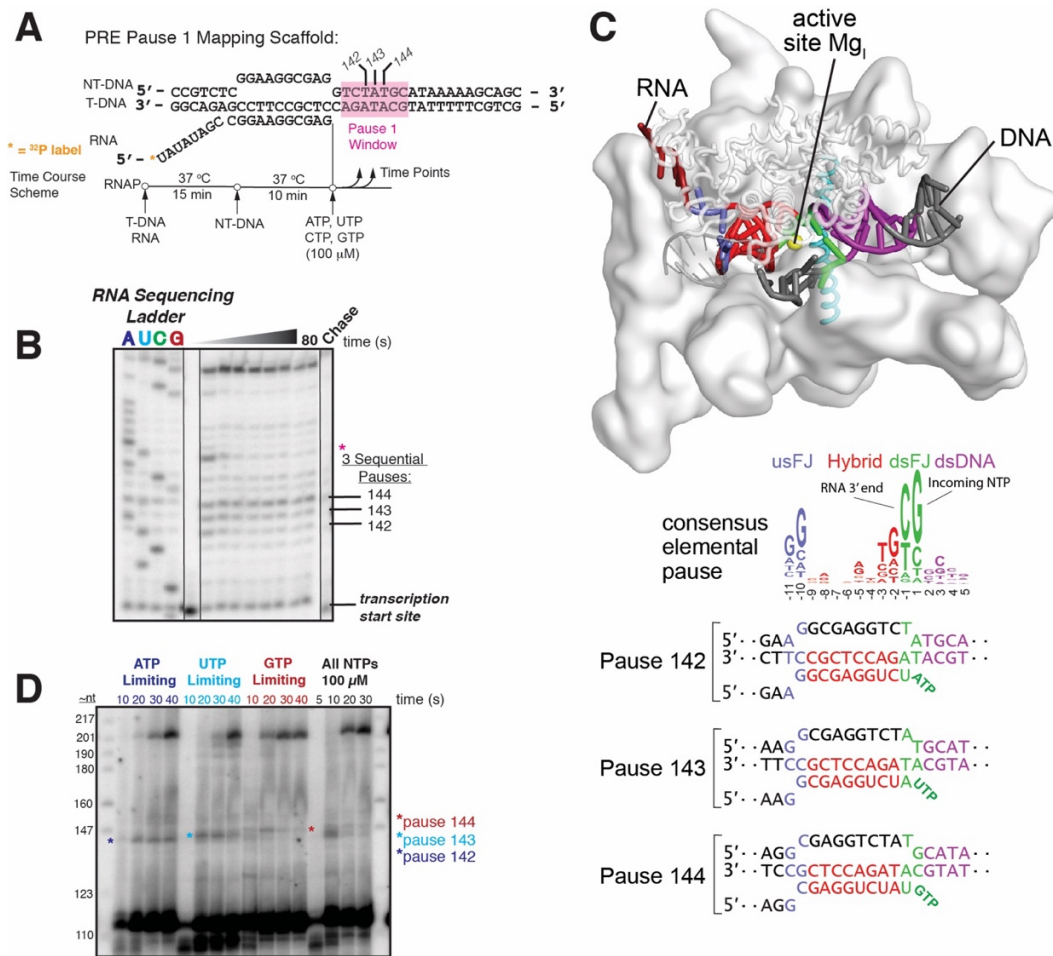


Figure 3.3 Locating the first critical pauses for PRE folding.

A. Experimental overview for mapping the first critical pause (“Pause 1”) in the PRE folding pathway. **B.** Pause mapping. RNA sequencing ladder generated using corresponding 3'-deoxyNTP. Time course samples were run alongside sequencing ladder for precise identification of the pause locations. Pause band numbering denotes nucleotide position on the PRE. All pause bands disappear after 5 min with saturating NTP (“Chase” lane), indicating these do not result from arrested transcription complexes and indeed represent transient PECs. The asterisk (*) denotes a pause RNA band downstream of the sequential pauses most likely introduced by shortening of the downstream DNA. The splice borders denote samples that were not adjacent to one another in the original gel image. **C.** Sequence comparison of PRE sequential pauses to the consensus elemental pause. (Top panel) The ePEC structure (pdb 6bjs)^{12,28} is shown above the consensus elemental pause sequence logo identified via NET-seq²⁷, represented as the nucleotide sequence of the non-template DNA. The ePEC structure and sequence logo have the following color scheme: upstream fork junction (usFJ, blue), hybrid (red), downstream fork junction (dsFJ, green), downstream DNA (dsDNA, purple). Numbering corresponds to position relative to the RNA 3' end at the RNAP active site. (Bottom panel) PRE pause sequences with color scheme to match that used for the pause consensus facilitate sequence comparison. **D.** Synchronized single-round *in vitro* transcription from the *alx* promoter to visualize the sequential pauses. Transcription elongation reactions were performed with one NTP at a limiting concentration (20 mM) to slow pause escape, permitting visualization of each pause in the sequential series. Each pause RNA band is denoted with an asterisk (*). Radiolabeled RNA samples were separated on a 10% denaturing polyacrylamide gel.

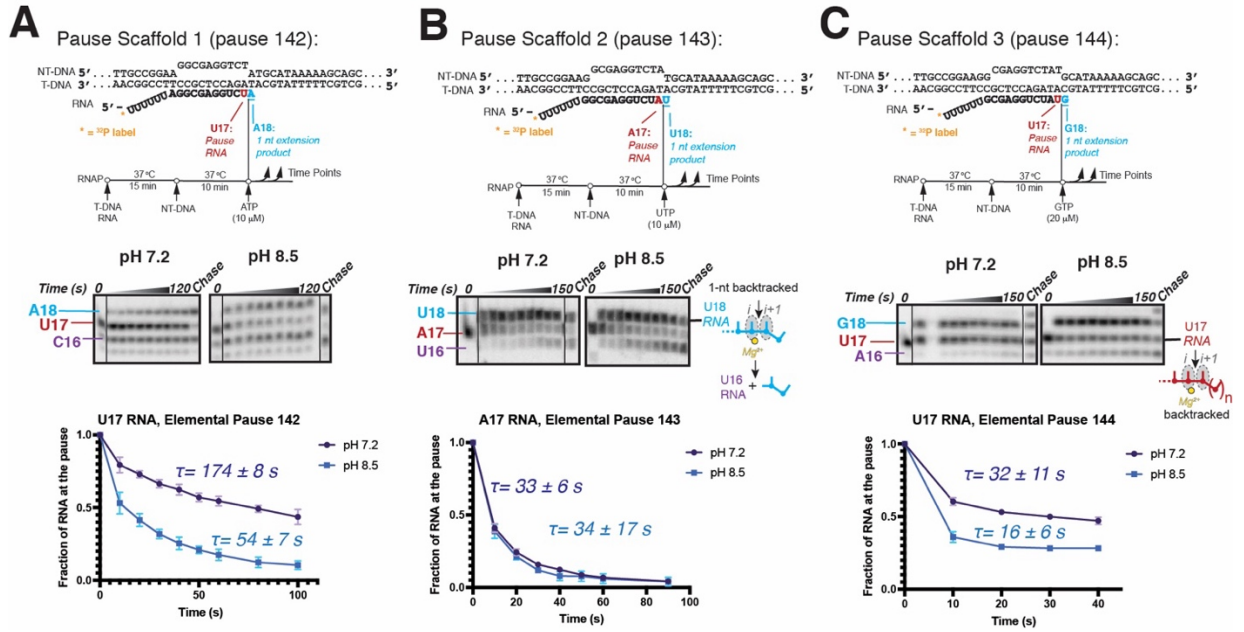


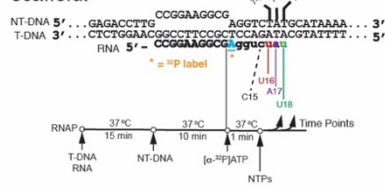
Figure 3.4 Measurements of elemental pause escape rate at the sequential pauses under neutral vs. alkaline conditions.

A. Elemental pause escape measurements at pH 7.2 and pH 8.5 for the first pause in the sequential series downstream of Hairpin S. (Top panel) Experimental overview of pause escape experiment. PECs were pre-formed by positioning RNAP on the nucleic acid scaffold shown. The rate of RNA extension was measured after addition of ATP. Time points were taken at regular intervals ranging from 0 s to 120 s. (Middle panel) Representative gel images from transcription time courses. “Chase” represents a sample taken after allowing elongation to proceed for 5 min with saturating ATP. (Bottom panel) Results from gel quantification of the pause RNA fraction over time at pH 7.2 vs. 8.5. Error is s.d. of triplicate. Average dwell time at the pause at each pH is specified. The splice borders denote samples that were not adjacent to one another in the original gel image. **B.** Elemental pause escape measurements at pH 7.2 and pH 8.5 for the second pause in the sequential series downstream of Hairpin S. (Top panel) The rate of RNA extension was measured after addition of UTP. Time points were taken at regular intervals ranging from 0 s to 150 s. (Middle panel) Representative gel images from transcription time courses. “Chase” represents sample taken after allowing elongation to proceed for 5 min with saturating UTP. Cartoon adjacent to gel image depicts a 1-nt backtracking and intrinsic cleavage event that occurs following incorporation of UMP by ECs positioned at the second pause. (Bottom panel) Results from gel quantification of the pause RNA fraction over time at pH 7.2 vs. 8.5. Error is s.d. of triplicate. Average dwell time at the pause at each pH is specified. The splice borders denote samples that were not adjacent to one another in the original gel image. **C.** Elemental pause escape measurements at pH 7.2 and pH 8.5 for the third pause in the sequential series downstream of Hairpin S. (Top panel) The rate of RNA extension was measured after addition of GTP. Time points were taken at regular intervals ranging from 0 s to 150 s (Middle panel) Representative gel images from transcription time courses. “Chase” represents sample taken after allowing elongation to proceed for 5 min with saturating GTP. Cartoon adjacent to gel image depicts that ECs positioned at the third pause enter a backtracked register; the precise translocation register could not be ascertained in this experiment, hence the “n” notation next to the extruded RNA piece. (Bottom panel) Results from gel quantification of the pause RNA fraction over time at pH 7.2 vs. 8.5. Error is s.d. of triplicate. Average dwell time at the pause at each pH is specified. The splice borders denote samples that were not adjacent to one another in the original gel image.

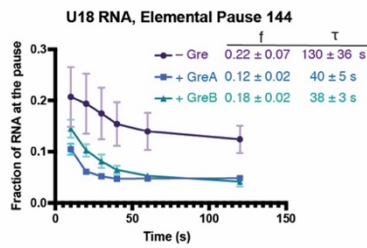
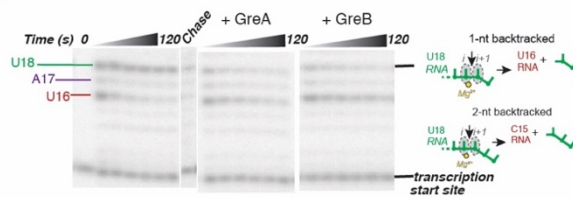
Figure 3.5 Measurement of pause kinetics at the third, backtracked sequential pause during active transcription elongation in presence and absence of Gre factors.

A. Experimental overview of sequential pause kinetic experiments. ECs were positioned upstream of the pauses on the nucleic acid scaffold shown. Transcription was initiated by addition of all 4 NTPs, with GTP present at a limiting concentration to increase the duration of the third pause. Time points were taken at regular intervals ranging from 0 s to 120 s **B.** Kinetic measurements for the third pause in the sequential series downstream of Hairpin S. (Top panel) Representative gel images from transcription time courses performed at pH 7.2. U16, A17, and U18 RNAs correspond to the first, second, and third sequential pauses, respectively. Cartoons adjacent to the rightmost gel panel depict the proposed backtracked registers adopted by ECs at the third pause. “Chase” represents a sample taken after allowing elongation to proceed for 5 min with saturating NTP. (Bottom panel) Results from gel quantification of the U18 pause RNA fraction over time at pH 7.2 in the presence and absence of Gre factors. Error is s.d. of triplicate. The table adjacent to the plot reports the pause efficiencies (f) and average pause dwell times (τ). **C.** Kinetic measurements for the third pause in the sequential series downstream of Hairpin S. (Top panel) Representative gel images from transcription time courses performed at pH 8.5. U16, A17, and U18 RNAs correspond to the first, second, and third sequential pauses, respectively. Cartoons adjacent to the rightmost gel panel depict the proposed backtracked registers adopted by ECs at the third pause. “Chase” represents a sample taken after allowing elongation to proceed for 5 min with saturating NTP. (Bottom panel) Results from gel quantification of the U18 pause RNA fraction over time at pH 8.5 in the presence and absence of Gre factors. Error is s.d. of triplicate. The table adjacent to the plot reports the pause efficiencies (f) and average pause dwell times (τ).

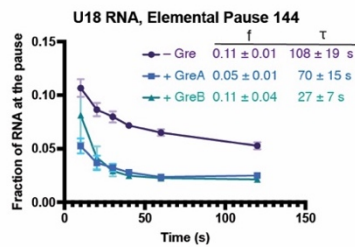
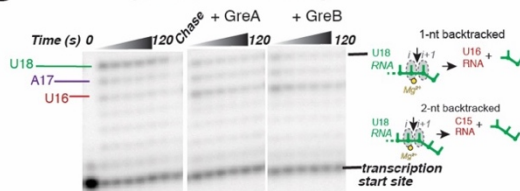
A Sequential pause kinetic measurements scaffold:



B pH 7.2, GTP limiting



C pH 8.5, GTP limiting



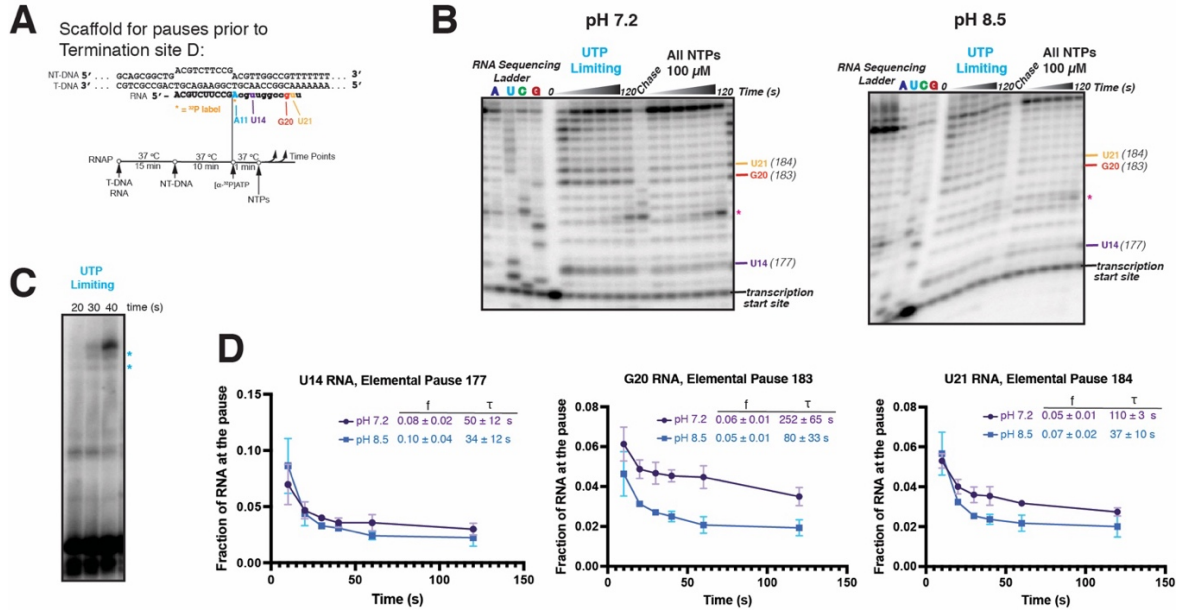


Figure 3.6 Measurement of pause kinetics at Pause 2 during active transcription elongation.

A. Experimental overview of Pause 2 mapping and pause escape kinetic experiments. ECs were positioned upstream of the pauses on the nucleic acid scaffold shown. Transcription was initiated by addition of all 4 NTPs. Time points were taken at regular intervals ranging from 0 s to 120 s. **B.** Representative gel images from transcription time courses performed at pH 7.2 and 8.5. Time course samples were run alongside an RNA sequencing ladder for precise mapping of pause locations. U14, G20 and U21 RNAs represent pauses that were enhanced in the UTP limiting condition and correspond to nt 177, 183, and 184 on the PRE, respectively. “Chase” represents a sample taken after allowing elongation to proceed for 5 min with saturating NTP. The band marked with an asterisk (*) of unknown origin accumulated at long reaction time and high NTP concentration. **C.** Gel image of synchronized single-round *in vitro* transcription from the *alx* promoter conducted at a limiting concentration of UTP. The UTP limiting condition increased the duration of multiple pauses prior to termination site D, which are denoted with an asterisk (*). The image for the UTP limiting time course is the same image from Fig. 3.3D. This image was reused here to iterate that the pauses observed under the UTP limiting condition in scaffold experiments are relevant in the context of the full-length RNA. **D.** Results from gel quantification of the U14, G20, and U21 RNA fractions over time at pH 7.2 and 8.5. Error is s.d. of triplicate. The table adjacent to the plots report the pause efficiencies (f) and average pause dwell times (τ).

5. References

1. Serganov, A. & Nudler, E. Review A Decade of Riboswitches. *Cell* **152**, 17–24 (2013).
2. Nechooshtan, G., Elgrably-weiss, M., Sheaffer, A., Westhof, E. & Altuvia, S. A pH-responsive riboregulator. *Genes Dev.* **23**, 2650–2662 (2009).
3. Bingham, R. J., Hall, K. S. & Slonczewski, J. L. Alkaline induction of a novel gene locus, *alx*, in *Escherichia coli*. *J. Bacteriol.* **172**, 2184–2186 (1990).
4. Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H. & Altuvia, S. Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.* **11**, 941–950 (2001).
5. Adeolu, M., Alnajar, S., Naushad, S. & Gupta, R. S. Genome-based phylogeny and taxonomy of the ‘*Enterobacteriales*’: proposal for *Enterobacterales* ord. nov. divided into the families *Enterobacteriaceae*, *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov., *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., *Morganellaceae* fam. nov., and *Budviciaceae* fam. nov. *Int. J. Syst. Evol. Microbiol.* **66**, 5575–5599 (2016).
6. Padan, E., Bibi, E., Ito, M. & Krulwich, T. A. Alkaline pH homeostasis in bacteria: New insights. *Biochim Biophys Acta* **1717**, 67–88 (2005).
7. Zilberstein, D., Agmon, V., Schuldiner, S. & Padan, E. *Escherichia coli* intracellular pH, membrane potential, and cell growth. *J. Bacteriol.* **158**, 246–252 (1984).
8. Zeinert, R., Martinez, E., Schmitz, J., Senn, K., Usman, B., Anantharaman, V., Aravind, L. & Waters, L.S. Structure-function analysis of manganese exporter proteins across bacteria. *J. Biol. Chem.* **293**, 5715–5730 (2018).
9. Nechooshtan, G., Elgrably-Weiss, M. & Altuvia, S. Changes in transcriptional pausing modify the folding dynamics of the pH-responsive RNA element. *Nucleic Acids Res.* **42**, 622–630 (2014).
10. Thomen, P., Lopez, P. J. & Heslot, F. Unravelling the mechanism of RNA-polymerase forward motion by using mechanical force. *Phys. Rev. Lett.* **94**, 128102 (2005).
11. Kang, J. Y., Mishanina, T. V, Landick, R. & Darst, S. A. Mechanisms of Transcriptional Pausing in Bacteria. *J. Mol. Biol.* **431**, 4007–4029 (2019).
12. Kang, J. Y., Mishanina, T.V., Bellecourt, M.J., Mooney, R.A., Darst, S.A. & Landick, R. RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. *Mol. Cell* **69**, 802-815.e5 (2018).
13. Nshogozabahizi, J. C., Aubrey, K. L., Ross, J. A. & Thakor, N. Applications and limitations of regulatory RNA elements in synthetic biology and biotechnology. *J. Appl. Microbiol.* **127**, 968–984 (2019).
14. Mishanina, T. V, Palo, M. Z., Nayak, D., Mooney, R. A. & Landick, R. Trigger loop of RNA polymerase is a positional , not acid – base , catalyst for both transcription and proofreading. *Proc. Natl. Acad. Sci.* **114**, E5103-E5112 (2017).

15. 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).
16. Theissen, G., Pardon, B. & Wagner, R. A quantitative assessment for transcriptional pausing of DNA-dependent RNA polymerases *in vitro*. *Anal. Biochem.* **189**, 254–261 (1990).
17. Ray-soni, A., Bellecourt, M. J. & Landick, R. Mechanisms of Bacterial Transcription Termination : All Good Things Must End. *Annu. Rev. Biochem.* **85**, 319–347 (2016).
18. Borukhov, S., Sagitov, V. & Goldfarb, A. Transcript cleavage factors from *E. coli*. *Cell* **72**, 459–466 (1993).
19. Orlova, M., Newlands, J., Das, A., Goldfarb, A. & Borukhov, S. Intrinsic transcript cleavage activity of RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 4596–4600 (1995).
20. Guo, X., Myasnikov, A.G., Chen, J., Crucifix, C., Papai, G., Takacs, M., Schultz, P. & Weixlbaumer, A. Structural Basis for NusA Stabilized Transcriptional Pausing. *Mol. Cell* **69**, 816–827 (2018).
21. 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. U. S. A.* **96**, 9545–9550 (1999).
22. Heilman-Miller, S. L. & Woodson, S. A. Effect of transcription on folding of the *Tetrahymena* ribozyme. *RNA* **9**, 722–733 (2003).
23. Wong, T. N., Sosnick, T. R. & Pan, T. Folding of noncoding RNAs during transcription facilitated by pausing-induced nonnative structures. *Proc. Natl. Acad. Sci.* **104**, 17995–18000 (2007).
24. Watters, K. E., Strobel, E. J., Yu, A. M., Lis, J. T. & Lucks, J. B. Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat. Struct. Mol. Biol.* **23**, 1124–1131 (2016).
25. Strobel, E. J., Cheng, L., Berman, K. E., Carlson, P. D. & Lucks, J. B. A ligand-gated strand displacement mechanism for ZTP riboswitch transcription control. *Nat. Chem. Biol.* **15**, 1067–1076 (2019).
26. Steinert, H., Sochor, F., Wacker, A., Buck, J., Helmling, C., Hiller, F., Keyhani, S., Noeske, J., Grimm, S., Rudolph, M.M., Keller, H., Mooney, R.A., Landick, R., Suess, B., Furtig, B., Wohnert, J. & Schwalbe, H. Pausing guides RNA folding to populate transiently stable RNA structures for riboswitch-based transcription regulation. *eLife* **6**, e21297 (2017).
27. Larson, M.H., Mooney, R.A., Peters, J.M., Windgassen, T., Dhananjaya, N., Gross, C.A., Block, S.M., Greenleaf, W.J., Landick, R. & Weissman, J.S. A Pause Sequence Enriched at Translation Start Sites Drives Transcription Dynamics *In Vivo*. *Science* **344**, 1042–1047 (2014).
28. Saba, J., Chua, X.Y., Mishanina, T.V., Dhananjaya, N., Windgassen, T.A., Mooney, R.A. & Landick, R. The elemental mechanism of transcriptional pausing. *eLife* **8**, e40981 (2019).

29. Abdelkareem, M., Saint-Andre, C., Takacs, M., Papai, G., Crucifix, C., Guo, X., Ortiz, J. & Weixlbaumer, A. Structural Basis of Transcription: RNA Polymerase Backtracking and Its Reactivation. *Mol. Cell* **75**, 298-309.e4 (2019).
30. Nudler, E., Mustaev, A., Lukhtanov, E. & Goldfarb, A. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* **89**, 33–41 (1997).
31. Kyzer, S., Kook, S. H., Landick, R. & Palangat, M. Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex. *J. Biol. Chem.* **282**, 19020–19028 (2007).
32. Tadigotla, V. R., Maoileidigh, D.O., Sengupta, A.M., Epshtein, V., Ebright, R.H., Nudler, E. & Ruckenstein, A.E. Thermodynamic and kinetic modeling of transcriptional pausing. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4439–4444 (2006).
33. Zamft, B., Bintu, L., Ishibashi, T. & Bustamante, C. Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 8948–8953 (2012).
34. Landick, R. The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.* **34**, 1062–1066 (2006).
35. Strobel, E. J., Watters, K. E., Nedialkov, Y., Artsimovitch, I. & Lucks, B. Distributed biotin – streptavidin transcription roadblocks for mapping cotranscriptional RNA folding. *Nucleic Acids Res.* **45**, 1–12 (2017).
36. Widom, J. R., Nedialkov, Y.A., Rai, V., Hayes, R.L., Brooks, C.L., Artsimovitch, I. & Walter, N.G. Ligand Modulates Cross-Coupling between Riboswitch Folding and Transcriptional Pausing. *Mol. Cell* **72**, 541–552 (2018).
37. Perdrizet, G. A., Artsimovitch, I., Furman, R., Sosnick, T. R. & Pan, T. Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proc. Natl. Acad. Sci.* **109**, 3323–3328 (2012).
38. Yakhnin, H., Yakhnin, A.V., Mouery, B.L., Mandell, Z.F., Karbasiafshar, C., Kashlev, M. & Babitzke, P. NusG-dependent RNA polymerase pausing and tylosin-dependent ribosome stalling are required for tylosin resistance by inducing 23s rRNA methylation in *Bacillus subtilis*. *mBio* **10**, e02665-19 (2019).
39. Yakhnin, A. V., Yakhnin, H. & Babitzke, P. RNA Polymerase Pausing Regulates Translation Initiation by Providing Additional Time for TRAP-RNA Interaction. *Mol. Cell* **24**, 547–557 (2006).
40. Yakhnin, A. V., FitzGerald, P.C., McIntosh, C., Yakhnin, H., Kireeva, M., Turek-Herman, J., Mandell, Z.F., Kashlev, M. & Babitzke, P. NusG controls transcription pausing and RNA polymerase translocation throughout the *Bacillus subtilis* genome. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 21628–21636 (2020).
41. Saenger, W. *Principles of Nucleic Acid Structure*. (Springer-Verlag, 1984).
42. Nixon, P. L. & Giedroc, D. P. Energetics of a strongly pH dependent RNA tertiary structure a frameshifting pseudoknot. *J. Mol. Biol.* **296**, 659–671 (2000).

43. Houck-Loomis, B., Durney, M.A., Salguero, C., Shankar, N., Nagle, J.M., Goff, S.P., & D'Souza, V. M. An equilibrium-dependent retroviral mRNA switch regulates translational recoding. *Nature* **480**, 561–564 (2011).
44. Huppler, A., Nikstad, L. J., Allmann, A. M., Brow, D. A. & Butcher, S. E. Metal binding and base ionization in the u6 RNA intramolecular stem-loop structure. *Nat. Struct. Biol.* **9**, 431–435 (2002).
45. Pechlaner, M., Donghi, D., Zelenay, V. & Sigel, R. K. O. Protonation-Dependent Base Flipping at Neutral pH in the Catalytic Triad of a Self-Splicing Bacterial Group II Intron. *Angew. Chemie - Int. Ed.* **54**, 9687–9690 (2015).
46. Baisden, J. T., Boyer, J. A., Zhao, B., Hammond, S. M. & Zhang, Q. Visualizing a protonated RNA state that modulates microRNA-21 maturation. *Nat. Chem. Biol.* **17**, 80–88 (2021).
47. Halder, A., Halder, S., Bhattacharyya, D. & Mitra, A. Feasibility of occurrence of different types of protonated base pairs in RNA: A quantum chemical study. *Phys. Chem. Chem. Phys.* **16**, 18383–18396 (2014).
48. Chawla, M., Sharma, P., Halder, S., Bhattacharyya, D. & Mitra, A. Protonation of base Pairs in RNA: Context analysis and quantum chemical investigations of their geometries and stabilities. *J. Phys. Chem. B* **115**, 1469–1484 (2011).
49. Kerpedjiev, P., Hammer, S. & Hofacker, I. L. Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams. *Bioinformatics* **31**, 3377–3379 (2015).

6. Supporting Information for: Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the *Escherichia Coli* pH-responsive riboswitch

Authors:

Christine Stephen¹ and Tatiana V. Mishanina^{*,1}

¹Department of Chemistry and Biochemistry University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093

List of contents:

Supplemental Figure 3.1: Secondary structure models of the *E. coli* pH-responsive riboswitch at pH 7.2 and 8.5

Supplemental Figure 3.2: Quantitation of synchronized single-round *in vitro* transcription from the *alx* promoter and transcription in the presence of pause-modulating factors

Supplemental Figure 3.3: Pause 1 mapping at limiting NTP concentrations and NET-seq data analysis

Supplemental Figure 3.4: Sequential pause escape measurements with ECs positioned directly at each pause and Gre-catalyzed RNA cleavage at the third pause

Supplemental Figure 3.5: Analysis of Pause 1 kinetics during active transcription elongation

Supplemental Figure 3.6: Pause 2 mapping at limiting GTP concentration and NET-seq data analysis

Supplemental Figure 3.7: Analysis of potential RNA structures in the vicinity of Pause 1 and test of *his* pause escape kinetics at different pH

Supplemental Figure 3.8: Analysis of previously reported mutations demonstrated to affect

Pause 1 longevity or location

Supplemental Table 3.1: PRE nucleotide sequence, plasmids and oligonucleotides used in this study

Supplemental Table 3.2: Summary of the sequential pause kinetic parameters from scaffold-based experiments positioning ECs at each pause (Fig. 3.4)

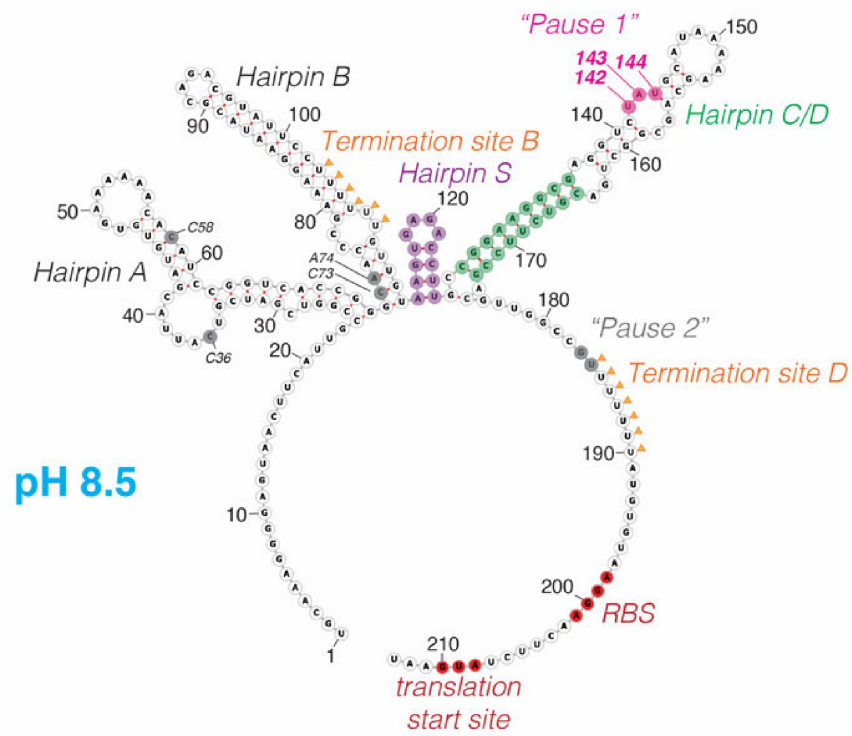
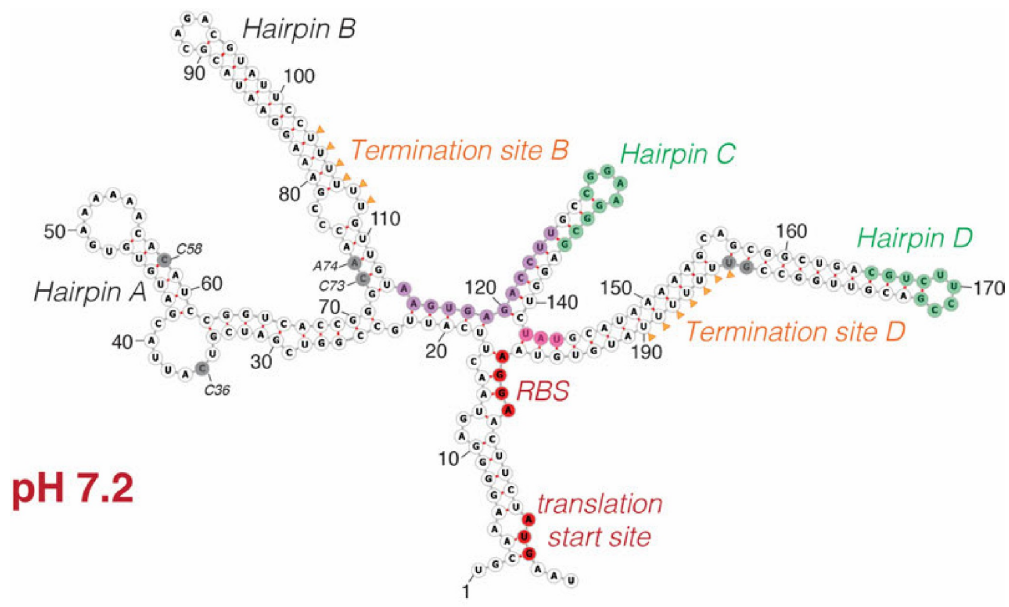
Supplemental Table 3.3: Summary of the sequential pause kinetic parameters from scaffold-based experiments during active transcription elongation (Fig. 3.5)

Supplemental Table 3.4: Summary of kinetic parameters for pauses prior to termination site D from scaffold-based experiments during active transcription elongation (Fig. 3.6)

Supplemental Discussion: Additional analysis of the EC translocation register at the sequential pauses (Figs. 3.4 and 3.5)

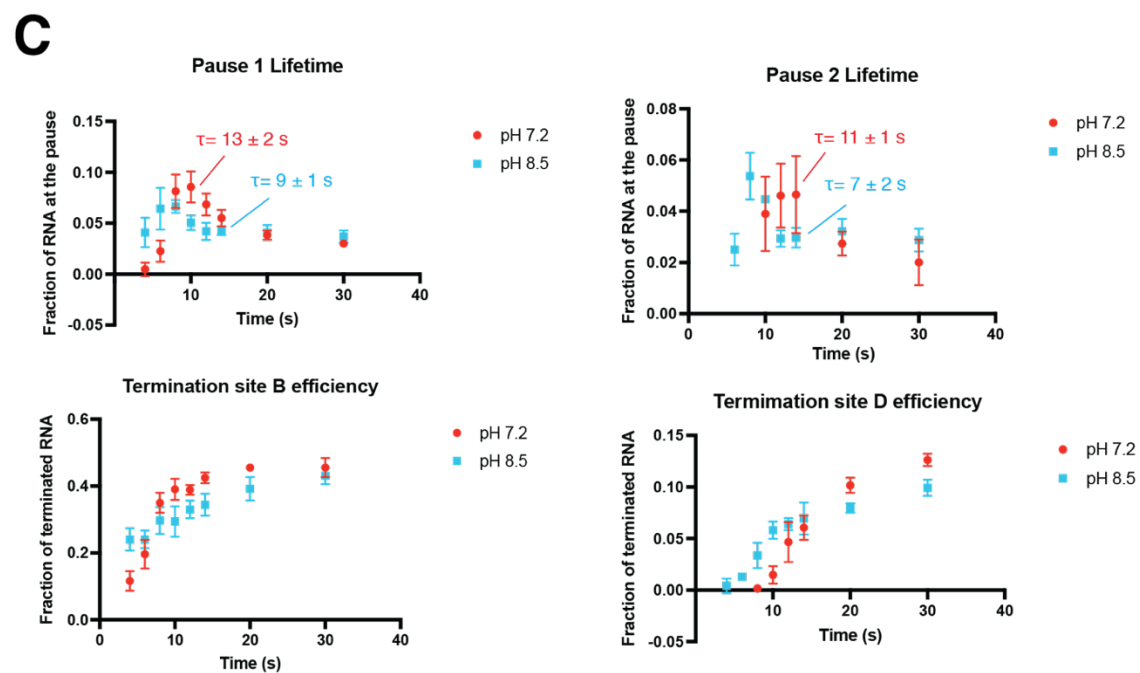
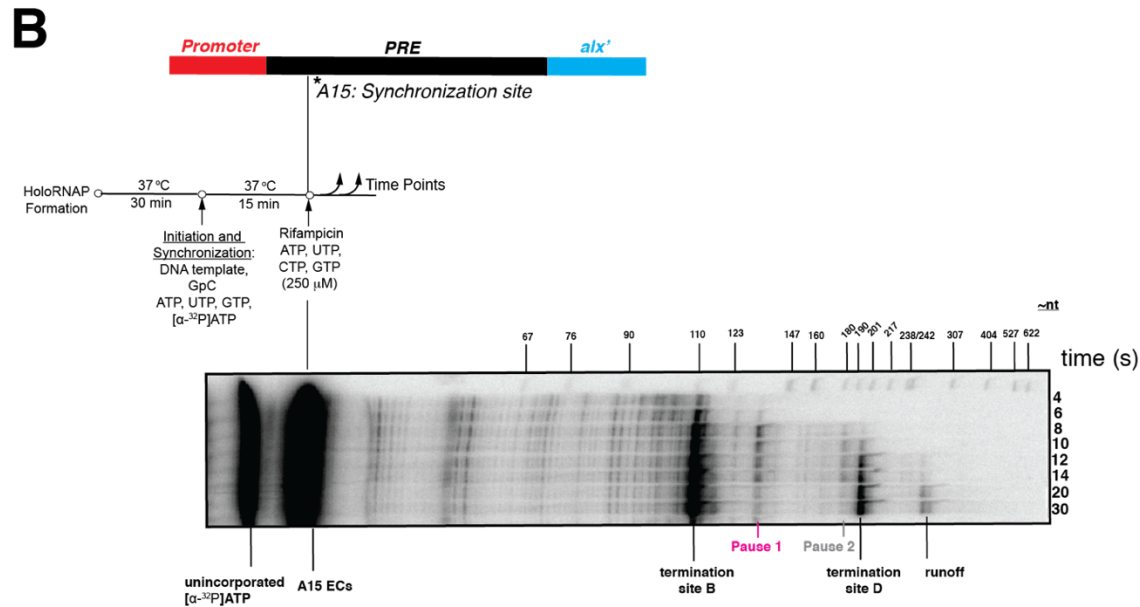
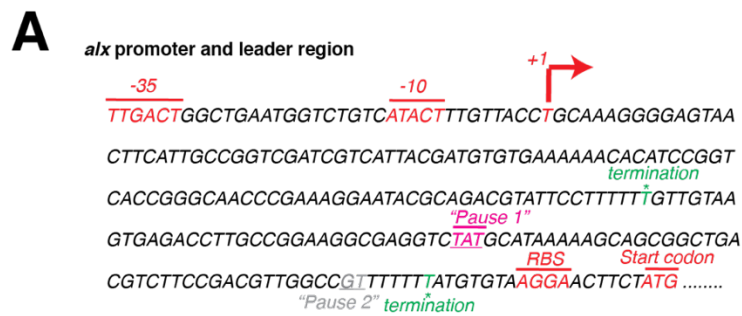
Supplemental Figure 3.1 Secondary structure models of the *E. coli* pH-responsive riboswitch at pH 7.2 and 8.5.

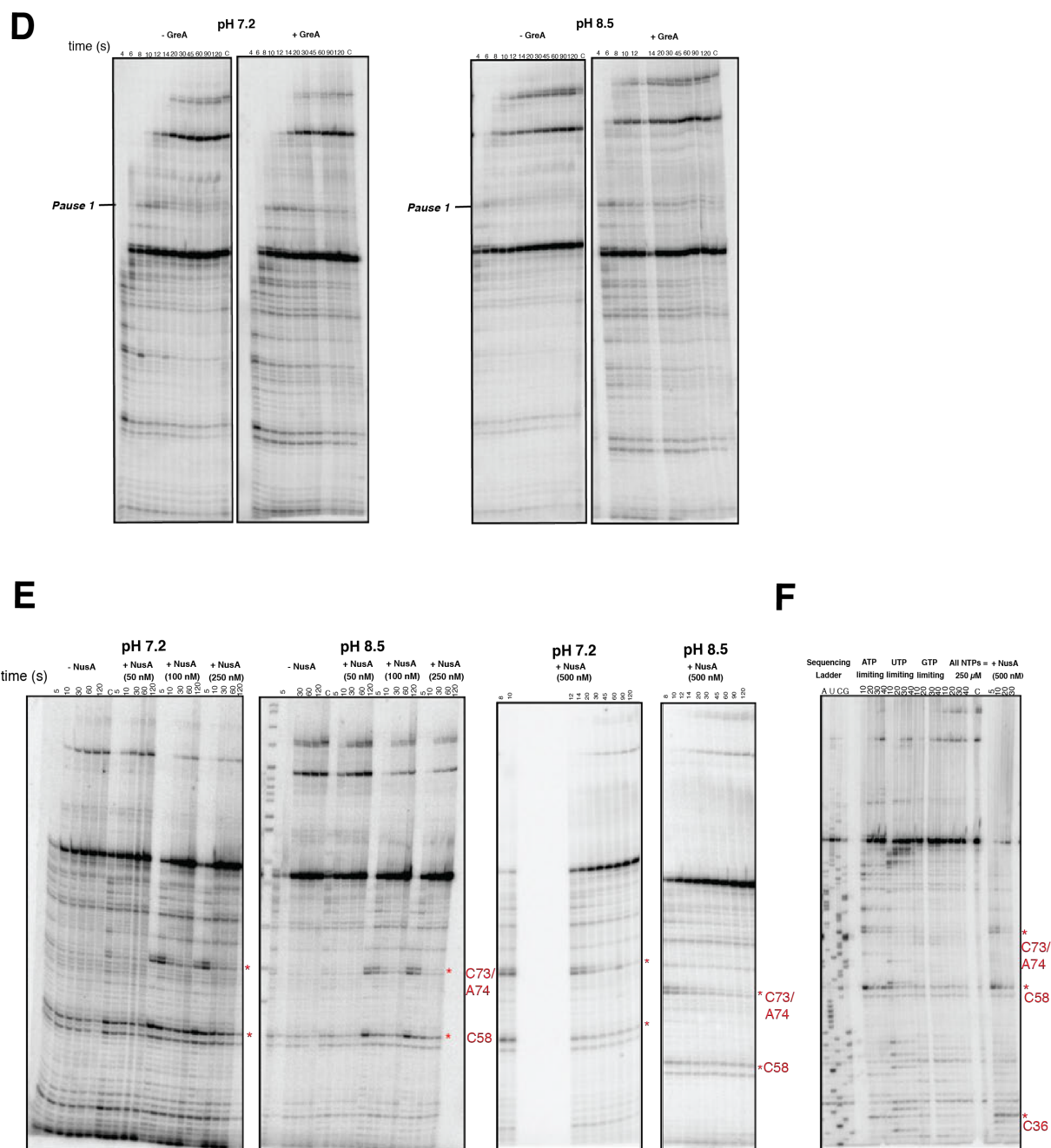
The PRE structure models presented here are based on previously published structure predictions, which were generated via chemical probing with dimethyl sulfate (DMS) followed by primer extension¹. The residues highlighted in purple and green denote regions that demonstrate alternate base pairing schemes at different pH. The residues highlighted in red represent the ribosome binding site (RBS) and the translation start site, which are base-paired in the RNA structure at neutral pH and are thus unavailable for ribosome binding and translation of the downstream gene (*a/x*). At alkaline pH, the RBS and translation start site become available in the RNA structure, prompting an increase in *a/x* gene expression. To help visualize the structural motifs in the vicinity of the newly mapped pauses, the residues in the 3' RNA end at Pause 1 and Pause 2 are highlighted in pink and grey, respectively. Also, residues C36, C58, and C73/A74 are highlighted in grey to visualize the newly mapped locations of putative hairpin-stabilized pauses. The poly-U tracts following hairpins B and D are marked with orange arrows to denote the intrinsic termination sites.



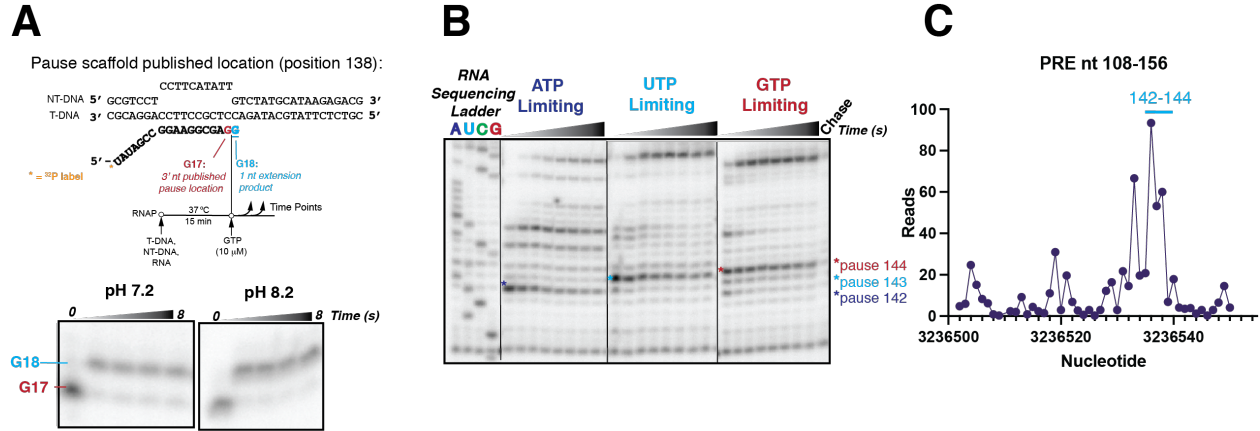
Supplemental Figure 3.2 Quantitation of synchronized single-round *in vitro* transcription from the *alx* promoter and transcription in the presence of pause-modulating factors.

A. DNA template used for synchronized, single-round *in vitro* transcription experiments. The -35 and -10 promoter elements, transcription start site (+1), RBS and start codon are denoted in red. The intrinsic termination sites downstream of Hairpins B and D are denoted in green. The newly mapped locations for Pauses 1 and 2 are denoted in pink and grey, respectively. **B.** Overview of experimental scheme for synchronized, single-round *in vitro* transcription reactions. The gel image is shown to highlight the A15 synchronization site and demonstrate an example of experimental results from a small-scale time course conducted at pH 7.2. The transcription time course samples were run alongside a radiolabeled DNA ladder (*MspI*-digested pBR322 plasmid) and analyzed on a 6% polyacrylamide gel. **C.** Results from gel quantification of promoter-initiated transcription time courses. (Top panels) Results from gel quantification of the “Pause 1” and “Pause 2” RNA fractions over time at pH 7.2 vs. 8.5. (Bottom panels) Results from gel quantification of RNAs resulting from intrinsic termination events at termination sites B and D. **D.** Results from synchronized, single-round *in vitro* transcription experiments to probe the effect of GreA on Pause 1 at pH 7.2 and 8.5. Gel images from the main text Fig. 3.2B are included for comparison of Pause 1 RNA species in the presence and absence of GreA. GreA was present at a final concentration of 500 nM in the experiment and RNA transcripts were visualized on 10% polyacrylamide gels. **E.** Results from synchronized, single-round *in vitro* transcription experiments to probe the effect of NusA on pausing. (Left panels) Transcription time courses at pH 7.2 and 8.5 in the presence of 50 nM, 100 nM, and 250 nM NusA. (Right panels) Transcription time courses at pH 7.2 and 8.5 in the presence of 500 nM NusA. RNA transcripts for all experiments with NusA were visualized on 10% polyacrylamide gels. **F.** Samples from a transcription time course with NusA at 500 nM run alongside an RNA sequencing ladder and samples from time courses conducted at various limiting NTP concentrations to increase pause durations and aid with pause mapping. RNA transcripts were visualized on a 10% polyacrylamide gel.





Supplemental Figure 3.2 Quantitation of synchronized single-round *in vitro* transcription from the *alx* promoter and transcription in the presence of pause-modulating factors. (Continued).

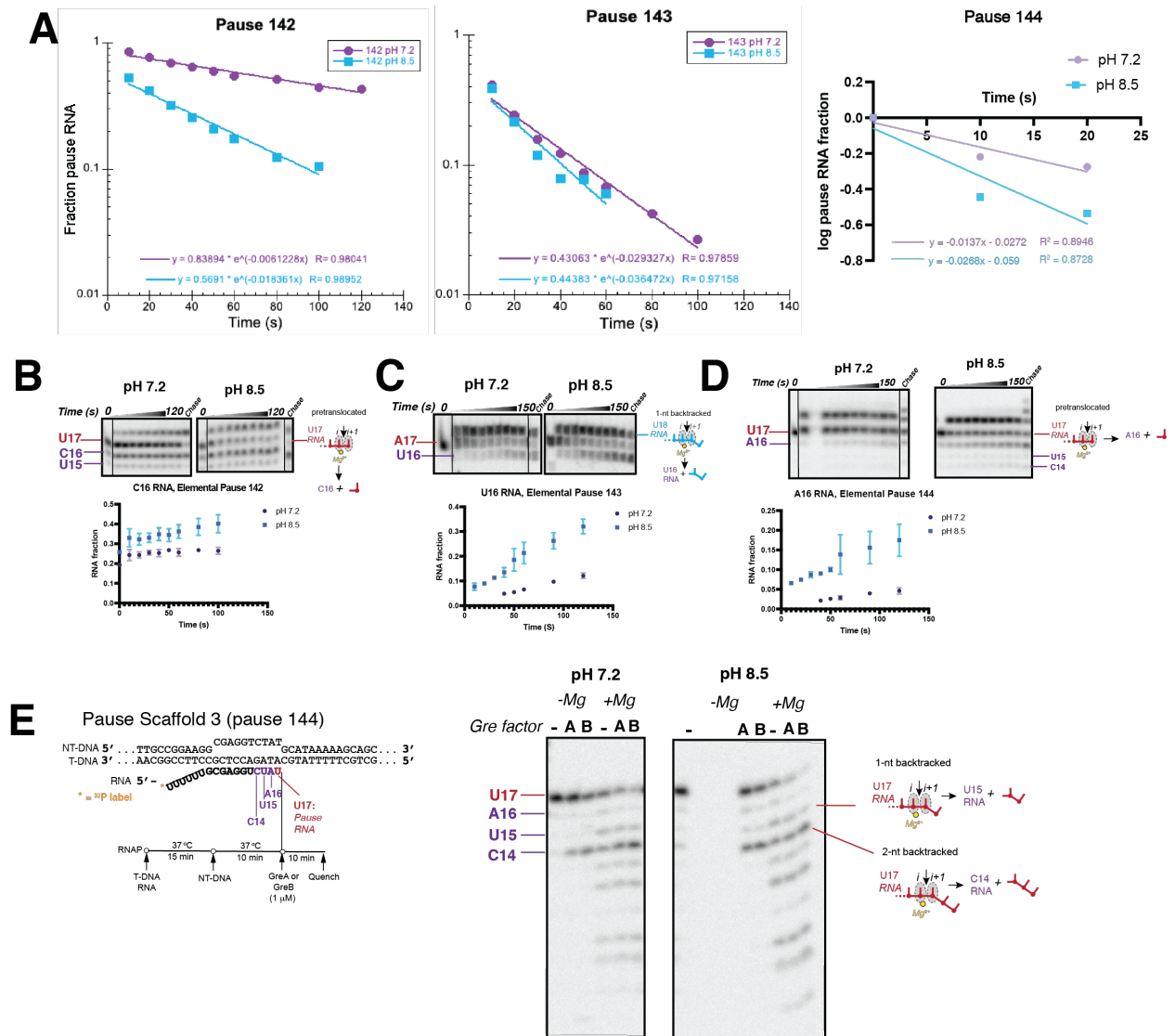


Supplemental Figure 3.3 Pause 1 mapping at limiting NTP concentrations and NET-seq data analysis.

A. Elemental pause escape measurements at pH 7.2 and 8.2 for the published location of Pause 1 in the PRE folding pathway. The experiments were conducted at pH 7.2 and 8.2 to match the conditions used in the original analysis of this putative pause¹. Upon addition of the incoming NTP, rapid RNA extension was observed at both pH, indicating the pause sequence was mapped incorrectly. G17 denotes position of the RNA band corresponding to the putative pause and G18 denotes the read-through RNA band. **B.** Additional transcription elongation time courses for pause mapping experiment. Time courses were performed with one NTP limiting to slow pause escape and clearly visualize the sequential pauses. RNA transcripts were visualized on a 15% polyacrylamide gel. **C.** Analysis of published NET-Seq data at the genomic location of the PRE in *E. coli*². The peaks representing the newly mapped pause locations at positions 142-144 on the PRE sequence are denoted in blue.

Supplemental Figure 3.4 Sequential pause escape measurements with ECs positioned directly at each pause and Gre-catalyzed RNA cleavage at the third pause.

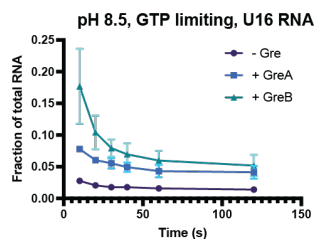
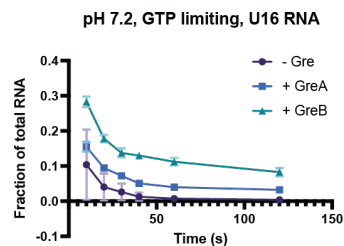
A. Fraction of RNA within each paused transcription complex is plotted as a function of time. Exponential fitting was performed on the initial, linear phase of pause decay. The equation for each fit is shown. For the third pause, an alternative method was used to obtain kinetic parameters due to poor results from the exponential fit. The alternative method used was similar to that applied for pause kinetic analysis on promoter-initiated transcription experiments (see materials and methods section “Quantification of pause kinetics and termination efficiency” for additional details). **B.** Analysis of the C16 RNA fraction resulting from intrinsic cleavage of ECs in the pre-translocated register at the first pause, which is depicted via a cartoon adjacent to the gel images. The fraction of the C16 RNA was plotted over time at both pH to visualize its accumulation during the transcription time course. A small fraction of U15 RNA appeared during the time course as well. **C.** Analysis of the U16 RNA fraction resulting from intrinsic cleavage with ECs in a 1-nt backtracked register following UMP incorporation at the second pause, which is depicted via a cartoon adjacent to the gel images. The fraction of the U16 RNA was plotted over time at both pH to visualize its accumulation during the transcription time course. **D.** Analysis of the A16 RNA fraction resulting from intrinsic cleavage of ECs in the pre-translocated register at the third pause, which is depicted via a cartoon adjacent to the gel images. The fraction of the A16 RNA was plotted over time at both pH to visualize its accumulation during the transcription time course. Additional minor RNA products, U15 and C14, appeared during the transcription time course at pH 8.5 and are likely the result of intrinsic cleavage by ECs in either a 1-nt or 2-nt backtracked register at the third pause. **E.** Test for putative backtracking event at the third pause using *E. coli* transcription factors GreA and GreB. (Left panel) Scaffold diagram and experimental scheme. (Right panel) Gel analysis of RNAs from the experiment, which was performed at pH 7.2 and 8.5 in the presence or absence of 5 mM MgCl₂ supplemented in the reconstitution buffer. For each condition, there was either no Gre factor present (-), GreA present (denoted by A), or GreB present (denoted by B). The identities of the A16, U15, and C14 RNAs are the result of intrinsic cleavage of ECs in the pre-translocated, 1-nt, and 2-nt backtracked registers, respectively. Cartoons adjacent to the gel image were added to help visualize the 1-nt and 2-nt backtracking events. RNA transcripts were visualized on a 15% polyacrylamide gel.



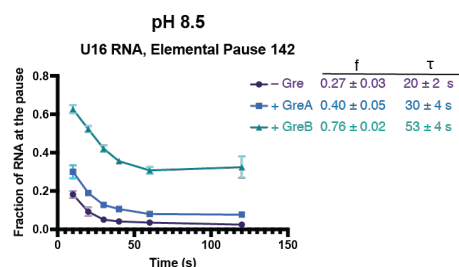
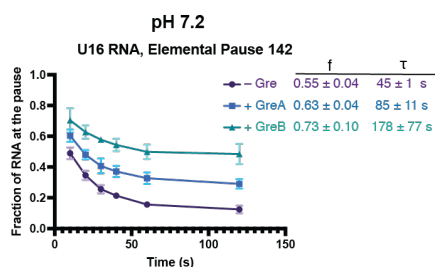
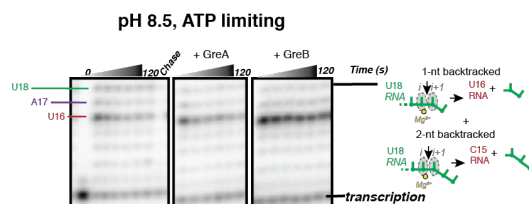
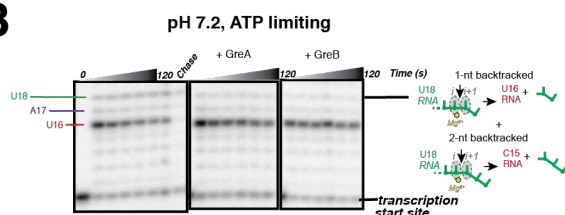
Supplemental Figure 3.5 Analysis of Pause 1 kinetics during active transcription elongation.

A. Plots of the U16 RNA fraction over time at pH 7.2 and 8.5 for pause escape measurements conducted at the GTP limiting condition in the presence and absence of Gre factors. **B.** Pause escape measurements for the first pause in the sequential series. (Top panels) Gel images from transcription time courses conducted at a limiting concentration of ATP in the presence and absence of Gre factors at pH 7.2 and 8.5. The U16, A17, and U18 RNAs are noted to label the RNAs corresponding to the first, second, and third sequential pause sites, respectively. RNA transcripts were analyzed on 15% polyacrylamide gels. (Bottom panels) Results from gel quantification of the pause RNA fraction over time at pH 7.2 vs. 8.5. Error is s.d. of triplicate. The table adjacent to the plot reports the pause efficiencies (f) and average pause dwell times (r) obtained in the presence and absence of Gre factors. **C.** Pause escape measurements for the second pause in the sequential series. (Top panels) Gel images from transcription time courses conducted at a limiting concentration of UTP in the presence and absence of Gre factors at pH 7.2 and 8.5. The U16, A17, and U18 RNAs are noted to label RNAs corresponding to the first, second, and third sequential pause sites, respectively. RNA transcripts were analyzed on 15% polyacrylamide gels. (Bottom panels) Results from gel quantification of the pause RNA fraction over time at pH 7.2 and 8.5. Error is s.d. of triplicate. The table adjacent to the plot reports the pause efficiencies (f) and average pause dwell times (r) obtained in the presence and absence of Gre factors. **D.** Results from gel quantification of the C15 RNA fraction over time at pH 7.2 and 8.5 from transcription time courses shown in Fig. S3.5C.

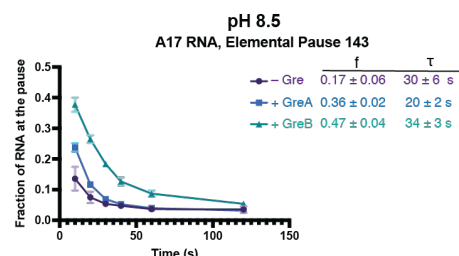
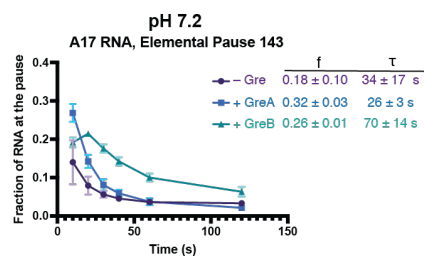
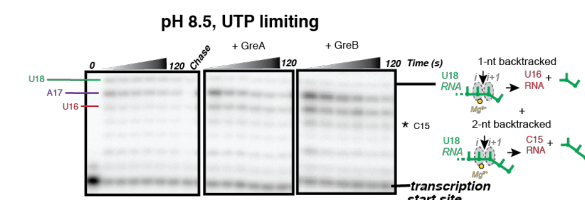
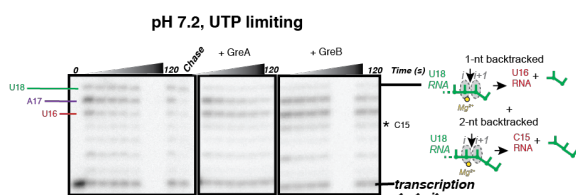
A



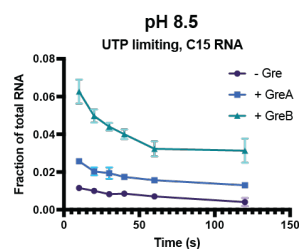
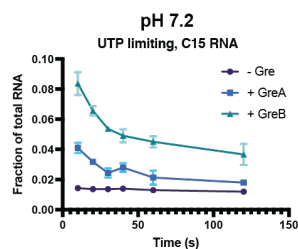
B

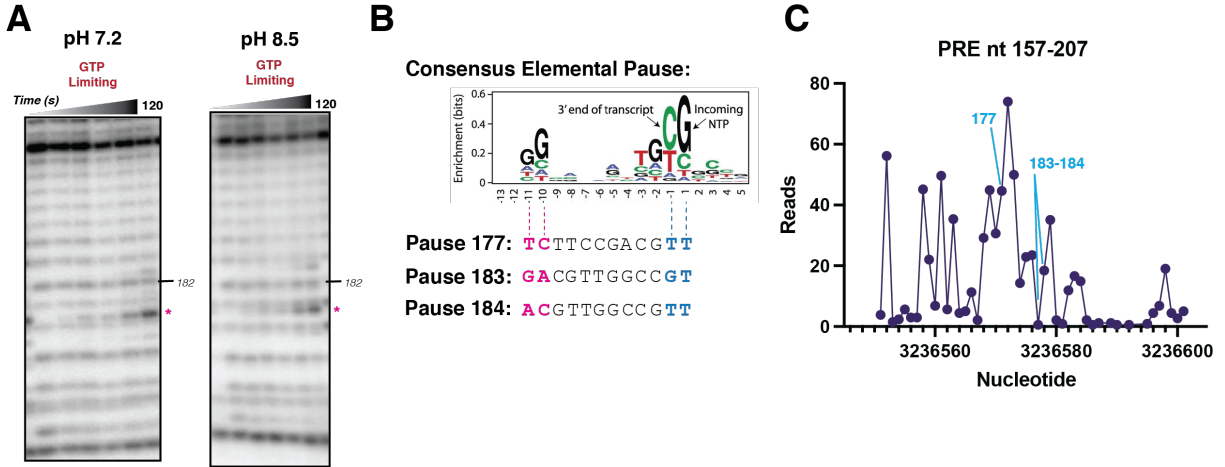


C



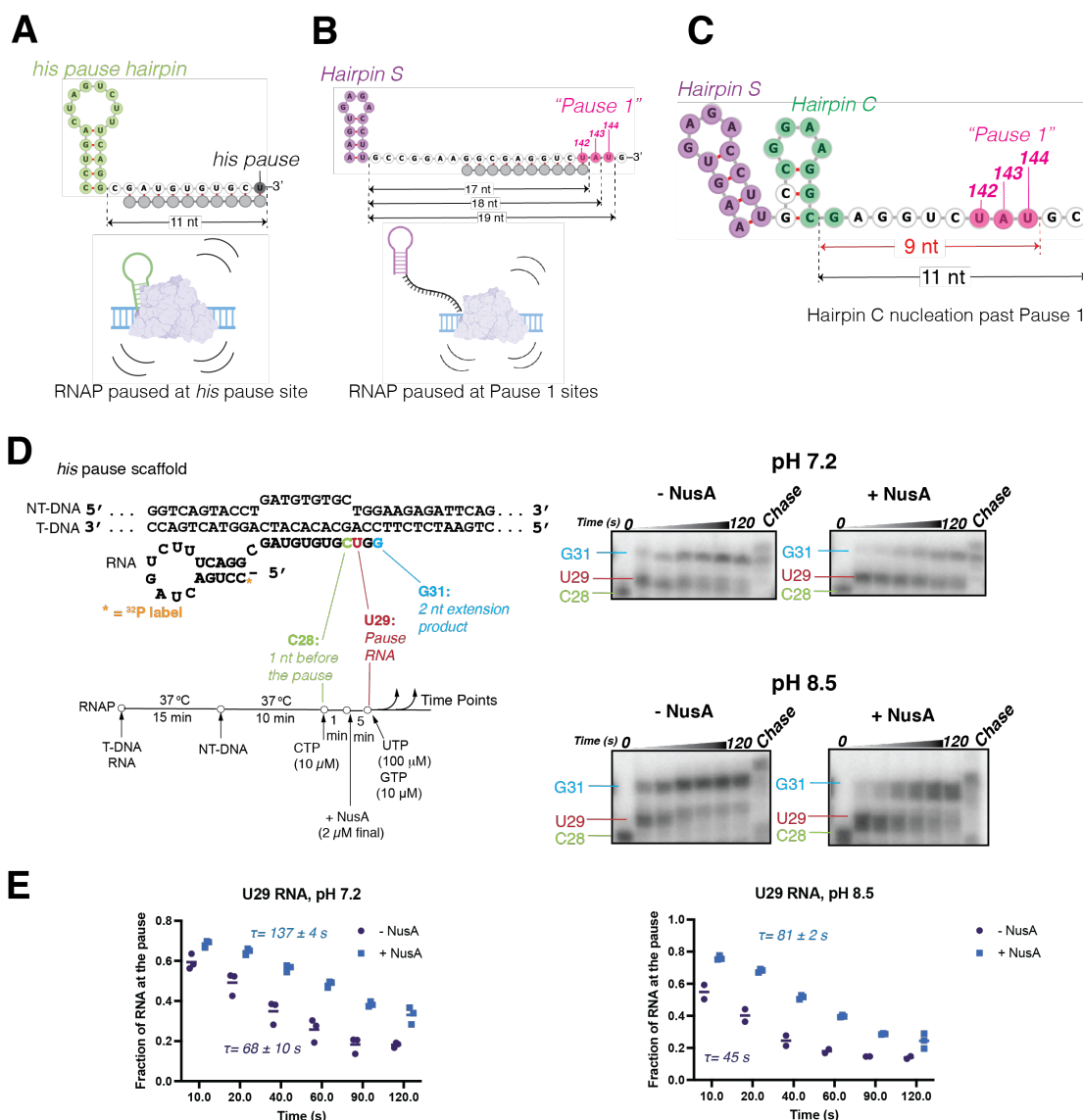
D





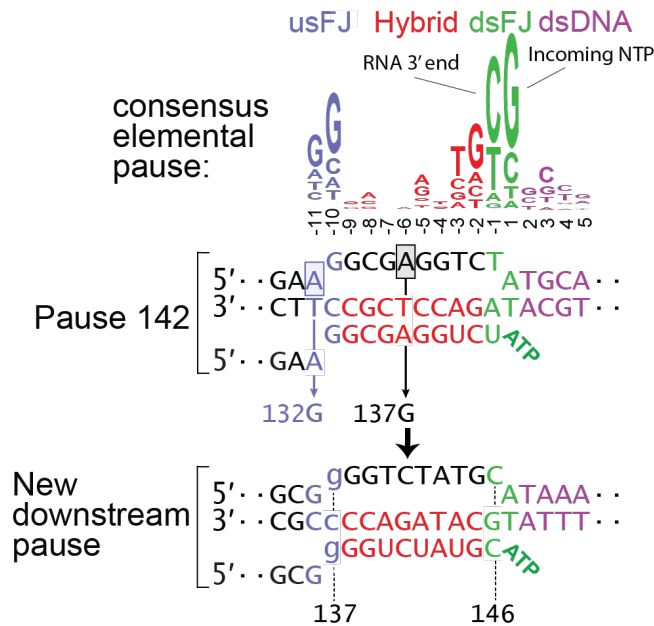
Supplemental Figure 3.6 Pause 2 mapping at limiting GTP concentration and NET-seq data analysis.

A. Additional results from Pause 2 mapping experiment conducted at a limiting concentration of GTP. The position of the reported Pause 2 location (182) is noted on the gel. The band marked with an asterisk (*) of unknown origin accumulated at long reaction time and high NTP concentration. **B.** Sequence comparison of the newly mapped pauses prior to termination site D to the consensus elemental pause. (Top panel) Consensus elemental pause sequence logo identified via NET-seq², shown as nucleotide sequence of the non-template DNA. Numbering corresponds to position relative to the RNA 3' end at the RNAP active site. (Bottom panel) Pause sequences prior to termination site D with numbered positions to facilitate comparison with the pause consensus sequence. **C.** Analysis of published NET-seq data at the genomic location of the PRE in *E. coli*². The peaks representing the newly mapped pause locations at positions 177, 183, and 184 on the PRE sequence are denoted in blue.



Supplemental Figure 3.7 Analysis of potential RNA structures in the vicinity of Pause 1 and test of *his* pause escape kinetics at different pH.

A. At a hairpin-stabilized pause such as the *his* pause, the nascent RNA hairpin is positioned ~11 nt away from the RNA 3' end at the active site. This spacing places the RNA hairpin within the RNA exit channel of RNA polymerase (RNAP, shown as a cartoon), stabilizing the paused RNAP and enhancing pause duration³. **B.** At the sequential pauses, the "Hairpin S" motif is situated 17-19 nt from the RNA 3' end at the active site, positioning it outside of the RNA exit channel. **C.** At the sequential pauses, the "Hairpin C" motif can only begin nucleation in the RNAP exit channel past the location of Pause 1. **D.** Testing pH dependence of pause escape kinetics at the *his* pause in the presence and absence of NusA (Left panel) Scaffold overview and experimental scheme. (Right panels) Gel images for *his* pause escape measurements at pH 7.2 and 8.5 in the presence and absence of NusA. RNA transcripts were analyzed on 15% polyacrylamide gels. **E.** Results from gel quantification of the pause RNA fraction over time at pH 7.2 and 8.5 in the presence and absence of NusA. Individual data points from each time course are shown. Three replicates were performed for all experiments at pH 7.2 and for the experiment at pH 8.5 with NusA. Two replicates were performed for the experiment at pH 8.5 without NusA present. Average dwell time at the pause at each pH is specified. An error is not reported for the pH 8.5 experiment without NusA since two replicates were performed.



Supplemental Figure 3.8 Analysis of previously reported mutations demonstrated to affect Pause 1 longevity or location.

Mutations that were previously reported to affect the duration and location of Pause 1⁴ are analyzed in the context of the newly mapped pause locations. (Top panel) Consensus elemental pause sequence logo identified via NET-seq², shown as nucleotide sequence of the non-template DNA. Numbering corresponds to position relative to the RNA 3' end at the RNAP active site. (Middle panel) The A132G mutation introduces a G at the -11 position at pause 142, increasing the similarity of the pause to the consensus sequence and explaining the increase in pause duration that was observed in prior work. (Bottom panel) The A137G mutation prompts a new pause site shifted downstream of the original Pause 1 location by introducing a -10 G at the 146 position creating a new sequence with a high degree of similarity to the consensus sequence.

Supplemental Table 3.1 PRE nucleotide sequence, plasmids and oligonucleotides used in this study.

Nucleotide sequence of the PRE RNA
UGCAAAGGGGAGUAACUUCAUUGCCGGUCGAUCGUCAUUACGAUGUGUGAAAAACA CAUCCGGUCACCGGGCAACCCGAAAGGAAUACGCAGACGUUUUUUUUGUUGUA AGUGAGACCUUGCCGGAAGGCGAGGUCUAUGCAUAAAAAGCAGCGGCUGACGUCUUC CGACGUUGGCCGUUUUUUUAUGUGUAAGGAACUUCUAUGAAU

Plasmids	Source
<u>pRM756</u> Wild-type <i>E. coli</i> RNAP ($\alpha 2\beta\beta'$ ω) overexpression plasmid. His10-ppx tag at <i>rpoC</i>-C-terminus	Landick Lab, University of Wisconsin- Madison
<u>pMIS4</u> overexpression plasmid for wild-type <i>E. coli</i> σ^{70} with an N-terminal His6 tag	Landick Lab, University of Wisconsin- Madison
<u>pMIS301</u> plasmid encoding wild-type PRE and the first 59 nt of <i>alx</i>, transcribed from the <i>alx</i> promoter	This work
<u>pRA46</u> overexpression plasmid for wild-type <i>E. coli</i> NusA with an N-terminal His6 tag	This work
<u>pIA578</u> overexpression plasmid for wild-type <i>E. coli</i> GreA with a C-terminal His6 tag	Landick Lab, University of Wisconsin- Madison
<u>pIA577</u> overexpression plasmid for wild-type <i>E. coli</i> GreB with a C-terminal His6 tag	Landick Lab, University of Wisconsin- Madison

Supplemental Table 3.1 PRE nucleotide sequence, plasmids and oligonucleotides used in this study.(Continued)

Oligonucleotides (5'-3')	Source
<u>pMIS301 forward primer for transcription template preparation</u> TGGATGAGCATCTGGTGGAT	IDT
<u>pMIS301 reverse primer for transcription template preparation</u> GCGACAACAACAGCGAATC	IDT
<u>Pause 1 mapping experiment template DNA</u> GCTGCTTTTTATGCATAGACCTCGCCTTCCGAGACGG	IDT
<u>Pause 1 mapping experiment non-template DNA</u> CCGTCTCGGAAGGCGAGGTCTATGCATAAAAAGCAGC	IDT
<u>Pause 1 mapping experiment RNA</u> UAUAUAGCCGGAAGGCGAG	IDT
<u>Sequential pause escape rate measurement template DNA</u> GCCGCTGCTTTTTATGCATAGACCTCGCCTTCCGGCAAGGTCTCACTTACAAG GACGC	IDT
<u>Sequential pause escape rate measurement non-template DNA</u> GCGTCCTTGTAAGTGAGACCTTGCCGGAAGGCGAGGTCTATGCATAAAAAGC AGCGGC	IDT
<u>Sequential pause 1 (nt 142 on PRE sequence) elemental pause RNA</u> UUUUUUAGGCGAGGUCU	IDT
<u>Sequential pause 2 (nt 143 on PRE sequence) elemental pause RNA</u> UUUUUUGGCGAGGUCUA	IDT

Supplemental Table 3.1 PRE nucleotide sequence, plasmids and oligonucleotides used in this study.(Continued)

<u>Sequential pause 3 (nt 144 on PRE sequence) elemental pause RNA</u> UUUUUUGCGAGGUCUAU	IDT
<u>RNA that positions ECs at nt 136 on PRE sequence</u> CCGGAAGGCG	IDT
<u>Reported Pause 1 location test template DNA</u> CGTCTCTTATGCATAGACCTCGCCTTCCAGGACGC	IDT
<u>Reported Pause 1 location test non-template DNA</u> GCGTCCTCCTTCATATTGTCTATGCATAAGAGACG	IDT
<u>Reported Pause 1 location test RNA (nt 138 on PRE sequence)</u> UAUAUAGCCGGAAGGCGAG	IDT
<u>Pause 2 mapping and escape rate measurement template DNA</u> TTCCTTACACATAAAAAACGGCCAACGTCGGAAGACGTCAGCCGCTGCAGG ACGC	IDT
<u>Pause 2 mapping and escape rate measurement non-template DNA</u> GCGTCCTGCAGCGGCTGACGTCTCCGACGTTGGCCGTTTTTTTATGTGTAAG GAA	IDT
<u>Pause 2 mapping RNA (places ECs at nt 173 on the PRE sequence)</u> ACGUCUCCG	IDT
<u><i>his</i> pause escape measurement template DNA</u> CCGGAAGATCTCTCTGAATCTCTTCCAGCACACATCAGGTACTGACC	IDT
<u><i>his</i> pause escape measurement non-template DNA</u> GGTCAGTACCTGATGTGTGCTGGAAGAGATTCAGAGAGATCTTCCGG	IDT
<u><i>his</i> pause escape measurement RNA</u> CCUGACUAGUCUUUCAGGCGAUGUGUG	IDT

Supplemental Table 3.2 Summary of the sequential pause kinetic parameters from scaffold-based experiments positioning ECs at each pause (Fig. 3.4).

Pause location	Rate constant (k), s ⁻¹	Average dwell time (τ), s	Pause half-life ($t_{1/2}$), s
Pause 142, pH 7.2	0.006 ± 0.0003	174 ± 8	121 ± 6
Pause 142, pH 8.5	0.02 ± 0.003	54 ± 7	38 ± 5
Pause 143, pH 7.2	0.03 ± 0.006	33 ± 6	23 ± 4
Pause 143, pH 8.5	0.03 ± 0.01	34 ± 17	24 ± 12
Pause 144, pH 7.2	0.03 ± 0.01	32 ± 11	22 ± 8
Pause 144, pH 8.5	0.06 ± 0.02	16 ± 6	11 ± 4

Note: For Pause 144, parameters were obtained using an alternative method due to poor results from the exponential fit (see materials and methods “Quantification of pause kinetics and termination efficiency” section for details).

Supplemental Table 3.3 Summary of the sequential pause kinetic parameters from scaffold-based experiments during active transcription elongation (Fig. 3.5).

Pause location	Rate constant (k), s ⁻¹	Average dwell time (r), s	Pause half-life ($t_{1/2}$), s	Pause efficiency (f)
Pause 142 pH 7.2 – Gre	0.02 ± 0.0002	45 ± 1	31 ± 1	0.55 ± 0.04
Pause 142 pH 7.2 + GreA	0.01 ± 0.002	85 ± 11	59 ± 7	0.63 ± 0.04
Pause 142 pH 7.2 + GreB	0.007 ± 0.004	178 ± 77	123 ± 53	0.73 ± 0.1
Pause 142 pH 8.5 – Gre	0.05 ± 0.004	20 ± 2	14 ± 1	0.27 ± 0.03
Pause 142 pH 8.5 + GreA	0.03 ± 0.005	29 ± 4	20 ± 3	0.40 ± 0.05
Pause 142 pH 8.5 + GreB	0.02 ± 0.001	53 ± 4	36 ± 3	0.75 ± 0.02
Pause 143 pH 7.2 – Gre	0.03 ± 0.01	34 ± 17	24 ± 11	0.18 ± 0.10
Pause 143 pH 7.2 + GreA	0.04 ± 0.005	26 ± 3	18 ± 2	0.32 ± 0.03
Pause 143 pH 7.2 + GreB	0.01 ± 0.003	70 ± 14	48 ± 10	0.26 ± 0.01
Pause 143 pH 8.5 – Gre	0.03 ± 0.007	30 ± 6	21 ± 4	0.17 ± 0.06
Pause 143 pH 8.5 + GreA	0.05 ± 0.004	20 ± 2	14 ± 1	0.36 ± 0.02
Pause 143 pH 8.5 + GreB	0.03 ± 0.003	34 ± 3	23 ± 2	0.47 ± 0.04
Pause 144 pH 7.2 – Gre	0.008 ± 0.002	130 ± 36	90 ± 25	0.22 ± 0.07
Pause 144 pH 7.2 + GreA	0.03 ± 0.003	40 ± 5	28 ± 4	0.12 ± 0.02
Pause 144 pH 7.2 + GreB	0.03 ± 0.002	38 ± 3	26 ± 2	0.18 ± 0.02
Pause 144 pH 8.5 – Gre	0.009 ± 0.002	108 ± 19	75 ± 13	0.11 ± 0.008
Pause 144 pH 8.5 + GreA	0.01 ± 0.004	70 ± 15	48 ± 11	0.05 ± 0.01
Pause 144 pH 8.5 + GreB	0.04 ± 0.008	27 ± 7	19 ± 5	0.11 ± 0.04

Supplemental Table 3.4 Summary of kinetic parameters for pauses prior to termination site D from scaffold-based experiments during active transcription elongation (Fig. 3.6).

Pause location	Rate constant (k), s ⁻¹	Average dwell time (r), s	Pause half-life ($t_{1/2}$), s	Pause efficiency (f)
Pause 177, pH 7.2	0.02 ± 0.006	50 ± 12	34 ± 8	0.08 ± 0.02
Pause 177, pH 8.5	0.03 ± 0.01	34 ± 12	23 ± 8	0.10 ± 0.04
Pause 183, pH 7.2	0.004 ± 0.0009	252 ± 65	175 ± 45	0.06 ± 0.007
Pause 183, pH 8.5	0.01 ± 0.008	80 ± 33	55 ± 23	0.05 ± 0.01
Pause 184, pH 7.2	0.009 ± 0.0002	110 ± 3	76 ± 2	0.05 ± 0.005
Pause 184, pH 8.5	0.03 ± 0.008	37 ± 10	26 ± 7	0.07 ± 0.02

Supplemental discussion

Analysis of the EC translocation register at the 144 pause in direct reconstitution experiments

In this section, we present evidence that ECs at the third pause are backtracked from direct reconstitution experiments (associated with Figs. 3.4 and S3.4). First, when ECs were reconstituted at the second pause and the third pause was encountered by incorporation of UMP, an RNA cleavage product (U16) appeared under both pH conditions, with the U16 RNA fraction increasing more rapidly at pH 8.5 vs. 7.2 (Figs. 3.4B and S3.4C). We propose that following extension of the A17 RNA to U18, the PECs entered a 1-nt backtracked register and cleaved 2 nt from the RNA 3' end, resulting in the observed increase of the U16 RNA fraction over time. It is well established that RNAP can perform intrinsic cleavage via hydrolysis in order to escape a backtrack, with hydrolysis activity increasing as pH alkalinizes⁵, thus explaining the greater increase in the U16 RNA at pH 8.5. However, due to limitations of the nucleic acid scaffold design, this experiment did not conclusively define the backtracked register of ECs at the third pause. Previous analyses of RNAP backtracking demonstrated that the extent of backtracking is precisely limited to 1-nt when the DNA-RNA complementarity is limited to 11-nt⁶. In this experiment, when ECs entered the third pause upon UMP incorporation and adopted the 1-nt backtracked register, there was exactly 11-nt of complementarity in the DNA-RNA hybrid. Therefore, a multi-nt backtracking event was disfavored since the upstream RNA segment contains a poly-U tract that is non-complementary to the template DNA.

In contrast to our results from the UMP incorporation experiment, we observed minimal intrinsic cleavage activity of backtracked PECs when ECs were positioned directly at the third pause to measure the GMP incorporation rate (Fig. 3.4C). For this experiment, the PECs were assembled in the absence of Mg^{2+} to prevent intrinsic cleavage prior to the addition of GTP.

Rather, Mg^{2+} was added alongside GTP and the dominant RNA cleavage product observed during the time course was the A16 RNA, which is 1-nt shorter than the U17 pause RNA (Figs. 3.4C and S3.4D). A much smaller fraction of shorter RNA cleavage products (U15 and C14, Fig. S3.4D) resulting from PEC intrinsic cleavage in 1-nt and 2-nt backtracked registers, compared to those observed in Fig. 3.4B, was observed likely for several reasons. First off, a reduction in the backtracked PEC population following GMP incorporation likely decreased the intrinsic cleavage rate – an observation consistent with a previous analysis of RNAP intrinsic cleavage⁶. Secondly, the nucleic acid scaffold design limits our analysis of the backtracked register since there is decreasing complementarity at the upstream end of the DNA-RNA hybrid as the PECs reverse translocate. Specifically, there is 11-nt of DNA-RNA complementarity when ECs are positioned directly at the third pause and enter the pre-translocated register. When the PECs backtrack by 1-nt, a 10-nt complementary DNA-RNA hybrid remains, with complementarity further decreasing as the PECs access multi-nt backtracked registers. This decrease in complementarity in the DNA- RNA hybrid likely disfavored the backtracked register to a certain extent – an observation consistent with a prior analysis of RNAP backtracking⁶.

The A16 RNA likely resulted from an intrinsic cleavage event with ECs in the pre-translocated register at the third pause (Fig. S3.4D). There are multiple probable explanations for the bias of ECs towards the pre-translocated register in this experiment. For example, since the nucleic acid scaffold disfavors backtracking to a certain extent, it is possible that the equilibria between the backtracked, pre-, and post-translocated states shifted to bias ECs toward the pre-translocated register. Furthermore, the identity of the RNA 3' nucleotide exerts an effect on the RNAP translocation register, with a 3' U favoring the pre-translocated register⁷. Because the RNA 3' nucleotide at the third pause is a U, the RNA 3' sequence identity likely further biased the translocation register equilibria, compounding the effect of decreasing upstream hybrid complementarity. These factors, combined with the decrease in population of

PECs following GMP incorporation, offer a probable explanation for the significant degree of intrinsic cleavage of PECs in the pre-translocated vs. the backtracked register at the third pause. Consistent with this observation, the A16 RNA was also observed (although to a lesser extent) when ECs were reconstituted at the third pause in the absence of the incoming nucleotide (Fig. S3.4E), an experiment described below in further detail.

Analysis of the EC translocation register at the 144 pause in direct reconstitution experiments with Gre factors

To definitively test whether the inactivated ECs observed at the third pause are backtracked, we reconstituted ECs directly at this pause and included the *E. coli* Gre factors GreA and GreB. GreA and GreB exhibit differing cleavage preferences: both factors can efficiently stimulate cleavage of 2-3 nt from the RNA 3' end to rescue RNAP from a 1-nt - 2-nt backtracked state but only GreB can stimulate cleavage of RNA segments greater than 3 nt⁸. In this experiment, ECs were assembled directly at the third pause at pH 7.2 and 8.5, mimicking the pH conditions used for pause escape measurements. To assess both the effect of Mg²⁺ and Gre factors on RNAP intrinsic cleavage at the third pause, the PECs were assembled in the presence and absence of 5 mM Mg²⁺.

At both pH 7.2 and 8.5, in the absence of additional Mg²⁺, we observed minimal RNA species resulting from RNAP intrinsic cleavage (likely due to small amount of Mg²⁺ carried over from RNAP storage buffer, Fig. S3.4E), which is expected due to the established role of Mg²⁺ in promoting intrinsic cleavage through either direct activation of the nucleophile (water) or via an indirect mechanism⁹. Our observation of minimal RNAP intrinsic cleavage in the limiting Mg²⁺ condition is also consistent with reported estimates for the K_D of Mg²⁺ in the endonucleolytic reaction. Although the reported values differ, the K_D for Mg²⁺ in the endonucleolytic reaction is estimated to be at least 10 mM^{10,11}, explaining the small degree of RNAP intrinsic cleavage

without Mg^{2+} supplemented in the reconstitution buffer.

When ECs were reconstituted at the third pause with 5 mM Mg^{2+} , many RNA species resulting from RNAP intrinsic cleavage became more prominent at both pH 7.2 and pH 8.5. A prominent RNA cleavage product was 3-nt shorter than the starting RNA at the pause (C14 RNA, Fig. S3.4E), resulting from intrinsic cleavage by PECs in the 2-nt backtracked register. There was a slightly smaller fraction of the C14 RNA observed at neutral pH: at pH 7.2, the C14 RNA represented ~18% of the total RNA fraction whereas at pH 8.5, the C14 RNA represented ~20% of the total RNA fraction. Besides the C14 RNA, there were additional RNA species resulting from intrinsic cleavage by PECs in a backtracked register ranging from 1- to 9-nt (Fig. S3.4E). The U14 RNA resulting from cleavage by PECs in a 1-nt backtrack represented ~10% of the total RNA at both pH 7.2 and pH 8.5. In contrast, the total fraction of RNA products resulting from intrinsic cleavage in the 3- to 9- nt backtracked registers increased markedly at alkaline pH. At pH 7.2, ~23% of the total RNAs resulted from intrinsic cleavage in the 3- to 9- nt backtracked registers, whereas at pH 8.5, ~42% of the total RNAs resulted from intrinsic cleavage in the 3- to 9- nt backtracked registers. This observation can be attributed to the expected increase in hydrolysis activity by RNAP as pH alkalinizes⁵. Altogether, the observed cleavage patterns at neutral and alkaline pH support that ECs positioned at the third pause sampled a wide range of backtracked registers. The C14 RNA fraction represented the largest fraction of any individual RNA cleavage product at both pH, suggesting that PECs favor the 2-nt backtracked register at the third pause.

When either GreA or GreB are added to PECs reconstituted in the absence of 5 mM Mg^{2+} , we observed a marked increase in only the C14 RNA, resulting from Gre factor-stimulated cleavage of PECs in the 2-nt backtracked register (Fig. S3.4E). We also observed a greater increase in the C14 RNA fraction in the presence of GreB compared to GreA at both pH (Fig.

S3.4E), indicating that GreB more efficiently cleaves PECs in the 2-nt backtracked register in this experiment. At pH 7.2, upon addition of GreA the C14 RNA represented ~12% of the total RNA fraction vs. ~30% with GreB. At pH 8.5, upon addition of GreA the C14 RNA represented ~27% of the total RNA fraction vs. ~40% with GreB. This result suggests that limiting Mg^{2+} decreased PEC sampling of backtracked registers greater than 2-nt. It is possible that Mg^{2+} exerts a stabilizing effect on longer, multi-nt backtracked registers and in its absence, the energetic barrier for accessing these registers increases. Consequently, the PECs enter the most energetically stable backtracked register based on sequence context alone. When the ECs reverse translocate on the destabilizing UA RNA-DNA hybrid at the third pause, the first stabilizing base pair they encounter is the CG pair in the RNA-DNA hybrid at position 14. This repositioning of C14 into the *i* site stabilizes the upstream end of the transcription bubble and further RNAP backtracking is not necessary to achieve an energetically stable configuration, explaining why the 2-nt backtracked register is favored in the limiting Mg^{2+} condition. Based on previous structural analysis of Gre factor-assisted cleavage, both GreA and GreB stimulated intrinsic cleavage in this experiment by providing additional coordination partners for trace Mg^{2+} carried over from RNAP storage buffer that is involved in catalysis and ordering the substrate for cleavage through interaction with backtracked RNA residues⁹.

When either GreA or GreB is added to PECs reconstituted in the presence of 5 mM Mg^{2+} , many RNA species resulting from cleavage in a backtrack became apparent (Fig. S3.4E). In particular, a clear increase in the C14 RNA was observed compared to the no factor 5 mM Mg^{2+} condition (Fig. S3.4E), further supporting Gre factor-assisted cleavage of PECs in the 2-nt backtracked register. Moreover, consistent with our results in the absence of 5 mM Mg^{2+} , we observed a greater increase in the C14 RNA fraction in the presence of GreB compared to GreA (Fig. S3.4E), supporting that GreB more efficiently cleaved PECs in the 2-nt backtracked register

in this experiment. At pH 7.2, the total RNA fraction resulting from cleavage of ECs in a 2-nt backtrack increased from ~18% in the absence of Gre factors to ~34% in the presence of GreA and 44% in the presence of GreB. At pH 8.5, the total RNA fraction resulting from cleavage of ECs in a 2-nt backtrack increased from ~20% in the absence of Gre factors to ~25% in the presence of GreA and ~34% in the presence of GreB. The C14 RNA fraction was generally smaller at alkaline pH vs. neutral pH due to increased RNA cleavage events of ECs sampling 3- to 9- nt backtracked registers, which was likely caused by an increased rate of hydrolysis under alkaline conditions. In addition to stimulating cleavage of ECs in the 2-nt backtracked register, Gre factors facilitated cleavage of ECs sampling 3- to 9- nt backtracks in the presence of 5 mM Mg^{2+} at both pH. Specifically, at pH 7.2 the total RNA fraction resulting from cleavage of ECs in 3- to 9- nt backtracks increased from ~23% in the absence of Gre factors to ~30% in the presence of GreA and ~32% in the presence of GreB. At pH 8.5, the total RNA fraction resulting from cleavage of ECs in 3- to 9- nt backtracks increased from ~42% in the absence of Gre factors, ~53% in the presence of GreA, and ~55% in the presence of GreB. Taken together, our results support that Gre factors facilitated cleavage of ECs sampling a wide range of backtracked registers. Moreover, the C14 RNA fraction represented the largest fraction of any individual RNA cleavage product at both pH in the presence and absence of Gre factors, further supporting that ECs favor the 2-nt backtracked register at the third pause.

A caveat to the above experiment is that ECs were reconstituted directly at the third pause and incubated with Gre factors for 10 min, providing the PECs ample time to sample various translocation registers and populate the most energetically stable configuration. The 2-nt backtrack may not be sampled as extensively when ECs encounter the third pause on a shorter time scale during active transcription elongation (e.g., Fig. 3.4B). Additionally, the extent of PEC backtracking may have been influenced by the scaffold design, as mentioned

previously.

Analysis of the EC translocation register at the 142 pause in direct reconstitution experiments

We also observed an intrinsic cleavage event when ECs were reconstituted at the first pause to measure the pause escape rate, producing the C16 RNA (Fig. S3.4B). We propose that the C16 RNA results from a pre-translocated cleavage event and not a backtracking and cleavage event following nucleotide incorporation to the second pause for multiple reasons. The RNA 3' end at the first pause contains a U, which favors the pre-translocated register⁷. Also, if the C16 RNA resulted from an intrinsic cleavage event following ATP incorporation, the C16 RNA fraction would increase with time as a larger fraction of PECs escape the pause, similarly to what was observed after UTP incorporation at the second pause (Fig. S3.4C).

Analysis of the EC translocation register at the 144 pause in an active elongation experiment

Here, we present evidence that a backtracking event occurs when ECs encounter the third sequential pause during active elongation (associated with Figs. 3.5 and S3.5). In experiments conducted at the GTP limiting condition, upon addition of Gre factors we observed a clear increase in the U16 RNA fraction (Figs. S3.5A) and a concomitant decrease in the U18 pause RNA fraction (Figs. 3.5B and 3.5C). This strongly suggests that the increased U16 RNA fraction resulted from Gre factor-stimulated cleavage of backtracked PECs. When ECs were positioned directly at the third pause in the presence and absence of Gre factors, we observed RNA cleavage products corresponding to a range of backtracked registers and both GreA and GreB clearly stimulated 3- nt cleavage of PECs in the 2-nt backtracked register (Fig. S3.4E). Since the U16 RNA is 2-nt shorter than the U18 RNA, its accumulation during the time course suggests PECs entered a 1-nt backtracked register at the third pause and Gre factors then stimulated a 2-nt cleavage (Figs. 3.5B and C). The U16 RNA remained prominent throughout

the time course since a 2-nt cleavage event repositioned the ECs at the first pause, thus slowing the extension of the U16 RNA. This apparent 2-nt cleavage event suggests that the translocation register equilibrium was affected in the context of active transcription elongation, since Gre factors did not markedly stimulate a 2-nt cleavage when ECs were positioned directly at the third pause (Fig. S3.4E).

Another possibility is that ECs at the third pause entered the 2-nt backtracked register and Gre factors stimulated a 3-nt cleavage; however, the RNA resulting from a 3-nt cleavage event was not directly observed since the ECs rapidly reentered elongation following cleavage. Following resumption of elongation after 3-nt cleavage, the ECs then encountered the first pause, explaining the increase in the U16 RNA fraction. Three lines of evidence support this possibility: (i) when ECs were positioned directly at the third pause, GreB more efficiently stimulated 3-nt cleavage (Fig. S3.4E); (ii) when the third pause was encountered during active transcription elongation, the U16 RNA fraction demonstrated a greater increase in the presence of GreB compared to GreA under both pH conditions (Figs. S3.5A and S3.5B), and (iii) the pause efficiency at the third pause demonstrated a greater increase in the presence of GreB compared to GreA (Figs. 3.5B and 3.5C), supporting that GreB more efficiently promoted RNA cleavage in ECs at the third pause and increased the flux of ECs through the sequential pauses. We also observed further evidence for this 3-nt cleavage event in the UTP limiting condition (10 μ M). Specifically, in the presence of GreB at both pH, we observed an increase in the C15 RNA fraction resulting from a 3-nt cleavage event following a 2-nt backtrack at the third pause (Figs. S3.5C and S3.5D). This increase in the C15 RNA fraction was only apparent in the UTP limiting condition since the incoming NTP at the C15 position is UTP. Also, the C15 RNA fraction exhibited a greater increase in the presence of GreB compared to GreA (Fig. S3.5D). This observation further supports that a subpopulation of PECs at the third pause enter a 2-nt backtracked register and GreB more efficiently stimulates 3-nt cleavage here. Taken together,

our results suggest that GreA and GreB incited both 2-nt and 3- nt cleavage of PECs sampling the 1-nt and 2-nt backtracked registers during active transcription elongation, respectively, and GreB more efficiently promoted 3-nt cleavage. More broadly, our results demonstrate that ECs enter a backtrack at the third pause, which holds potential relevance for PRE RNA folding.

References

1. Nechooshtan, G., Elgrably-weiss, M., Sheaffer, A., Westhof, E. & Altuvia, S. A pH- responsive riboregulator. *Genes Dev.* **23**, 2650–2662 (2009).
2. Larson, M.H., Mooney, R.A., Peters, J.M., Windgassen, T., Dhananjaya, N., Gross, C.A., Block, S.M., Greenleaf, W.J., Landick, R. & Weissman, J.S. A Pause Sequence Enriched at Translation Start Sites Drives Transcription Dynamics *In Vivo*. *Science* **344**, 1042–1047 (2014).
3. Kang, J. Y., Mishanina, T.V., Bellecourt, M.J., Mooney, R.A., Darst, S.A. & Landick, R. RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. *Mol. Cell* **69**, 802-815.e5 (2018).
4. Nechooshtan, G., Elgrably-Weiss, M. & Altuvia, S. Changes in transcriptional pausing modify the folding dynamics of the pH-responsive RNA element. *Nucleic Acids Res.* **42**, 622–630 (2014).
5. Orlova, M., Newlands, J., Das, A., Goldfarb, A. & Borukhov, S. Intrinsic transcript cleavage activity of RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 4596–4600 (1995).
6. Turtola, M., Mäkinen, J. J. & Belogurov, G. A. Active site closure stabilizes the backtracked state of RNA polymerase. *Nucleic Acids Res.* **46**, 10870–10887 (2018).
7. Hein, P. P., Palangat, M. & Landick, R. RNA transcript 3'-proximal sequence affects translocation bias of RNA polymerase. *Biochemistry* **50**, 7002–7014 (2011).
8. Borukhov, S., Sagitov, V. & Goldfarb, A. Transcript cleavage factors from *E. coli*. *Cell* **72**, 459–466 (1993).
9. Abdelkareem, M., Saint-Andre, C., Takacs, M., Papai, G., Crucifix, C., Guo, X., Ortiz, J. & Weixlbaumer, A. Structural Basis of Transcription: RNA Polymerase Backtracking and Its Reactivation. *Mol. Cell* **75**, 298-309.e4 (2019).

10. Sosunova, E., Sosunov, V., Kozlov, M., Nikiforov, V., Goldfarb, A. & Mustaev, A. Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 15469–15474 (2003).
11. Zhang, J., Palangat, M. & Landick, R. Role of the RNA polymerase trigger loop in nucleotidyl transfer, transcript cleavage, and transcriptional pausing. *Nat Struct Mol Biol.* **17**, 99–104 (2010)

7. Acknowledgements

Chapter 3, in full, is a reprint of the material as it appears in Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the *Escherichia coli* pH-responsive riboswitch. Stephen, Christine; Mishanina, Tatiana, Journal of Biological Chemistry, 2022. The dissertation author was the primary investigator and author of this paper.

Chapter 4 Structurally distinct manganese-sensing riboswitch aptamers regulate diverse expression platform architectures

Title:

**Structurally distinct manganese-sensing riboswitch aptamers regulate diverse
expression platform architectures**

Authors:

Christine N. Stephen¹, Danea E. Palmer¹, Clarisa Bautista, and Tatiana V. Mishanina^{1*}

¹Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA,
92093, USA

* To whom correspondence should be addressed. Tel: +1 858 246-5883; Email:

tmishanina@ucsd.edu

0. Abstract

Manganese (Mn)-sensing riboswitches protect bacteria from Mn toxicity by upregulating expression of Mn exporters. The Mn aptamers share key features but diverge in other important elements, including within the metal-binding core. Although X-ray crystal structures of isolated aptamers exist, these structural snapshots lack crucial details about how the aptamer communicates the presence or absence of ligand to the expression platform. In this work, we investigated the Mn-sensing translational riboswitches in *E. coli* (*mntP* and *alx*), which differ in aptamer secondary structure, nucleotide sequence, and pH-dependence of Mn response. We performed co-transcriptional RNA chemical probing, allowing us to visualize RNA folding intermediates that form and resolve *en route* to the final folded riboswitch. For the first time, we report that sampling of metal ions by the RNA begins before the aptamer synthesis and folding are complete. At a single-nucleotide resolution, we pinpoint the transcription window where “riboswitching” occurs in response to Mn binding and uncover key differences in how the *alx* and

mntP riboswitches fold. Finally, we describe riboswitch-specific effects of pH, providing insights into how two members of the same riboswitch family differentially sense two distinct environmental cues: concentration of Mn and pH.

1. Introduction

Transition metal ions play key cellular roles (e.g., serving as enzyme cofactors) but become harmful in excess due to misincorporation into biomolecules and metal-catalyzed formation of reactive oxygen species(1, 2). To match the need for metal ion transporters, cells use the metal's intracellular concentration to regulate expression of the metal's transporters at the transcription and/or translation level(3). Examples of this regulation include metal-binding transcription regulators(4, 5) or direct interaction of the metal with a riboswitch encoded in the mRNA(6–8). Riboswitches are 5'-untranslated regions of mRNA that promote or inhibit expression of the downstream gene by changing their folding upon binding a ligand(9). A riboswitch is comprised of a ligand-binding aptamer overlapping with an expression platform, whose folding determines the transcriptional and/or translational fate of the mRNA(9). Many classes of metal-sensing riboswitches have been identified, such as magnesium(7), nickel(8), cobalt(8), and manganese (Mn)(6, 10).

The *yybP-ykoY* Mn-sensing riboswitch family is the largest metal-sensing family, with >1,000 unique representatives in different bacterial phyla(11). Crystal structures, single-molecule FRET (smFRET), and molecular dynamics (MD) simulations of isolated Mn-sensing aptamers provided valuable insights into the overall RNA fold, metal binding site, and RNA dynamics in the presence or absence of Mn(12, 13). However, these studies lack the expression platform of the riboswitch – the “business end” that determines the gene expression outcome. Consequently, it remains unknown how Mn binding to the aptamer actuates the differential folding of the downstream RNA to affect its transcription or translation. Further, studies to date were performed

on pre-synthesized RNA, missing the aspect of active transcription that is critical to riboswitch folding(14–17).

We addressed these gaps by investigating the two Mn-sensing translational riboswitches in *E. coli* which regulate translation of the *alx* and *mntP* genes, with the *alx* riboswitch upregulating gene expression in response to alkaline pH and increased Mn concentration ([Mn])(18, 19) (Fig. 4.1A). Binding of Mn to either the *mntP* or *alx* aptamers promotes a translationally active conformation in the expression platform, increasing expression of the Alx and MntP Mn exporters(10, 19, 20) (Fig. 4.1A-B). Multiple structures of Mn-sensing riboswitch aptamers established their core structural features, consisting of two coaxially stacked “legs” connected by a four-way junction (4WJ)(6, 12, 21) (Fig. 4.1C). The right leg (L1, P1.1, P1.2, and P2; Fig. 4.1C) extrudes a conserved adenosine in L1 to form a cross-helix A-minor interaction with P3.1 in the left leg (L3, P3.1, P3.2, and P4; Fig. 4.1C). The L3 loop in the left leg stacks below the conserved adenosine in the A-minor interaction, forming a phosphate-rich metal-binding hot spot (Fig. 4.1C). X-ray crystal structures show two metal-binding sites in the aptamer, with one occupied by Mg^{2+} (M_A) and the other by Mn^{2+} (M_B)(6, 12, 21) (Fig. 4.1C). The P1.1 “switch helix” is a key element of the *yybP-ykoY* aptamer as its 3'-half alternatively pairs within the aptamer or the expression platform depending on whether Mn is bound, with consequences to folding of the rest of the expression platform.(12) (see Fig. 4.1B for *alx* and *mntP* alternative P1.1 pairing).

Despite sharing conserved features for Mn-sensing, the *alx* and *mntP* riboswitches differ in their global architectures (Fig. 4.2A). The *alx*-like riboswitches are exclusively present in the *Enterobacteriales* order, members of which are often found in environments that can alkalinize, such as water, soil, and in association with other organisms(22, 23). Upon alkaline pH challenge, the *E. coli* cytosol alkalinizes and then recovers to pH 7.8(24); in turn, cytosol alkalinization increases cytoplasmic [Mn](19), potentially explaining *alx* translational activation at alkaline pH. Initial studies of the *alx* and *mntP* riboswitches put forward alternative secondary structure models

based on chemical probing with pre-synthesized RNA(10, 18), prompting us to ask (i) how these alternative structures fold and respond to Mn during transcription and (ii) why alkaline pH uniquely affects *alx* structure. To answer these fundamental questions, we applied modern RNA chemical probing methods to *alx* and *mntP*, where ribonucleotides are modified by either a selective 2'-hydroxyl (SHAPE) reagent or dimethyl sulfate (DMS) and detected by high-throughput mutational profiling (MaP)(25).

In this work, we performed DMS-MaP in a co-transcriptional format using established methods(26, 27), getting to the heart of how Mn binding to the aptamer is “communicated” to the expression platform. Unexpectedly, we uncovered evidence that both *alx* and *mntP* aptamer folding intermediates begin sampling for divalents before the full aptamer domain has emerged from RNAP, challenging the notion that a fully folded aptamer is required for ligand sensing. Further, a critical RNA loop that participates in Mn sensing diverges in sequence between *alx* and *mntP* and may hold the key to how pH differentially affects riboregulation by these riboswitches. Using gene reporter assays, we show that alkaline pH alone cannot rearrange the *alx* riboswitch into a translationally active conformation; instead, the *alx* riboswitch senses the increased [Mn] brought about by alkaline pH. Lastly, our co-transcriptional chemical probing data uncovered distinct structural features in the *alx* and *mntP* expression platforms, which are potentially coupled to the cellular need for the Alx and MntP exporter proteins under increased cytosolic [Mn].

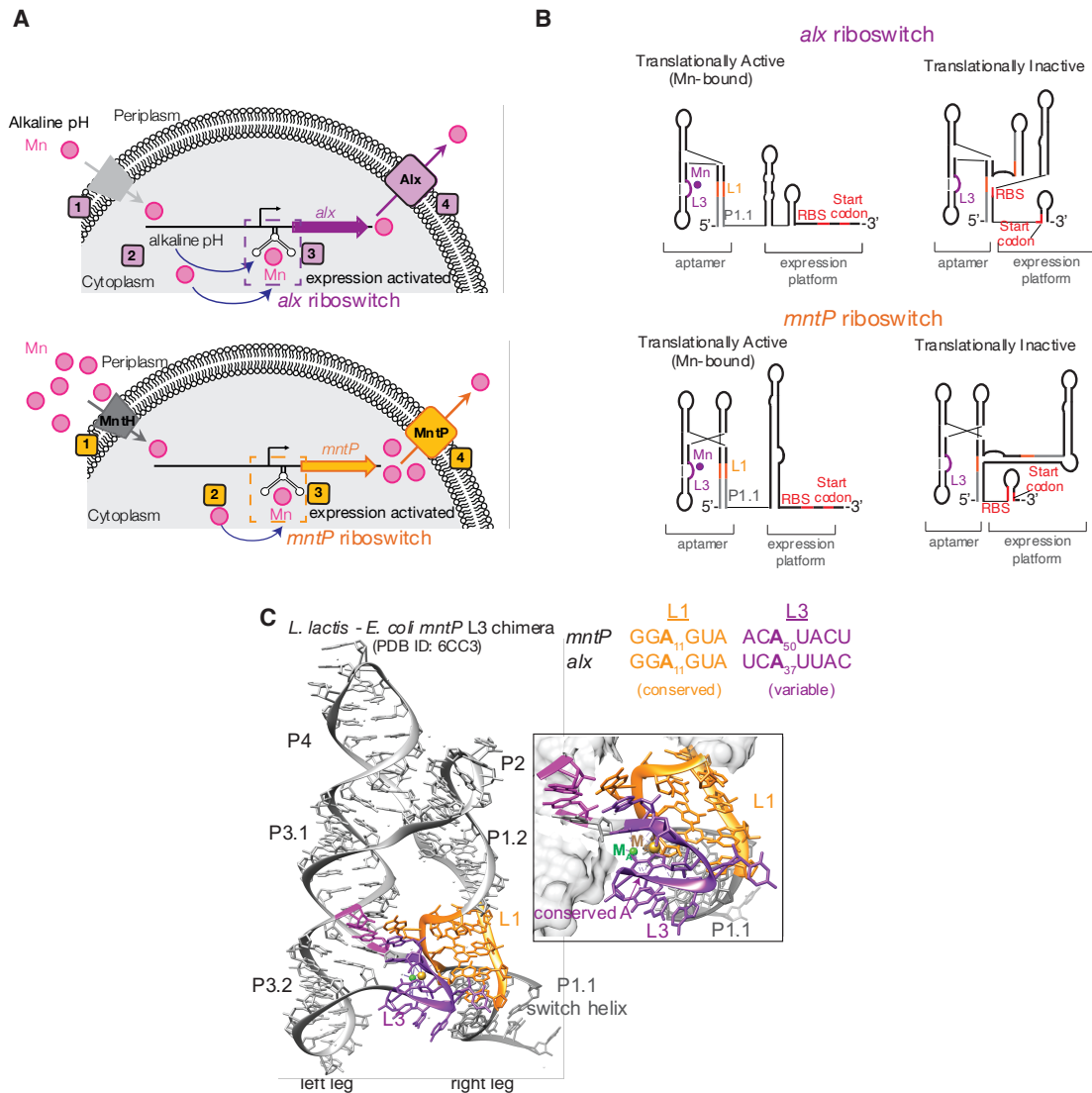


Figure 4.1 Introduction to the *E. coli* *alx* and *mntP* Mn-sensing translational riboswitches.

A. Cartoon model of *alx* and *mntP* riboregulation of their respective genes in *E. coli*. (Top panel) At alkaline pH, the cytosolic [Mn] increases in *E. coli*, resulting in the *alx* riboswitch turning “on” expression of Alx, an Mn exporter. (Bottom panel) An increase in environmental [Mn] results in the *mntP* riboswitch turning “on” expression of MntP, the major Mn exporter in *E. coli*. **B.** Cartoon models of the *alx* and *mntP* riboswitches in their translationally active vs. inactive conformations. L3 (purple), L1 (orange), P1.1 (gray), and RBS/start codon (red) are highlighted. **C.** *L. lactis* Mn-sensing aptamer crystal structure with the *E. coli* *mntP* L3 sequence (PDB ID: 6CC3). “Left leg” and “right leg” refer to coaxially stacked helices in the aptamer. L3 (purple), L1 (orange), P1.1 (gray), and G-C receptor base pair in cross-helix A-minor (magenta) are highlighted. Mg²⁺ in the M_A site and Cd²⁺ in the M_B site (captured instead of Mn²⁺) are green and yellow spheres, respectively. Sequence alignment of *alx* and *mntP* L1 and L3 are orange and yellow, respectively.

2. Materials and Methods

The source and catalog number (if applicable) for all key reagents are provided as a Supplementary Table 4.1.

All oligonucleotides were purchased from either Integrated DNA Technologies or Eton Bioscience. A summary of all oligonucleotides used in this work including sequences, modifications, and purifications is presented in Supplementary Table 4.2.

Procedures for protein preparations, construction of plasmids and bacterial strains, and preparation of DNA templates are provided in the Supplemental Methods. Details for all DNA templates, strains, and plasmids are tabulated in Supplementary Tables 4.3-4.6.

Single-round *in vitro* transcription for TECprobe-VL experiments

To prepare the holoRNAP, *E. coli* RNAP (6 μ M) and σ^{70} (12 μ M) were combined in 1x TECprobe transcription buffer (100 mM Tris-Cl, pH 7.2 or 8.5, 50 mM KCl, and 1 mM DTT) and incubated for 30 min at 37 °C. The single-round *in vitro* transcription reactions for TECprobe-VL experiments were assembled by first combining 1x transcription buffer, 40 nM DNA template (*alx* riboswitch, Template 3, Supplementary Table 4.3) or 50 nM DNA template (*mntP* riboswitch, Template 6, Supplementary Table 4.3) with randomly distributed internal biotin-11 dNMPs (AAT BioQuest/Biotium), 0.5 μ M holoRNAP, and 500 μ M ATP, UTP, CTP, and GTP. The transcription buffer omitted magnesium, which was a part of the start solution added later to initiate transcription elongation. Tris-Cl was added at a final concentration of 100 mM to minimize pH changes during the chemical probing reaction, as described previously(27). The reaction was incubated for 10 min at 37 °C to form open promoter complexes. Next, SAVPhire Monomeric Streptavidin (Sigma-Aldrich) was added to the reaction at a final concentration of 5 μ M, followed by incubation for 10 min at 37 °C to allow streptavidin binding to biotinylation sites on the DNA template. Transcription was initiated by adding 10X start solution (100 mM MgCl₂, 100 μ g/ml rifampicin) to the reaction to

a final concentration of 10 mM MgCl₂ and 10 µg/ml rifampicin to inhibit reinitiation. After addition of start solution, the final reaction volume was 60 µl. To test the effect of the Mn ligand on riboswitch structural rearrangements, MnCl₂ was added to a final concentration of 1 mM in the transcription reactions and was introduced alongside the start solution. The transcription reaction was incubated for 30 s at 37 °C before proceeding to the chemical probing, which is described in the next section.

Single-round *in vitro* transcription for the single-length TECprobe-SL experiments followed the same protocol as outlined above for variable length TECprobe-VL experiments. The DNA templates used for TECprobe-SL experiments contained a terminal biotin modification that was introduced by a 5' biotinylated reverse primer in the PCR (*alx* riboswitch, Template 7, Supplementary table 4.3).

RNA chemical probing during TECprobe-VL experiments

Benzoyl cyanide (BzCN, Sigma-Aldrich) probing was performed by splitting the 60-µl transcription reaction into 25 µl aliquots and mixing with either 2.78 µl of 400 mM BzCN dissolved in anhydrous DMSO (Invitrogen) to a final concentration of 17.7 mM [(+) sample] or with anhydrous DMSO [(-) sample]. Since the BzCN probing reaction is complete within ~1 s, no additional incubation was required before quenching the reactions with 250 µl of TRIzol reagent (Invitrogen).

DMS probing was performed by splitting the 60-µl transcription reaction into 25 µl aliquots and mixing with either 2.78 µl of 6.5% v/v DMS (Sigma-Aldrich) in 200 proof ethanol [(+) channel] or with 200 proof ethanol [(-) channel]. The samples were incubated for 5 min at 37 °C. The DMS probing reaction was quenched by adding 6.67 µl of 14.5 µM 2-mercaptoethanol and incubating the sample at 37 °C for 1 min. *In vitro* transcription reactions were then quenched by addition of 250 µl of TRIzol reagent.

RNA purification post-TECprobe-VL probing

The RNA was purified from the quenched *in vitro* transcription reactions using the following protocol. First, 50 μ l of chloroform was added to the samples and the samples were mixed by vortexing followed by centrifugation at 18,000g and 4 °C for 5 min. Next, the aqueous layer was transferred to a new tube and nucleic acids were precipitated by addition of 105 μ l cold isopropanol and 1 μ l GlycoBlue Coprecipitant (Invitrogen). The samples were incubated at room temperature for 15 min and centrifuged at 18,000g and 4 °C for 15 min. The supernatant was discarded and the pellet containing the precipitated nucleic acids was washed by addition of 200 μ l 70% ethanol. Following the 70% ethanol wash, the samples were centrifuged at 18,000g and 4 °C for 2 min and the supernatant was discarded. The pelleted nucleic acids were resuspended in 25 μ l of 1x DNase buffer (10 mM Tris-Cl, pH 7.5, 2.5 mM MgCl₂) and mixed with 0.75 μ l of Turbo DNase (Invitrogen) to remove residual DNA template. The samples were incubated at 37 °C for 15 min and DNase digestion was quenched by addition of 250 μ l TRIzol reagent (Invitrogen). A second TRIzol extraction was performed as described above, except the pellet containing the DMS-modified RNA transcripts was resuspended in 5 μ l of 10% (v/v) DMSO.

RNA 3' adapter ligation onto the TECprobe-VL purified transcripts

The 9N_VRA3 adapter oligonucleotide (0139, Supplementary Table 4.2) was pre-adenylated with the 5' DNA Adenylation Kit (New England Biolabs) using the following protocol. A 100 μ l master mix was prepared that consisted of the 1X DNA Adenylation Buffer (New England Biolabs), 100 μ M ATP, 5 μ M 9N_VRA3 oligo, and 5 μ M Mth RNA Ligase (New England Biolabs). The 100 μ l master mix was split into 4 aliquots (25 μ l each) in PCR tubes and incubated in a thermal cycler set to 65 °C for 1 h. To clean up the adenylation reactions, reactions were pooled to 50 μ l followed by addition of 500 μ l TRIzol (Invitrogen) and 100 μ l of chloroform to each 50 μ l pool. The samples were vortexed to mix and centrifuged at 18,000g and 4 °C for 5 min. The aqueous phase was transferred to a new tube and the adapter was ethanol precipitated. The

pellet was resuspended in Tris-Cl, pH 8.0, and its concentration was quantified with the Qubit ssDNA Assay Kit (Invitrogen) and a Qubit 3 Fluorometer. The pre-adenylated adapter was diluted to 1 μ M, divided into small aliquots, and stored at -20 °C.

The RNA 3' adapter ligation reactions were performed using the following protocol. Purified RNA redissolved in 5 μ l of 10% (v/v) DMSO was combined with 1x T4 RNA Ligase Buffer (New England Biolabs), 0.5 U/ μ l SuperscriptaseIN (Invitrogen), 0.05 μ M pre-adenylated 9N_VRA3 adapter, 15 % PEG 8000, and 5 U/ μ l T4 RNA Ligase 2, truncated KQ (New England Biolabs). The final volume for the ligation reactions was 20 μ l and the reactions were incubated at 25 °C for 2 h.

SPRI bead purification of adapter-ligated RNA

A solid-phase reversible immobilization (SPRI) bead purification protocol was used to deplete excess 9N_VRA3 adapter oligonucleotide. The SPRI beads were prepared with SpeedBead Magnetic Carboxylate Beads (Cytiva) using the protocol from Jolivet and Foley(28). The ligation reactions were combined with 17.5 μ l of water, 40 μ l anhydrous isopropanol, and 22.5 μ l of the SPRI beads. The samples were vortexed and incubated at room temperature for 5 min. The samples were then placed on a magnetic separation rack and the beads collected on the tube wall before disposing of the supernatant. Keeping the tubes on the magnetic stand, the beads were washed twice with 200 μ l of 70% ethanol. After discarding the second ethanol wash, the tubes were placed in a dry bath set to 37 °C to evaporate the residual ethanol. The beads were resuspended in 25 μ l of elution buffer (20 mM Tris-Cl pH 8.0, 50 mM KCl, 1 mM DTT, 0.1 mM EDTA, 0.1% Triton X-100, 10 mM MgCl₂) and the samples were incubated for 3 minutes at room temperature to elute the RNA. The samples were placed on a magnetic rack and beads collected on the tube wall before aliquoting the supernatant containing the RNA into a fresh tube.

TECprobe-VL cDNA synthesis and cleanup

The reverse transcription reactions were performed via capture of RNA transcripts on streptavidin beads using a 5' biotinylated reverse transcription primer (0159, Supplementary table 4.2), which anneals to the 9N_VRA3 sequence that was appended to the intermediate riboswitch transcripts. Dynabeads MyOne Streptavidin C1 beads (Invitrogen) were equilibrated as described previously(27). 5 μ l of 500 nM RT primer was combined with the purified RNA samples and annealed in a thermal cycler using the following program: 70 °C for 5 min, ramp to 50 °C at 0.1 °C /s, 50 °C for 5 min, ramp to 40 °C at 0.1 °C /s, 40 °C for 5 min, cool to 25 °C. The equilibrated streptavidin beads were placed on a magnetic stand to allow the beads to collect on the tube wall and the supernatant was discarded. The streptavidin beads were then resuspended with the samples containing RNA annealed to the 5' biotinylated RT primer and incubated at room temperature for 15 min on an end-over-end rotator. The samples were placed on a magnetic stand and the beads were allowed to collect on the tube wall before discarding the supernatant. The beads were resuspended in 19.5 μ l of reverse transcription master mix and incubated in a thermal cycler set to 42 °C for 2 min before addition of the reverse transcriptase. Upon addition of 0.5 μ l of SuperScript II Reverse Transcriptase (Invitrogen) to each sample, the 20 μ l RT reactions contained 1x first strand buffer (50 mM Tris-Cl, pH 8.0, 75 mM KCl), 500 μ M dNTPs, 10 mM DTT, 2 % formamide, 10 ng/ μ l Extreme Thermostable Single-Stranded DNA Binding Protein (ET SSB) (New England Biolabs), 0.1 % Triton-X-100, and 3 mM MnCl₂. The reverse transcription reactions were incubated at 42 °C for 180 min followed by a 70 °C incubation for 50 min to inactivate the SuperScript II and then cooled to 12 °C. Next, the RNAs were degraded by adding 1.25 U RNase H (New England Biolabs) and 12.5 U RNase I_f (New England Biolabs) to each sample and the samples were incubated at 37 °C for 20 min followed by a 70 °C incubation for 20 min to inactivate the RNases. The samples were placed on a magnetic stand and the beads were allowed to collect on the tube wall. The supernatant was discarded and the beads were

resuspended in 75 μ l of Tris Wash Buffer (10 mM Tris-Cl, pH 8.0, 0.05% Triton X-100) to wash the beads. The samples were placed on a magnetic stand and the wash was discarded. The bead-bound cDNA was resuspended in 25 μ l of the Tris Wash Buffer and stored at -20 °C.

Chemical probing of equilibrium-refolded intermediate transcripts

Single-round *in vitro* transcription using randomly biotinylated DNA templates was executed as described for TECprobe-VL experiments above. To quench transcription, 600 μ l of TRIzol was added to the reactions. Next, 120 μ l of chloroform was added to the samples and the samples were vortexed to mix followed by centrifugation at 18,000g and 4 °C for 5 min. The aqueous layer was transferred into a new tube and nucleic acids were isopropanol precipitated as described for TECprobe-VL experiments except a larger volume of isopropanol was used to account for the larger sample volume. The pellets containing the intermediate RNA transcripts were resuspended in 30 μ l of 1x TECprobe transcription buffer (100 mM Tris-Cl, pH 7.2 or 8.5, 50 mM KCl, 1 mM DTT), placed on a heat block set to 95 °C for 2 min, and snap-cooled on ice for 1 min. To each sample, 30 μ l of 2x equilibration buffer (200 mM Tris-Cl, pH 7.2 or 8.5, 100 mM KCl, 2 mM DTT, 20 mM MgCl₂) was added, which contained either no Mn or 2 mM MnCl₂ (1 mM final in 60 μ l reaction), and the samples were incubated at 37 °C for 20 min. The RNAs were then probed with DMS as described for TECprobe-VL experiments, and the DMS-probed RNA was purified as described for TECprobe-VL experiments.

Preparation of TECprobe-VL dsDNA libraries for sequencing

cDNA libraries were PCR-amplified for high throughput sequencing using the following protocol. Separate PCR reactions were performed for the (+) and (-) samples, which contained 1x Q5 Reaction Buffer, 1x Q5 GC Enhancer, 200 μ M dNTPs, 250 nM RPI Indexing Primer (0163-0178,0235-0248, Supplementary table 4.2), 250 nM dRP1_NoMod.R reverse primer (0162, Supplementary Table 4.2), 10 nM SC1Brdg_MINUS or SC1Brdg_PLUS channel barcode oligo (0160,0161, Supplementary Table 4.2), 12 μ l of bead-bound cDNA library, and 0.02 U/ μ l Q5 High-

Fidelity DNA Polymerase (New England Biolabs). Amplification was performed with the following thermal cycler program: 98 °C for 30 s, [98 °C for 10s, 62 °C for 20 s, 72 °C for 20s] x 22 cycles, 72 °C for 2 min, hold at 12 °C.

High-throughput DNA sequencing of TECprobe-VL libraries

Sequencing of the TECprobe-VL and equilibrium refolded intermediate transcript libraries was performed by Novogene Co. on a Novaseq X Plus system with 2 x 150 Paired End (PE) reads and a 20% PhiX spike in. The TECprobe-VL and equilibrium refolded intermediate transcript libraries were sequenced at a depth of ~30-60 million PE reads.

Sequencing of TECprobe-SL libraries was performed by Novogene Co. on a Novaseq X Plus system with 2 x 150 PE reads. No PhiX spike in was added since these libraries were sequenced in a partial lane. The TECprobe-SL libraries were sequenced at a depth of ~10 million PE reads. Sequencing data for TECprobe-VL and TECprobe-SL experiments is publicly available in the Sequencing Read Archive (SRA) under project number PRJNA1197522 and BioSamples are tabulated in Supplementary Table 4.7.

TECprobe-VL data processing and visualization

Sequence read pre-processing, target generation using `cotrans_preprocessor`, and `shapemapper2` run script generation were performed exactly as described in the published protocol using the publicly available TECprobe data processing tools (<https://github.com/e-strobel-lab/TECtools>). TECprobe visualization tools was used to assemble reactivity matrices and analyze the correlation between replicate data sets (https://github.com/e-strobel-lab/TECprobe_visualization). The `compile_sm2_output` script was used to assemble the reactivity profiles for intermediate transcripts into matrix format and to normalize the dataset. The normalized reactivity matrices were visualized in heatmap format using GraphPad Prism.

For the *alx* riboswitch, replicate data sets were required to obtain adequate sequencing depth for longer transcripts, as there was a strong bias in representation towards shorter

transcripts for the amplified cDNA libraries. Plots with fraction of aligned reads per transcript length and alignment rates for *a/x* co-transcriptional TECprobe-VL experiments are presented in SI Fig. 4.21. Additionally, the presence of two intrinsic termination sites in the *a/x* riboswitch sequence (110 and 190 nt) results in a clear drop in sequencing depth for intermediate transcripts past these sites (SI Fig. 4.19A). For the co-transcriptional DMS-MaP libraries, 5-7 independent replicates were pooled per chemical probing condition to increase data quality for longer transcript lengths (see SI Figs 4.22-4.25 for individual and concatenated heatmaps). Plots with fraction of aligned reads per transcript length and alignment rates for *a/x* equilibrium refolded TECprobe-VL experiments are presented in SI Fig. 4.26. For the equilibrium refolded *a/x* TECprobe-VL experiments, 2 independent replicates were pooled per chemical probing condition to increase data quality for longer transcript lengths (see SI Figs 4.27-4.30 for individual and concatenated heatmaps). To further increase sequencing depth when necessary, multiple independent PCR amplifications were performed with the same cDNA library and sequenced; reads were subsequently pooled together for data processing.

For the *mntP* riboswitch, two replicate data sets were concatenated and analyzed together after assessing the correlation between replicates using the `plot_cotrans_correlation` script available in TECprobe visualization tools. Plots with fraction of aligned reads per transcript length and alignment rates for co-transcriptional *mntP* TECprobe-VL experiments are presented in SI Fig. 4.31. Plots with fraction of aligned reads per transcript length and alignment rates for equilibrium refolded *mntP* TECprobe-VL experiments are presented in SI Fig. 4.32. To further increase sequencing depth when necessary, multiple independent PCR amplifications were performed with the same cDNA library and sequenced; reads were subsequently pooled together for data processing.

***In vitro* synthesis of full-length riboswitch RNAs with T7 RNAP and cleanup for SHAPE-MaP and DMS-MaP**

RNA for *in vitro* SHAPE/DMS-MaP was synthesized using the HiScribe T7 High Yield RNA Kit (New England Biolabs). 100 μ L reactions were assembled according to manufacturer instructions and incubated at 37 °C overnight. RNA was purified and concentrated using the RNA Clean and Concentrator kit (Zymo Research). For the *alx* riboswitch and *alx* aptamer, RNAs were gel-purified to isolate full-length transcripts away from premature transcription termination products using a denaturing PAGE (8% acrylamide, 8 M urea, 0.5x TBE (Fisher Scientific)). For gel-purification, RNAs were diluted 1:5 in buffer and mixed with an equal volume 2x formamide loading dye (95% formamide, 0.5 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol), heated at 95 °C for 2 mins, and loaded on a pre-heated gel. Gel was run at 50 W, stained with 1x SYBR Green II RNA Gel Stain (Invitrogen) in 1X TBE for 10 mins, and visualized using a Edvotek TruBlu 2 Blue Transilluminator. Bands corresponding to full-length RNA were excised and eluted in 10 mL TE buffer (10 mM Tris pH 8.0, 1 mM EDTA pH 8.0) at 4 °C overnight. The eluted solution of RNA was spun to pellet acrylamide, and RNA was concentrated using an Amicon Ultra Centrifugal filter (3 kDa molecular weight cutoff). RNAs were ethanol precipitated, stored at -20 °C, and resuspended in RNase-free H₂O. For the *mntP* riboswitch, there are no terminated transcripts, eliminating the need for gel purification. After transcription and cleanup, *mntP* RNAs were ethanol precipitated, stored at -20 °C, and resuspended in RNase free H₂O.

SHAPE-MaP and DMS-MaP on the full-length equilibrium refolded *alx* and *mntP* riboswitches

5 pmol of RNA in added RNase-free H₂O to a volume of 6 μ L was heat denatured for 2 mins at 95 °C, cooled at 4 °C for 2 mins, and then 3 μ L buffer was added for a final volume of 9 μ L. The final buffer concentration was 100 mM HEPES, 100 mM NaCl, and 5 mM MgCl₂ at either pH 7.2 or pH 8.4 and supplemented with 1 mM MnCl₂ where indicated. RNA was incubated at 37

°C for 20 mins to allow refolding. For modification with 1M7, 9 μ L of RNA was mixed with 1 μ L of 100 mM 1M7 (“+” modified sample) or 1 μ L DMSO (“-” negative control), incubated for 75 s at 37 °C, and placed on ice. For the denatured control, 5 pmol RNA (3 μ L) was mixed with 5 μ L formamide and 1 μ L 10x denaturing buffer (500 mM HEPES, 40 mM EDTA) at pH 7.2 or 8.4. RNA was denatured for 1 min at 95 °C, mixed with 1 μ L of 100 mM 1M7, incubated at 95 °C for 1 min, and placed on ice. All RNAs were purified and eluted in 50 μ L of RNase-free H₂O using MicroSpin G-25 columns (Cytiva) and stored at -20 °C. For modification with DMS, 9 μ L of RNA was mixed with 1 μ L of 6.5% v/v DMS (“+” modified sample) or 1 μ L ethanol (“-” negative control), incubated for 5 minutes at 37 °C, quenched with 2.4 μ L of 30% v/v 2-mercaptoethanol solution at 37 °C for 1 min before placing on ice. For the denatured control, 5 pmol RNA (3 μ L) was mixed with 5 μ L formamide and 1 μ L 10x denaturing buffer (500 mM HEPES, 40 mM EDTA) at pH 7.2 or 8.4. Then RNA was denatured for 1 min at 95 °C, mixed with 1 μ L of 6.5% v/v DMS, incubated at 95 °C for 1 min, quenched with 2.4 μ L of 30% 2-mercaptoethanol solution at 37 °C for 1 min before placing on ice. All RNAs were purified and eluted in 50 μ L of RNase-free H₂O using MicroSpin G-25 columns (Cytiva) and stored at -20 °C.

Reverse transcription (RT) of 10 μ L of RNA was performed by annealing 1 μ L RT primer (0277) for 5 min at 65 °C. Then, 8 μ L RT buffer was added (50 mM Tris-Cl pH 8, 75 mM KCl, 10 mM DTT, 0.5 mM dNTPs, and MnCl₂ added to 15 mM before use). Samples were incubated for 2 min at 42 °C before introducing Superscript II reverse transcriptase (Thermo Fisher Scientific). RT reactions were incubated for 3 h at 42 °C. The reverse transcriptase was inactivated for 15 min at 70 °C before storing the reactions at 4 °C. cDNA was purified and eluted in 50 μ L of RNase-free H₂O using MicroSpin G-25 columns (Cytiva) and stored at -20 °C. The first PCR (step 1) amplified cDNA through 20 amplification cycles using primers 0278 and 0279. Excess primer was removed using the Monarch PCR & DNA Cleanup Kit (New England Biolabs). The second PCR (step 2) further amplified the first PCR amplicon through an additional 10 cycles. The step 2 PCR

used primers 0149 and an RPI indexing primer (0163-0178 and 0235-0248). Step 2 PCR was cleaned up using Agencourt AMPure XP beads (Beckman Coulter), and final DNA concentration was measured using the Invitrogen Quant-IT dsDNA High-Sensitivity HS Assay kit quantified with a BioTek Synergy H1 Multimode Microplate Reader before pooling libraries for sequencing. Sequencing of these pooled libraries was performed by Novogene Co. on a NovaSeq 6000 system with 100M 150 PE reads. Sequencing data for SHAPE-MaP and DMS-MaP experiments is publicly available in the Sequencing Read Archive (SRA) under project number PRJNA1197522 and BioSamples are tabulated in Supplementary Table 4.8.

RNA structure prediction and visualization

RNA structure prediction for co-transcriptionally probed riboswitch folding intermediates was performed using the RNAstructure V6.4(29) Fold command with default settings, adding the DMS outputs as a constraint, and forcing the 3'-most 11 to 14 nt to be single-stranded to account for the *E. coli* RNAP footprint. The set of minimum free-energy structures were visualized in VARNA(30) to assess agreement with the DMS probing data, and structure diagrams were colored in accordance with DMS reactivity. Additionally, existing X-ray crystal structures for Mn-sensing aptamer domains were visualized with UCSF Chimera(31) and used to guide modeling of conserved secondary and tertiary structure motifs within the *alx* and *mntP* aptamers. RNA structure prediction for equilibrium refolded riboswitch folding intermediates was performed as described above; however, the 3' most nt were not kept single-stranded since the intermediate RNA transcripts were extracted from stalled transcription elongation complexes before DMS probing.

The denatured and equilibrium-refolded full-length *mntP* riboswitch, *alx* riboswitch, and *alx* aptamer RNA structures were predicted using the RNAstructure V6.4(29) Fold command with default settings, adding the SHAPE or DMS outputs as a constraint. The minimum free energy structures were visualized in VARNA(30) and colored in accordance with nucleotide reactivity.

β -galactosidase assays

Overnight culture of each strain was diluted 1:100 into LBK with appropriate antibiotic at either pH 6.8 or pH 8.4, with or without 1 mM supplemented MnCl_2 . The cultures were grown to mid-log phase and supplemented with 1 mM IPTG where necessary for expression of T7 RNAP. β -galactosidase assays were performed using the Miller method(32). Briefly, bacterial growth was measured using OD_{600} before cultures were placed on ice. 100 μL cells were lysed via vortexing in 900 μL sodium phosphate buffer (100 mM sodium phosphate pH 7, 10 mM KCl, 2 mM MgSO_4 , 38 mM 2-mercaptoethanol), 20 μL 0.1% SDS, and 20 μL chloroform. Substrate *O*-nitrophenyl-b-D-galactopyranoside (ONPG) was added and the reaction was stopped with 1 M Na_2CO_3 when a yellow color appeared, or after 35 min. After centrifuging 5 min at 16,000 g, absorbance was measured at 420 nm (A_{420}) to calculate activity using Miller units. Each culture was tested in duplicate and A_{420} values were averaged. Each final reported value is an average of three replicates with standard deviation.

Sequence analysis for *yybP-ykoY* riboswitches

The sequences for *yybP-ykoY* riboswitches were downloaded from the Rfam database (Rfam reference number RF00080) and aligned using the Kalign multi sequence alignment program (<https://www.ebi.ac.uk/jdispatcher/msa>). The sequence alignment and consensus L3 sequence was visualized using Jalview(33). The *yybP-ykoY* riboswitch sequences found within the Enterobacteriales order were accessed from both the Rfam database and from *yybP-ykoY* regulon collection (https://regprecise.lbl.gov/collection_rfam.jsp?riboswitch_id=13). These sequences were aligned with the Muscle multi sequence alignment program (<https://www.ebi.ac.uk/jdispatcher/msa>). The sequence alignment and consensus L3 sequence was visualized using Jalview(33).

3. Results

The Mn-sensing aptamers of the *alx* and *mntP* riboswitches differ structurally

In Fig. 4.2A, we present secondary structure models of the *alx* and *mntP* riboswitch aptamer domains based on prior experimental data for the full-length riboswitches(10, 18), which we updated to reflect the overall aptamer architecture and conserved base-pairs identified by X-ray crystal structures(6, 12, 21). The model of the *mntP* aptamer in Fig. 4.2A is supported by an X-ray crystal structure of the aptamer in the apo state (PDB ID: 4Y1M)(6). Since no detailed structural information exists for the *alx* aptamer, we mapped its secondary structure with DMS/SHAPE-MaP using the traditional approach of *in vitro* synthesis by T7 RNA polymerase (RNAP) followed by denaturation and refolding.

pH and Mn effect on the alx aptamer structure. The aptamer RNA was refolded at either neutral (7.2) or alkaline pH (8.5), consistent with pH values used in published work for DMS probing, *in vitro* transcription, and gene reporter fusion experiments(18, 19, 34). The aptamer RNAs were probed with either 1-methyl-7-nitroisatoic anhydride (1M7) or DMS. The resulting secondary structure models of the *alx* aptamer agree with the published 2D predictions(18) and 3D structure of an *L. lactis* chimera with the *alx* L3 sequence(21) (Fig. 4.2B and SI Fig. 4.1A-C). Comparison of the nucleotide (nt) reactivities as a function of refolding pH revealed clear flexibility differences in L3 and L1, with multiple nt (e.g., A37 and A11) showing decreased reactivity towards 1M7 or DMS at alkaline pH (Fig. 4.2C).

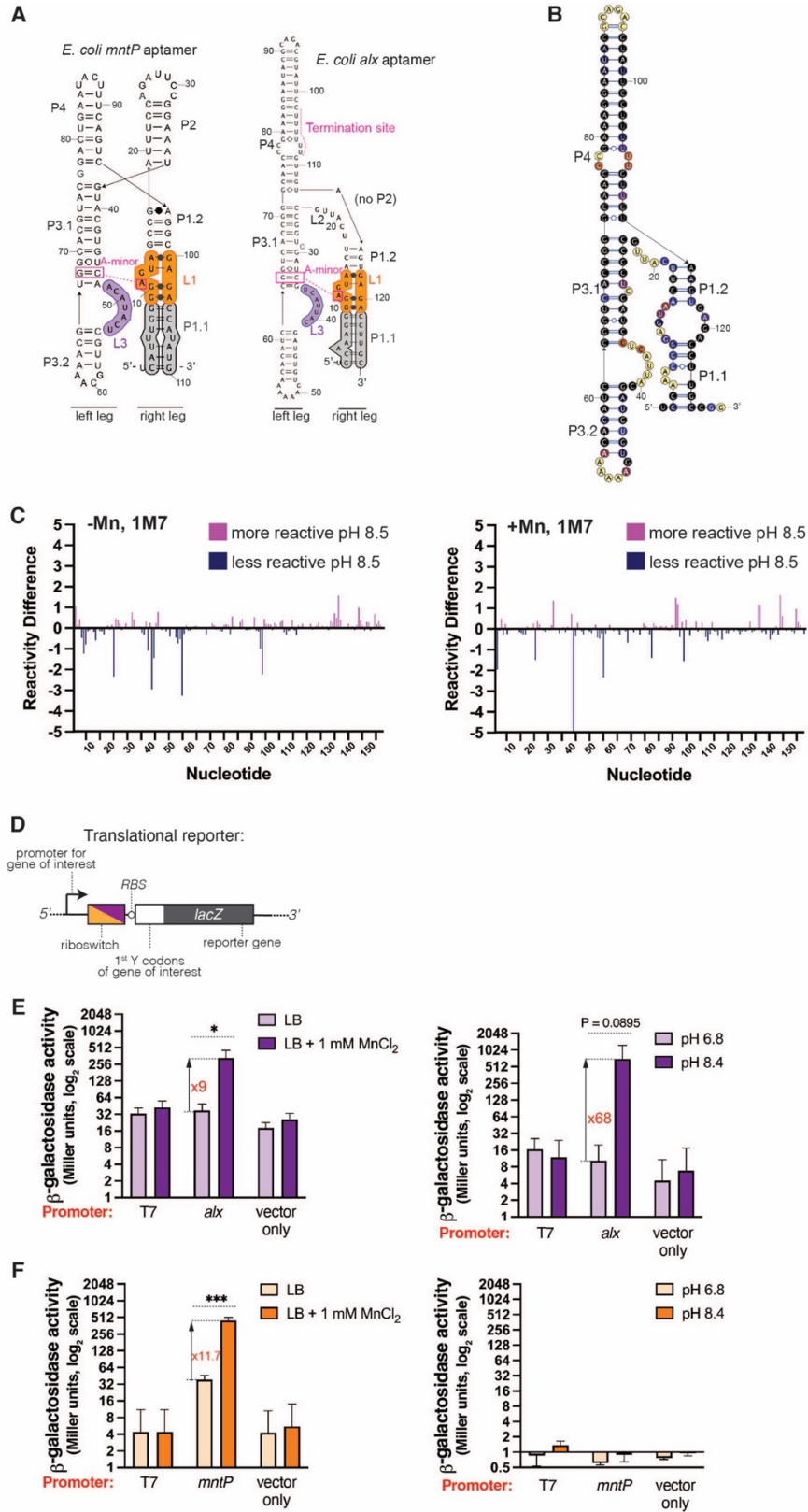
We next refolded the *alx* aptamer RNA at either neutral (7.2) or alkaline pH (8.5) in the presence of 1 mM MnCl₂ and probed with either DMS or 1M7. While the total physiological [Mn] (free and bound) in *E. coli* is estimated to be 15-21 μ M(35–37), we chose a saturating concentration of ligand to maximize the ligand-induced shift in the RNA conformational ensemble, consistent with prior chemical probing experiments on other riboswitches(27, 38). MgCl₂ was

present at 5 mM in the RNA refolding buffer, thus Mg remained in excess of Mn, as is the case *in vivo*(39). We observed L3 generally becomes less reactive towards 1M7 in the presence of Mn, consistent with reduced flexibility upon Mn binding (SI Fig. 4.1D). Interestingly, in the presence of Mn, pH-dependent reactivity changes for multiple L1/L3 nt are abolished (e.g., A37 and A11, Fig. 4.2C), supporting that these nt are buried in the metal-sensing core with Mn and consistent with published Mn-bound aptamer structures(6, 12, 21) (see Supplementary Note 1 and SI Fig. 4.1E for additional analysis).

Structural comparison of the alx and mntP aptamers. Despite belonging to the same family, the *alx* and *mntP* riboswitch aptamers harbor intriguing structural differences. First, the *alx* aptamer contains an intrinsic termination site in the P4 helix(34), which occurs before the metal-sensing pocket forms (Fig. 4.2A). The *alx* aptamer lacks the P2 helix, resulting in a three-way junction (3WJ) rather than a 4WJ in the *mntP* aptamer and all other aptamers studied by X-ray crystallography. Lastly, the L3 sequences in *alx* and *mntP* metal-sensing cores differ (Fig. 4.1C), raising the possibility that L3 tunes Mn affinity of the aptamer.

Figure 4.2 Aptamer architectures and riboswitch-controlled translational response *in vivo*.

A. *mntP* and *alx* aptamer structure models showing the *mntP* 4WJ and *alx* 3WJ elements. The key Mn-sensing loops (L1 is orange and L3 is purple) and the P1.1 switch helix (gray) are highlighted. **B.** Secondary structure model of equilibrium-refolded (pH 7.2 –Mn) *alx* aptamer modeled based on reactivity with 1M7. **C.** Individual nucleotide reactivity differences when modified with 1M7 at neutral (7.2) vs alkaline pH (8.5), measured in the presence or absence of Mn. **D.** Translational reporter construct design for *alx* and *mntP* translational reporters with the native or T7 promoter. **E.** Translational response of *alx* to supplemented Mn (left) or alkaline pH (right) as a function of promoter. Statistical analysis done using a parametric unpaired t-test to assess significance. **F.** Translational response of *mntP* to supplemented Mn (left) or alkaline pH (right) as a function of promoter. Statistical analysis done using a parametric unpaired t-test to assess significance.



Transcription by native RNA polymerase is required for *alx* and *mntP* riboswitch-mediated translational regulation *in vivo*

We next wondered whether *alx* and *mntP* riboregulation is kinetically controlled. To test this, we constructed *alx* and *mntP* translational *lacZ* reporters cloned with either the T7 promoter or native *alx/mntP* promoter into single copy plasmids (Fig. 4.2D). The effect of extracellular alkaline pH or elevated [Mn] on gene expression was measured by β -galactosidase assays in either an *E. coli* strain (MC4100) lacking *alx* (Δalx), to preclude negative autoregulation(19), or the Δalx strain with T7 RNAP expressed from the *lac* promoter (chromosomally encoded, $\Delta alxT7$). The plasmids containing the native promoter translational reporters were transformed into the Δalx strain and the plasmids containing the T7 promoter translational reporters were transformed into the $\Delta alxT7$ strain. The strains were cultivated in i) neutral pH (LBK pH 6.8) or alkaline pH (LBK pH 8.4) media, and ii) neutral pH (LBK pH 6.8) media with supplemented $MnCl_2$.

We observed the expected changes in *alx* and *mntP* translation in response to pH and supplemented $MnCl_2$ when reporter fusions were transcribed from their native promoters(19) (Fig. 4.2E-F). While only *alx* translation was induced by alkaline pH, translation of both genes increased with supplemented $MnCl_2$ (Fig. 4.2E-F). Strikingly, transcription by T7 RNAP abolished the response to alkaline pH and/or Mn (Fig. 4.2E-F), demonstrating that faster transcription of the *alx* and *mntP* riboswitches (~ 220 nt/s; T7 RNAP(40) vs. ~ 20 -90 nt/s; *E. coli* RNAP(41)) prevents riboswitching in response to alkalinization and/or elevated [Mn] in the cytosol. In sum, these data establish that transcription by *E. coli* RNAP is required for ligand-coupled structural rearrangements in the *alx* and *mntP* expression platforms.

Analysis of the *mntP* aptamer folding during transcription using co-transcriptional chemical probing

To investigate both pH and Mn-mediated structural changes in the *alx* and *mntP* riboswitch folding intermediates, we performed co-transcriptional chemical probing (variable length Transcription Elongation Complex RNA structure probing TECprobe-VL method, see SI Fig. 4.2 for construct designs)(27). Co-transcriptional chemical probing experiments were performed at both neutral (7.2) and alkaline pH (8.5) $\pm$ 1 mM MnCl_2 (referred to as “+Mn” or “-Mn” condition, respectively) with 10 mM MgCl_2 present to prevent compromised transcription fidelity from Mn-coordination by RNAP, as observed for DNA polymerases(42) and reverse transcriptases(43). Since published work established that $\text{pH} \geq 8$ renders all 4 nucleobases accessible to DMS modification(44), we first present data for *alx* and *mntP* DMS-probed folding intermediates at pH 8.5.

The two legs of the Mn aptamer fold differently. We initially mapped *mntP* aptamer folding transitions and determined the specific transcript lengths where helices in the aptamer become stably folded (Fig. 4.3A-B and SI Fig. 4.3). P3.1, P3.2, and P4 in the left leg of the aptamer fold first, as detected with coordinated drops in reactivity of nucleotides that are base-paired in these helices (Fig. 4.3C). In stark contrast, the right leg of the aptamer folds into transient, low-melting temperature (T_m) stem-loops that form and resolve until the synthesis of the aptamer is completed (~nt 124, assuming a 14-nt RNAP footprint, Fig. 4.3B). While the 5' RNA stretch for these aptamer elements (L1, P1.1, and P1.2) is transcribed early on, their 3' pairing partner does not emerge from RNAP until late in transcription. Thus, the weak stem-loop intermediates may prevent RNA from being trapped in stable but not biologically relevant folds, consistent with other riboswitches(16, 27). P2 in the right leg folds towards the end of aptamer transcription, even though the sequence encoding P2 emerges at transcript length ~50 nt, indicating that folding of the entire right leg occurs at the end of aptamer synthesis (Fig. 4.3C).

RNA folding intermediates sample for Mn. Surprisingly, we observed Mn-dependent differences in reactivity of several nucleotides before complete *mntP* aptamer has been synthesized (Fig. 4.3D). For example, A48, the first nucleotide in the L3 loop, demonstrates markedly high reactivity $\pm$ Mn, with its reactivity trending higher +Mn starting at transcript length $\sim$ 115 nt – before the aptamer synthesis finished $\sim$ 124 nt (Fig. 4.3D). We hypothesize that this is indicative of an RNA folding intermediate sampling divalent metal ions in solution in search for Mn, with its L3 and partially formed L1/P1.1. Such partial binding of Mn was observed in a crystal structure and molecular dynamics (MD) simulations of the full-length *X. oryzae* aptamer, referred to as “conformer 2”, which lacks the ordered stacking pattern in L3 and contains a partially dehydrated Mn in the M_B site(12) (Fig. 4.3D). The partially dehydrated Mn lacks half of the contacts seen in “conformer 1” (coordinated to fully dehydrated Mn); however, the inner sphere contact between the phosphate of the first nucleotide in L3 and the metal is observed (A48 in *E. coli mntP* aptamer vs. A46 in *X. oryzae* aptamer)(12). Therefore, A48 can contact a partially dehydrated Mn before the full aptamer is synthesized, explaining its Mn-dependent reactivity differences in *mntP* aptamer folding intermediates. U54 in L3 and A103 in L1 also show Mn-dependent reactivity differences in aptamer intermediates, with their reactivities trending higher – Mn (Fig. 4.3D). These reactivity differences can be explained by greater conformational flexibility in the metal-binding site (see Fig. 4.3D for U54 and A103 positioning in *X. oryzae* conformers 1 and 2). Additionally, we observed multiple RNAP pauses preceding *mntP* aptamer completion that are enhanced by the presence of Mn and may mediate partial Mn binding/ sampling of divalents in the environment by the RNA (see Supplementary Note 2 and SI Figs. 4.4-4.5 for further analysis).

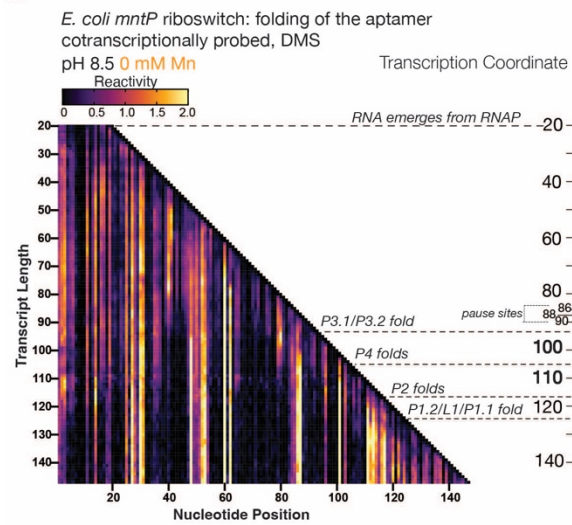
To identify the intermediates whose folding is truly controlled by transcription kinetics vs. thermodynamic stability of the folds, we applied a modified TECprobe-VL workflow wherein intermediate RNA transcripts are extracted from the stalled transcription elongation complexes

and equilibrium refolded before DMS probing. These experiments identified clear structural differences between co-transcriptionally folded and equilibrium refolded *mntP* intermediates and provided additional evidence of divalent sampling by aptamer folding intermediates (see Supplementary Note 3 and SI Fig. 4.6 for further analysis).

Figure 4.3 Co-transcriptional DMS probing of the *mntP* riboswitch aptamer at alkaline pH.

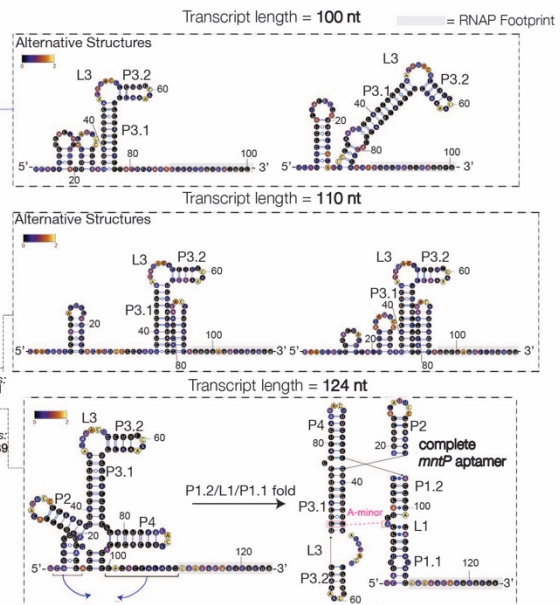
A. TECprobe-VL DMS reactivity matrix for the *mntP* aptamer at pH 8.5. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together. **B.** Secondary structures of *mntP* aptamer folding intermediates colored by DMS reactivity of the indicated transcript at pH 8.5. **C.** Transcript length-dependent reactivity changes showing folding transitions of the *mntP* aptamer at pH 8.5 $\pm$ 1 mM Mn. DMS data for 0 mM Mn (orange) and 1 mM Mn (purple) conditions from Fig. 4.3A and SI Fig. 4.3, respectively. *E. coli mntP* apo aptamer (PDB ID: 4Y1M) shown with labeled loops and helical domains. Nt highlighted in orange correspond to the nt positions with reactivities plotted in the adjacent graph. **D.** Plots show transcript length- and Mn-dependent reactivity changes in *mntP* aptamer folding intermediates at pH 8.5 $\pm$ 1 mM Mn. DMS data for 0 mM Mn (orange) and 1 mM Mn (purple) conditions from Fig. 4.3A and SI Fig. 4.3, respectively. Vertical dotted line marks completion of aptamer folding. *X. oryzae* aptamer metal-sensing core (PDB ID: 6N2V) with Mg²⁺ (green sphere) and Mn²⁺ (purple sphere) highlighted. Conformer 1 shows Mn²⁺ fully bound and Conformer 2 shows Mn²⁺ partially bound. Nt highlighted in purple correspond to nt reactivity plots above.

A

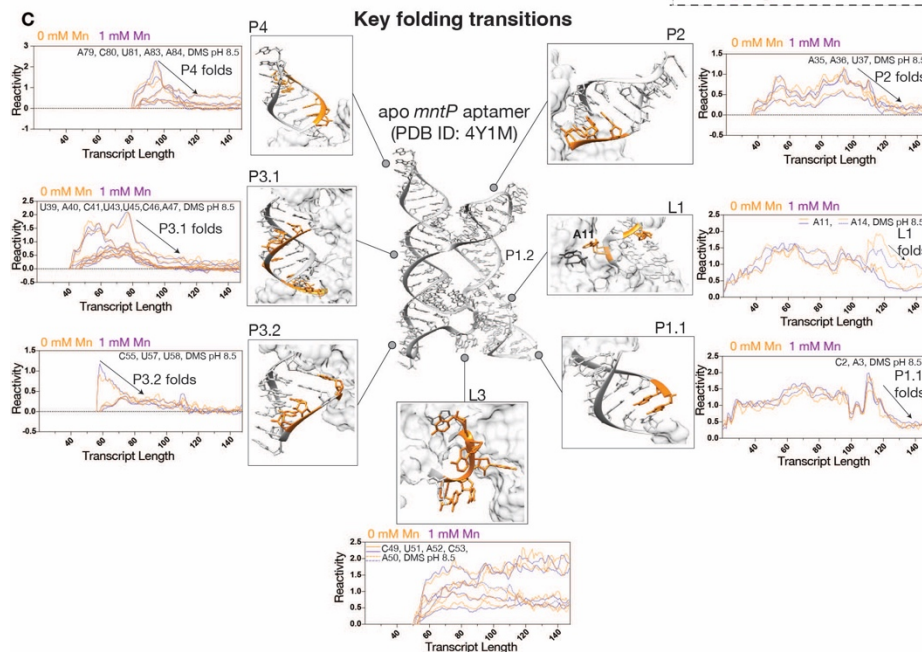


B

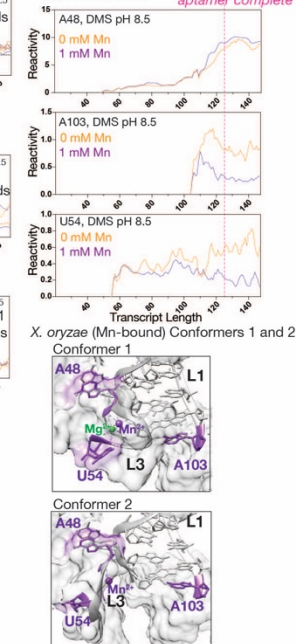
Aptamer folding intermediates



C



D Ligand-dependent differences



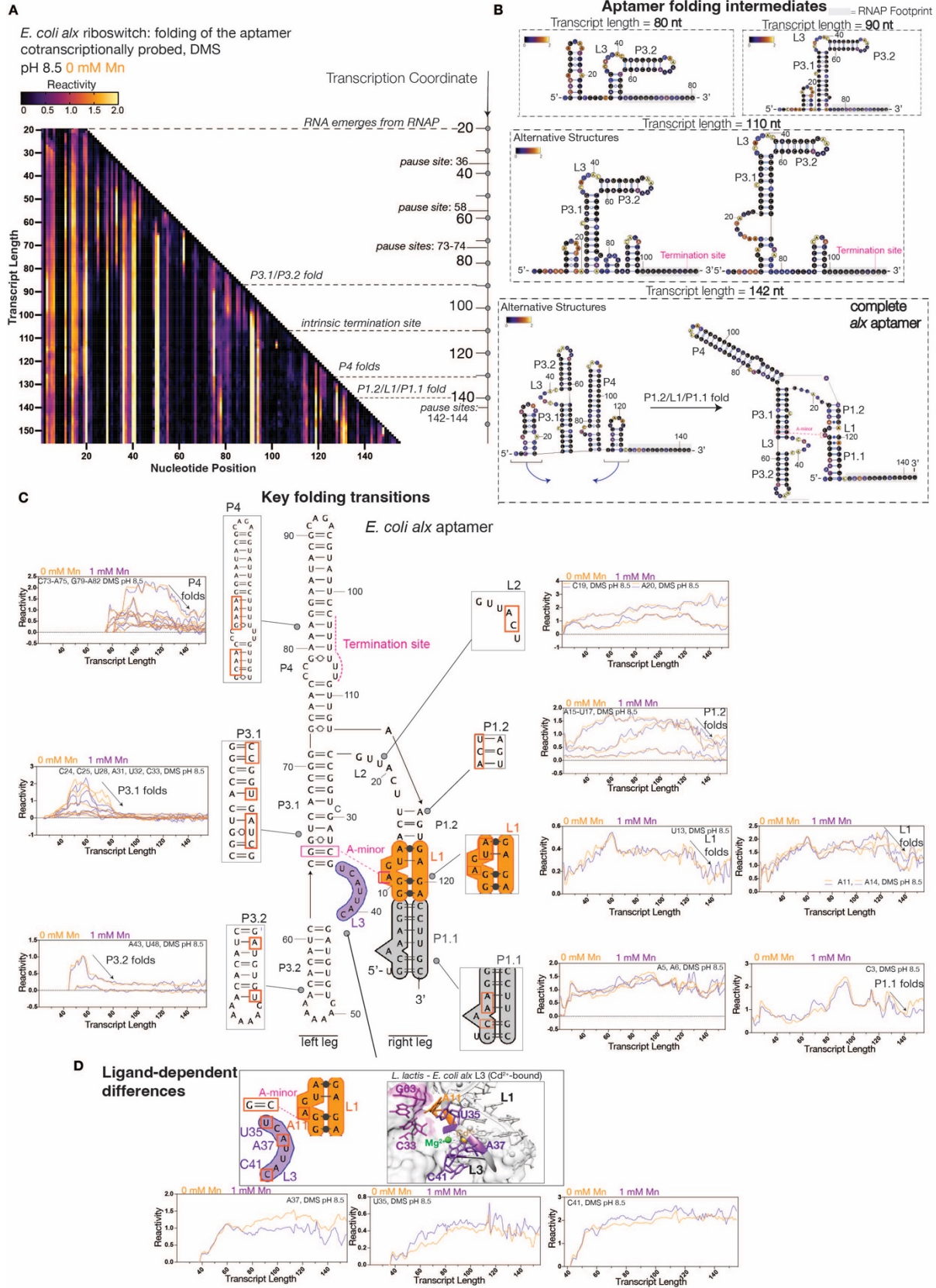
Analysis of the *alx* aptamer folding during transcription using co-transcriptional chemical probing

We next investigated the *alx* aptamer folding transitions at pH 8.5 (Fig. 4.4A-B and SI Fig. 4.7). Here, P3.1, P3.2, and P4 in the left leg fold first as the aptamer is extended from ~80 nt to ~125 nt (Fig. 4.4B-4.4C). Like in the *mntP* aptamer, the 5' end of the right leg folds into transient, low- T_m stem-loops that resolve once the full-length aptamer RNA has emerged from RNAP (Fig. 4.4B), highlighting a potentially conserved mechanism of the *yybP-ykoY* aptamer folding.

*L3 flexibility differs between *alx* and *mntP* aptamers.* Like with *mntP*, we captured Mn-dependent reactivity differences for L3 in *alx* aptamer folding intermediates (Fig. 4.4D). For example, the reactivity of U35, the first nt in the *alx* L3, trends higher +Mn starting at transcript lengths as short as ~50 nt (Fig. 4.4D). This trend mimics that seen with the first nt in *mntP* L3, A48, with a few key differences. First, the absolute reactivity values $\pm$ Mn are greater for A48 in *mntP* L3 (Fig. 4.3D). Second, the increased reactivity +Mn occurs much earlier for *alx* aptamer intermediate transcript lengths (~50 nt for *alx* vs. ~115 nt for *mntP*). The absolutely conserved adenosine in *alx* L3 that directly coordinates Mn (A37) is more reactive –Mn for aptamer folding intermediates starting at transcript length ~70 nt (Fig. 4.4D). The higher A37 reactivity –Mn is consistent with greater L3 flexibility in this condition and further supports that the L3 loop can begin sampling for divalents before the aptamer domain is fully folded. Curiously, the absolutely conserved adenosine in *mntP* L3 (A50) does not demonstrate this Mn-dependent reactivity difference (Fig. 4.3C), pointing to a difference in L3 dynamics between *alx* and *mntP* aptamer folding intermediates. Similar to *mntP*, we observed clear structural differences between co-transcriptionally folded and equilibrium refolded *alx* intermediates (see Supplementary Note 4 and SI Fig. 4.8 for further analysis).

Figure 4.4 Co-transcriptional DMS probing of the *alx* riboswitch aptamer at alkaline pH.

A. TECprobe-VL DMS reactivity matrix for the *alx* aptamer at pH 8.5. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from seven independent replicates that were concatenated and analyzed together. **B.** Secondary structures of *alx* aptamer folding intermediates, which are colored by the DMS reactivity of the indicated transcript at pH 8.5. **C.** Transcript length-dependent reactivity changes showing folding transitions of the *alx* aptamer at pH 8.5 ± 1 mM Mn. DMS data for 0 mM Mn (orange) and 1 mM Mn (purple) conditions from Fig. 4.4A and SI Fig. 4.7, respectively. Secondary structure model of *alx* aptamer with labeled structural domains. L1 (orange), L3 (purple), P1.1 (gray), and intrinsic termination site (pink) are highlighted. Nt boxed in red correspond to adjacent nt reactivity plots. **D.** *L. lactis* Mn-sensing aptamer with *E. coli alx* L3 sequence (PDB ID: 6CC1). L3 nt (purple), cross-helix A-minor nt: A11 (orange) and G63-C33 (magenta), Mg^{2+} (green sphere), and Cd^{2+} (yellow sphere) are highlighted. Plots show transcript length- and Mn-dependent reactivity changes in *alx* aptamer folding intermediates at pH 8.5 ± 1 mM Mn. DMS data for 0 mM (orange) and 1 mM (purple) conditions from Fig. 4.4A and SI Fig. 4.7, respectively.



Mn binding is a prerequisite for *alx* riboswitch pH response

Early work characterized the *alx* riboswitch as a pH sensor(18); however, recent studies have called this into question(19). Since the intracellular [Mn] increases in *E. coli* at alkaline pH(19), we wondered whether the *alx* riboswitch is actually sensing the increase in [Mn] as a consequence of cytosol alkalinization. To test this, we performed translational *lacZ* reporter assays on the A37U *alx* mutant, where the absolutely conserved adenosine that directly coordinates Mn in L3 is replaced with a uridine. This point mutation was previously shown to ablate Mn-sensing in the *X. oryzae* aptamer(12) making it the best candidate mutation to decouple the effect of Mn and alkaline pH on *alx* riboregulation *in vivo*. Strikingly, the 68-fold increase in *alx* reporter translation observed upon a shift from pH 6.8 to 8.4 for the wild-type riboswitch was abolished by the A37U mutation (Fig. 4.5A). These data provide the first evidence that the *alx* riboswitch is not a true pH sensor; instead, the riboswitch senses increased cytosolic [Mn], which is an indirect effect of cytosol alkalinization. Published *alx* translational reporter data showed an additive effect of alkaline pH and Mn(19). We thus hypothesize that alkaline pH enhances the affinity of the *alx* aptamer for Mn, rather than alters the *alx* riboswitch folding, unifying observations from early studies on this system(18, 45) with a recently published genetic study(19). A combination of alkaline pH and high [Mn] produced higher *mntP* reporter translation compared to the neutral pH condition with added Mn, suggesting that alkaline pH may also enhance Mn affinity for the *mntP* aptamer, but to a lesser degree than for the *alx* aptamer(19). To explore this intriguing point further, we DMS-probed the *alx* and *mntP* folding intermediates at pH 7.2 $\pm$ Mn with the TECprobe-VL method, as described in the following section.

pH effects on the *alx* and *mntP* aptamer folding: L3

We largely observed the same co-transcriptional folding transitions for the two legs in the *alx* and *mntP* aptamers at pH 7.2 as described for pH 8.5 (SI Fig. 4.9A-B). Although the *alx* and *mntP* aptamers exhibit the same global folding transitions at pH 7.2, key pH-dependent reactivity

differences in their metal-sensing cores emerged. Here, we focused on A and C nucleotide reactivities since G and U nucleotides are protonated at the respective N1 and N3 positions at neutral pH and are thus unreactive towards DMS ($pK_a \sim 9.2(46)$). First, the absolutely conserved adenosine in the *alx* L3 (A37) exhibits a general Mn-dependent reactivity decrease during aptamer folding at both pH (Fig. 4.5B). However, the magnitude of this reactivity decrease is clearly pH-dependent, with A37 consistently less reactive at pH 8.5 +Mn compared to other conditions tested (Fig. 4.5B), suggesting a potentially more stable A37-Mn coordination at alkaline pH. In stark contrast, the absolutely conserved adenosine in *mntP* L3 (A50) shows no drop in reactivity at pH 8.5 +Mn; instead, A50 lacks clear pH-dependent reactivity trends during aptamer folding (Fig. 4.5C). The first L3 nucleotide in the *mntP* aptamer (A48) is more reactive +Mn vs. – Mn regardless of pH (Fig. 4.5C), supporting that A48 stacked with A11 in an Mn-bound, ordered metal-sensing core increases its DMS reactivity. Interestingly, the absolute reactivity values for A48 are markedly lower at pH 7.2 $\pm$ Mn (Fig. 4.5C), suggesting that A48 has an upwardly shifted pK_a close to physiological pH ($\sim$ pH 7.4-7.8 for *E. coli*(23)). We thus hypothesize that deprotonation of A48 at alkaline pH promotes its stacking with A11 in the cross-helix A-minor interaction, favoring docking of the helical legs in the aptamer (see Supplementary Note 5 and SI Figs. 4.9C-D, 4.10, and 4.11 for further analysis of pH-dependent structural changes in *alx/mntP* L1 and L3 in co-transcriptionally folded vs. equilibrium refolded aptamer intermediates).

We wondered whether L3 sequences in *Enterobacteriales* *yybP-ykoY* family members show a unique trend that could explain the pH-dependent reactivity differences in the *alx* and *mntP* aptamers. We aligned sequences of *yybP-ykoY* riboswitch family members within the *Enterobacteriales* order, revealing two distinctive groups: an *alx*-like group (consensus: UCAUUAC, Fig. 4.5D) and an *mntP*-like group (consensus: ACAUACU, Fig. 4.5D). We then compared the above sequences to the *yybP-ykoY* riboswitch family L3 consensus sequence generated from 1,063 aligned sequences (consensus: ACAUACU, Fig. 4.5D). Generally, the first

nucleotide in L3 is an A – an important position since this nucleotide stacks with the conserved A11 in L1 that forms the cross-helix A-minor interaction. Therefore, this position being strictly a U in *alx*-like members suggests a functional role. Since the N3 of U has a close-to-physiological pKa of ~9.2(46), it is possible that a shift from neutral to alkaline pH alters its stacking with the conserved A11 and consequently Mn-coupled structural rearrangements in the expression platform.

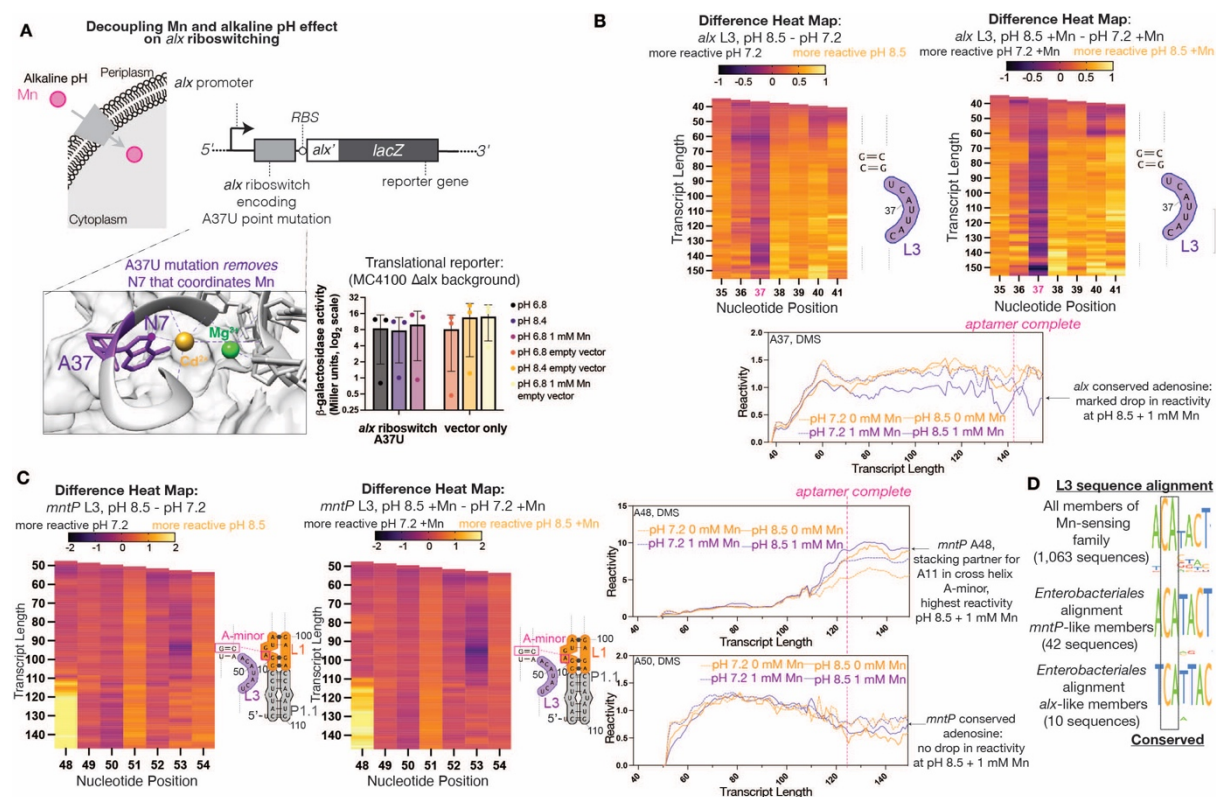


Figure 4.5 Decoupling the effects of alkaline pH and Mn on *alx* riboswitching *in vivo* and pH-dependent reactivity changes in the divergent *alx* and *mntP* L3 sequences.

A. (Top panel) *E. coli* cell with increased cytosolic [Mn] resulting from an alkaline pH challenge (left) and construct design for *alx* translational reporter (right). (Bottom panel) *L. lactis* Mn-sensing aptamer with *E. coli* *alx* L3 sequence (PDB ID: 6CC1, left). Conserved A37 (purple), Mg^{2+} (green sphere), and Cd^{2+} (yellow sphere) are highlighted. (Bar graph) Loss of *alx* translational response to either Mn or pH upon mutation of the Mn-coordinating adenosine. **B-C.** Reactivity difference heatmaps and reactivity changes in aptamer intermediates for *alx* L3 nt (**B**) and *mntP* L3 nt (**C**) during aptamer folding at pH 7.2 vs. 8.5 $\pm$ 1 mM Mn. *alx* L3 and *mntP* L3 (purple), L1 (orange), L3 (purple), P1.1 (gray), and cross-helix A-minor (pink) are highlighted. DMS data for 0 mM Mn (orange) and 1 mM Mn (purple) conditions from Fig. 4.4A, SI Fig. 4.7, and SI Fig. 4.9B (**B**, *alx*) and Fig. 4.3A, SI Fig. 4.3, and SI Fig. 4.9A (**C**, *mntP*). Vertical dotted line marks completion of aptamer folding. **D.** L3 consensus sequences from *yybP-ykoY* family members in the Rfam database (top) and *mntP*-like (middle)/ *alx*-like (bottom) members in the *Enterobacteriales* order.

Analysis of *mntP* expression platform folding during transcription

Cross-helix A-minor disruption at the riboswitching point. The expression platform for the *mntP* riboswitch contains a single helix that begins folding at transcript length ~175 nt and was detected by coordinated drops in reactivity for its constituent nucleotides (Fig. 4.6A and Fig. S4.12A). In the absence of Mn, the aptamer remains stably folded for most of expression platform synthesis until the transcript is ~205 nt long (Fig. 4.6B). At this critical decision point – the riboswitching point – the L1 and P1.1 elements of the aptamer are disrupted if no Mn is available, allowing for alternative pairing of P1.1 switch helix in the expression platform (Fig. 4.6C). The disruption of P1.1 in the absence of Mn is consistent with previously published single-molecule fluorescence experiments with an isolated aptamer that show that this helix is less stable without Mn(12). Analysis of single-nucleotide trajectories revealed that A11 reactivity transiently spikes at ~205 nt, whether or not Mn is included in the transcription reaction, but the amplitude of the spike is Mn-dependent (Fig. 4.6D). The spike in A11 reactivity coincides with P1.1 strand invasion and its reduced amplitude +Mn indicates A11 remains trapped in the long-range A-minor interaction in Mn-bound aptamer. The L1/P1.1 reactivity difference heatmap $\pm$ Mn highlights A11 as the nucleotide most sensitive to the presence of the ligand, serving as a marker for the riboswitching point (Fig. 4.6D). We also assessed reactivity changes for A48 in L3 during expression platform folding, since A48 stacks with A11 and helps anchor the cross-helix A-minor interaction. A48 reactivity drops $\pm$ Mn starting at transcript length ~185 nt, with a larger decrease –Mn (Fig. 4.6E). A48 reactivity is consistently lower –Mn for the remainder of expression platform folding (Fig. 4.6E), likely due to broken A48-A11 stacking.

We observed similar reactivity trends in *mntP* expression platform folding at pH 7.2 as pH 8.5 (e.g., A11 reactivity spike, SI Fig. 4.12B-C) and uncovered additional evidence of P1.1/L1 disruption in the vicinity of the riboswitching point (see Supplementary Note 6 and SI Fig. 4.12D). Taken together, our observations highlight that a global structural change occurs in the *mntP*

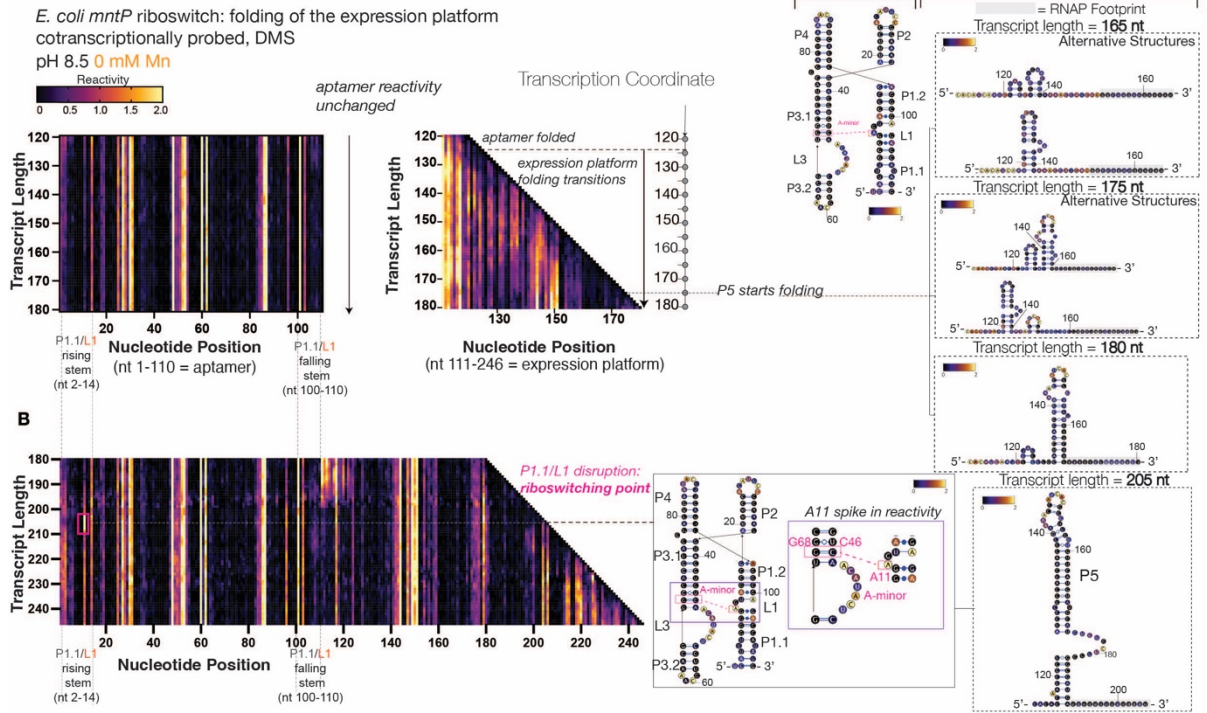
expression platform at transcript lengths ~200-205 nt, which is likely prompted by the start of strand invasion at P1.1 in the aptamer (see Supplementary Note 6 and SI Fig. 4.12D for additional analysis of pH- and Mn-dependent reactivity changes in *mntP* L1, L3, and P1.1 during expression platform folding).

Co-transcriptional binding of Mn is required for riboswitching. We next evaluated the differences between co-transcriptionally folded and equilibrium refolded *mntP* intermediate transcripts that contain both the aptamer domain and expression platform, to assess whether co-transcriptional folding is required to observe Mn-coupled structural rearrangements in the expression platform. While we observed similar P5 folding transitions (SI Fig. 4.13A-B), the Mn-dependent spike in A11 reactivity is abolished when the RNA is equilibrium refolded $\pm$ Mn at either pH 7.2 or 8.5 (SI Fig. 4.13C). These data support that co-transcriptional Mn-binding is critical for stabilization of the A-minor motif that docks the two legs of the aptamer together and for prevention of the P1.1/L1 invader strand from disrupting the aptamer (see Supplementary Note 7 and SI Fig. 4.13D for reactivity analysis of the A11 stacking partner, A48, in equilibrium refolded intermediates). Further supporting the above observations, we found that T7-synthesized, equilibrium refolded *mntP* riboswitch adopts a translationally inactive conformation $\pm$ Mn (DMS/1M7-probed, described further in Supplementary Note 8 and SI Fig. 4.14).

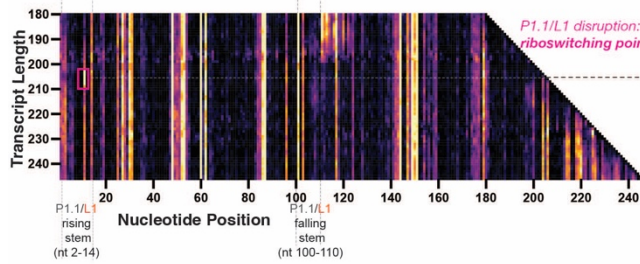
Figure 4.6 Co-transcriptional DMS probing of the *mntP* riboswitch during expression platform folding.

A-B. (Left panel) TECprobe-VL DMS reactivity matrices of *mntP* riboswitch at pH 8.5. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together. (Right panel) Secondary structures of *mntP* expression platform folding intermediates colored by DMS reactivity of the indicated transcript at pH 8.5. **C.** (Top panel) Truncated secondary structure models of *mntP* riboswitch with a stabilized L1/P1.1 (Mn-bound) or disrupted L1/P1.1 (Mn-unbound). L1 (orange), L3 (purple), P1.1 (gray), and cross-helix A-minor interaction (pink) are highlighted. (Bottom panel) *X. oryzae* aptamer metal-sensing core (PDB ID: 6N2V) with Mg²⁺ (green sphere) and Mn²⁺ (purple sphere) highlighted. Conformer 1 shows Mn²⁺ fully bound and Conformer 2 shows Mn²⁺ partially bound. A11 (orange) and A48 (purple) are highlighted. **D-E.** (Top panels) Transcript length- and Mn-dependent reactivity changes in A11 (**D**) and A48 (**E**) showing cross-helix A-minor disruption. (Bottom panels) Reactivity difference heatmaps for P1.1/L1 (**D**) and L3 (**E**) at pH 8.5 ± 1 mM Mn. DMS data for 0 mM and 1 mM Mn conditions from Fig. 4.6A-B and SI Fig. 4.12A, respectively. Vertical or horizontal dotted lines mark indicated RNA folding transitions.

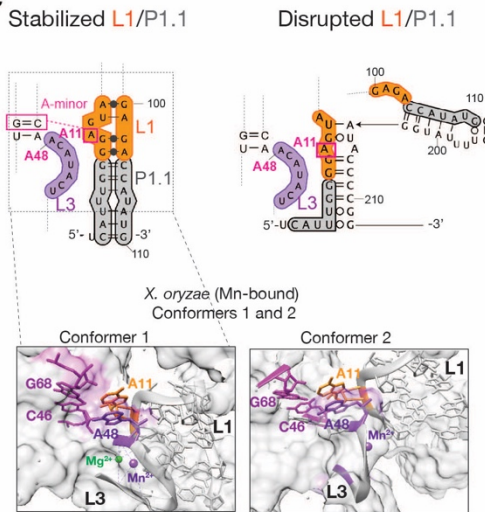
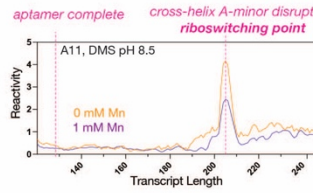
A

Key *mntP* expression platform folding transitions

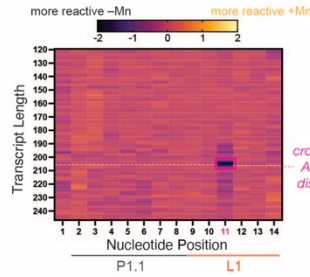
B



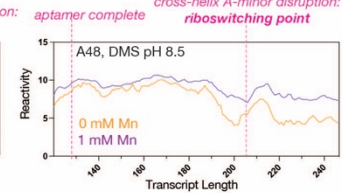
C

D *mntP* riboswitching point: cross-helix A-minor disruption

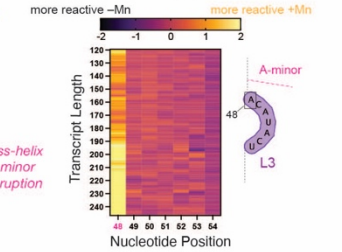
Difference heat map: pH 8.5 +Mn - pH 8.5



E Altered L3 stacking after cross-helix A-minor disruption



Difference heat map: pH 8.5 +Mn - pH 8.5



Analysis of *alx* expression platform folding during transcription

*Distinct features of the *alx* expression platform folding.* The expression platform for the *alx* riboswitch contains multiple helices that begin folding at transcript length ~155 nt, as detected by coordinated drops in reactivity for the constituent nucleotides in each helix (Fig. 4.7A and SI Fig. 4.15A). The *alx* expression platform contains two helices (termed here P5 and P6) in its translationally inactive conformation that are predicted to fold $-Mn$, resulting in the ribosome binding sequence (RBS) base-pairing with a disrupted P1.1/L1 (Fig. 4.7B-D). On the other hand, when P1.1/L1 remain base-paired in the Mn-bound *alx* aptamer, an alternative helix in the expression platform (termed here P5/P6) is predicted to fold, and the downstream RBS-containing strand is single-stranded (Fig. 4.7E-G).

Unlike the *mntP* riboswitch, the *alx* P1.1/L1 rising stem is highly dynamic during the entirety of *alx* expression platform folding both $\pm Mn$, which was detected as consistent reactivity towards DMS for multiple nt (Fig. 4.7A). The strand predicted to invade P1.1/L1 in the *alx* aptamer emerges at transcript length 155 nt, resulting in formation of the P5 helix (Fig. 4.7A). Once the expression platform is extended to transcript length 190 nt, the P6 intrinsic terminator helix folds (Fig. 4.7A). Interestingly, the DMS reactivities for nucleotides in the *alx* expression platform were largely similar $\pm Mn$, resulting in similar secondary structure predictions for expression platform folding intermediates (SI Fig. 4.15B-D).

Despite the global structural similarities in the *alx* expression platform secondary structural models $\pm Mn$, we detected multiple Mn-dependent reactivity differences for specific nucleotides that indicate a modest ligand-dependent shift in the conformational ensemble. P5 stem-loop nucleotides are more reactive $-Mn$ (Fig. 4.7C), consistent with a population bias towards the inactive conformer. A137 (bulged in active conformer vs. base-paired in inactive conformer) demonstrates markedly high reactivity in general and exhibits higher reactivity $+Mn$ (Fig. 4.7F), indicating a shift in the conformational ensemble towards the *alx* active conformer.

pH effects during folding of the alx expression platform. We observed similar DMS reactivity patterns in the *alx* expression platform at neutral vs. alkaline pH, such as folding of P5 and P6 (SI Fig. 4.16A-B), indicating that a shift to alkaline pH impacts metal-sensing rather than directly affecting folding of the expression platform and thus riboswitching. However, we observed interesting pH-dependent reactivity trends for A137, which was more reactive +Mn at both pH and most reactive overall at pH 8.5 +Mn (SI Fig. 4.16B), supporting that the active conformer is sampled more frequently under this condition.

We next assessed pH- and Mn-dependent reactivity trends in the *alx* aptamer during expression platform folding (Fig. 4.8A-C). Contrasting *mntP*, A11 remains accessible for DMS methylation across all conditions tested with no significant pH and Mn-dependent differences (Fig 4.8B and SI Fig. 4.16C). Since A11 is conserved in *alx* and *mntP* L1, it is curious that A11 exhibits distinctive reactivity trends between the two riboswitches. Unlike in *mntP*, the absolutely conserved A37 in *alx* aptamer remains least reactive at pH 8.5 +Mn for the duration of expression platform folding (supportive of more stable Mn binding at alkaline pH), with multiple transient spikes in reactivity indicating a change in L3 dynamics at pH 8.5 +Mn during expression platform folding (Fig. 4.8C and SI Fig. 4.16C). Additionally, we observed interesting pH- and Mn-dependent reactivity trends for *alx* expression platform nt in equilibrium refolded intermediates (see Supplementary Note 9 and SI Fig. 4.17).

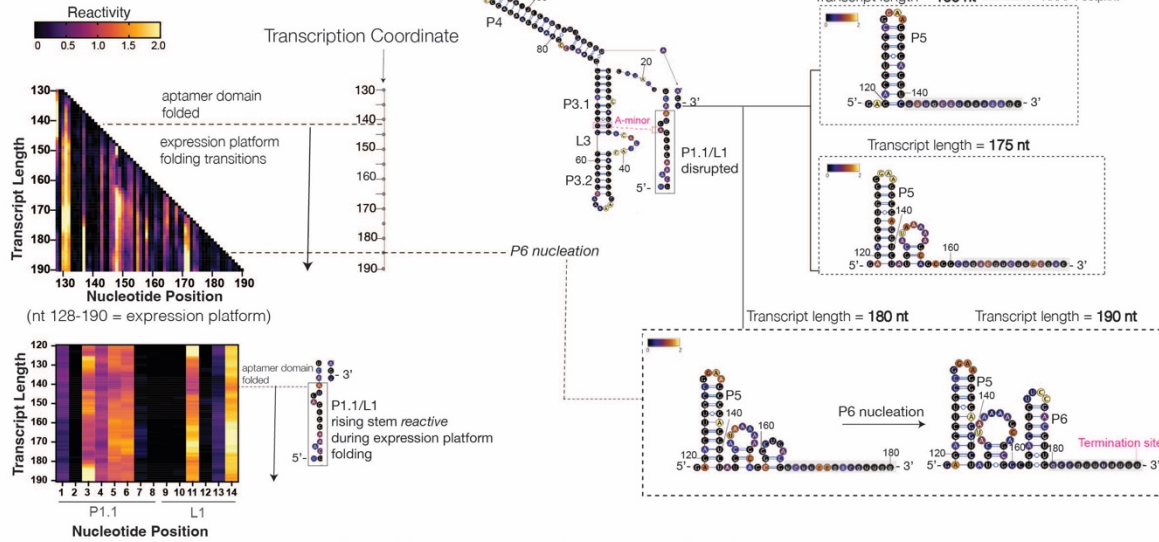
Figure 4.7 Co-transcriptional DMS probing of the *a/x* riboswitch during expression platform folding.

A. (Left panel) TECprobe-VL DMS reactivity matrix of the *a/x* expression platform (top) and L1/P1.1 rising stem (bottom) at pH 8.5. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from seven independent replicates that were concatenated and analyzed together. (Right panel) Secondary structures of *a/x* folding intermediates colored by DMS reactivity of the indicated transcript at pH 8.5. **B.** Truncated secondary structure model of the *a/x* aptamer with a disrupted L1/P1.1 (Mn-unbound). L1 (orange), L3 (purple), and P1.1 (gray) are highlighted. **C.** Secondary structure model of *a/x* expression platform at transcript length 190 nt –Mn. L1 falling stem (orange) and P1.1 falling stem (gray) paired in P5 are highlighted. Adjacent to nt boxed in pink are plots with their respective transcript-length and Mn-dependent reactivity changes. DMS data for 0 mM (orange) and 1 mM Mn (purple) conditions are from Fig. 4.7A and SI Fig. 4.15A, respectively. **D.** Secondary structure model of full-length, translationally inactive *a/x* expression platform with same aptamer color scheme as in panels A-B. RBS/start codon (red) and P6 intrinsic terminator stem (pink) are highlighted. **E.** Truncated secondary structure model of the *a/x* aptamer with a stabilized L1/P1.1 (Mn-bound) with same aptamer color scheme as in panel A. Cross-helix A-minor motif (pink) is highlighted. **F.** Secondary structure model of *a/x* expression platform at transcript length 190 nt +Mn. Adjacent to nt boxed in pink are plots with their respective transcript-length and Mn-dependent reactivity changes. DMS data for 0 mM (orange) and 1 mM Mn (purple) conditions are from Fig. 4.7A and SI Fig. 4.15A, respectively. **G.** Secondary structure model of full-length, translationally active *a/x* expression platform with same aptamer color scheme as in panel D. RBS/start codon (red) and intrinsic termination site in inactive conformer (pink) are highlighted.

A

Key *alx* expression platform folding transitions

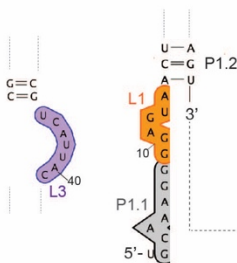
E. coli alx riboswitch: folding of the expression platform cotranscriptionally probed, DMS
pH 8.5 0 mM Mn



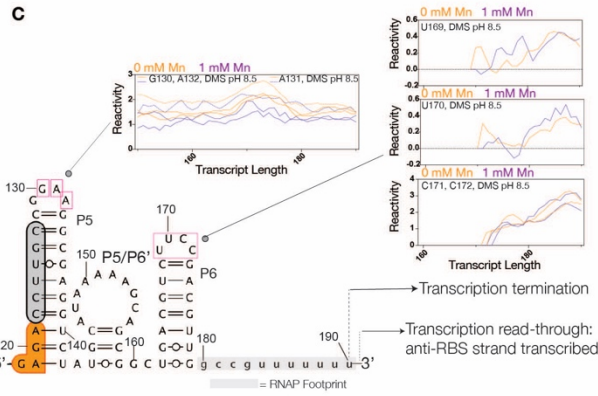
Mn-dependent differences in *alx* expression platform structure

B

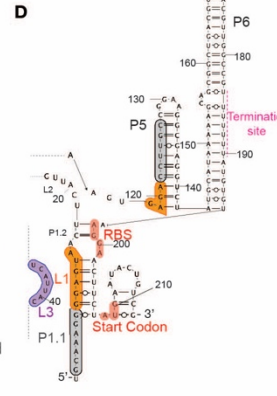
Disrupted L1/P1.1 0 mM Mn
(aptamer core)



C

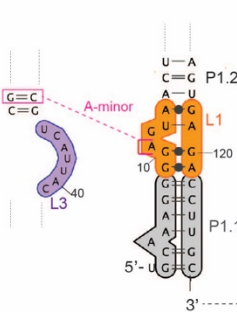


D

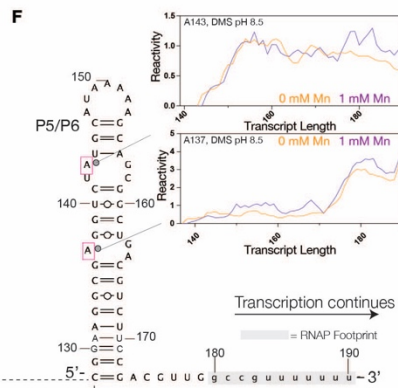


E

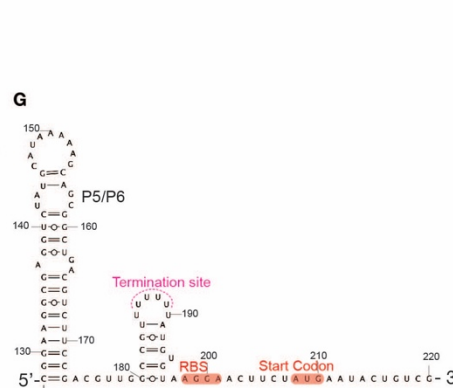
Stabilized L1/P1.1 1 mM Mn
(aptamer core)



F



G



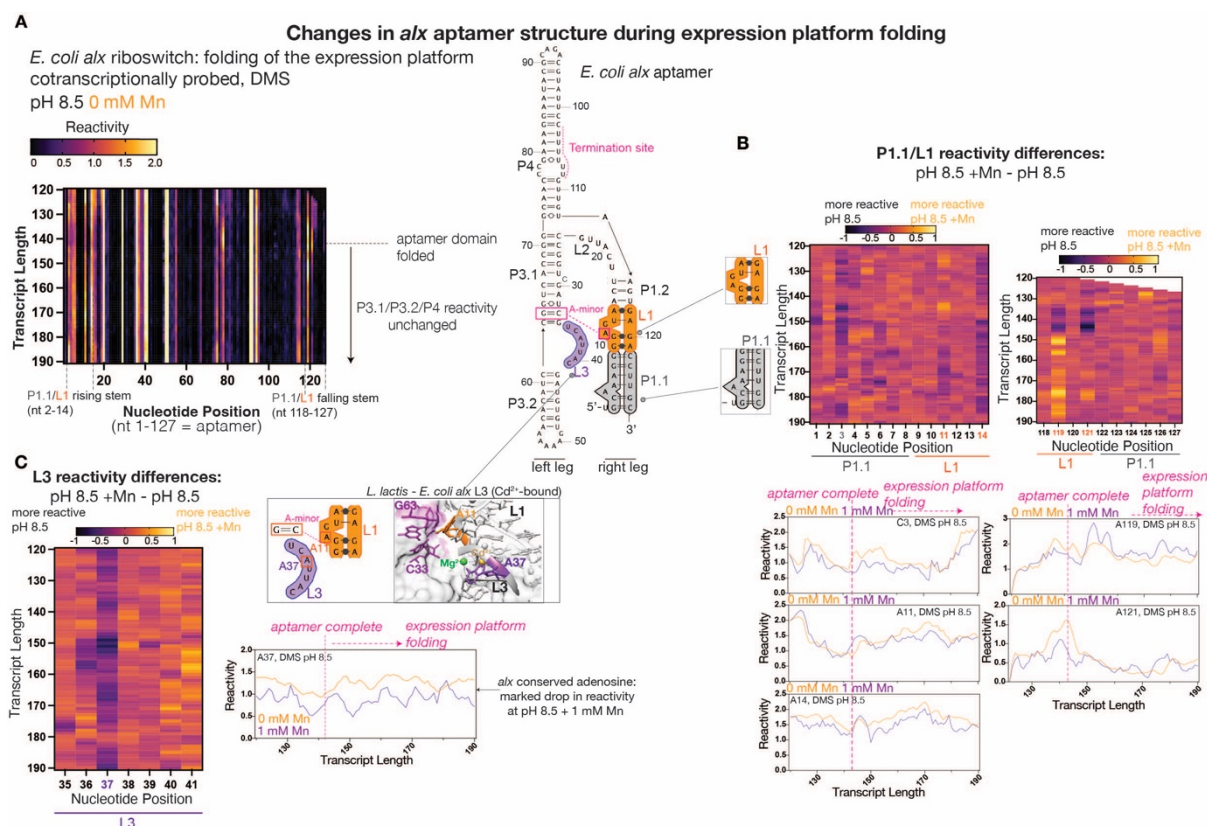


Figure 4.8 Changes in the *alx* aptamer during expression platform folding.

A. (Left panel) TECprobe-VL DMS reactivity matrix of the *alx* riboswitch aptamer during expression platform folding at pH 8.5. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from seven independent replicates that were concatenated and analyzed together. (Right panel) Secondary structure model of *alx* aptamer with labeled structural domains. L1 (orange), L3 (purple), P1.1 (gray), and intrinsic termination site (pink) are highlighted. **B-C.** Reactivity difference heatmaps for L1/P1.1 (**B**) and L3 (**C**) at pH 8.5 $\pm$ 1 mM Mn. Adjacent to heatmaps are transcript length- and Mn-dependent reactivity changes in L1/P1.1 nt (**B**) and L3 nt (**C**). DMS data for 0 mM and 1 mM Mn conditions from Fig. 4.8A and SI Fig. 4.15A, respectively. Vertical or horizontal dotted lines/arrows mark indicated RNA folding transitions. *L. lactis* Mn-sensing aptamer with *E. coli alx* L3 sequence (PDB ID: 6CC1) shown in panel B with A37 (purple), cross-helix A-minor nt: A11 (orange) and G63-C33 (magenta), Mg²⁺ (green sphere), and Cd²⁺ (yellow sphere) highlighted.

The *mntP* and *alx* RBS regions are dynamic in co-transcriptionally displayed RNA

The RNA surrounding the *mntP* RBS did not demonstrate a clear Mn-dependent reactivity trend, with DMS-constrained structure models predicting low- T_m hairpins interspersed with single-stranded regions (Fig. 4.18A). We thus conclude that the RNA around the RBS is dynamic, and any Mn-dependent structural differences are likely averaged out. Prior smFRET study observed

such RBS dynamics for the preQ₁ translational riboswitch with ligand-dependent accessibility bursts in the RBS region(47). Therefore, characterization of the *mntP* RBS dynamics requires complementary, e.g., single-molecule methods. Likewise, heatmap visualization of the *mntP* RBS nucleotides at pH 7.2 or 8.5 $\pm$ Mn or in equilibrium refolded intermediates at both pH showed no clear reactivity difference that would indicate increased RBS accessibility +Mn (SI Fig. 4.18B-C).

For *alx*, the presence of the P6 intrinsic terminator stem resulted in a stark decrease in representation for longer transcripts containing the RBS (SI Fig. 4.19A), precluding accurate modeling of the RNA structure in this region. To boost representation of full-length *alx*, we performed a modified co-transcriptional chemical probing experiment with a terminally biotinylated DNA template and the fast-acting SHAPE reagent benzoyl cyanide (BzCN, reaction time $\sim$ 1s) to better capture the short-lived co-transcriptional, Mn-dependent structural changes in the RBS. Consistent with *mntP*, the single-length BzCN probing experiment revealed no clear Mn-dependent differences in the reactivities of RBS nucleotides (Fig. 4.19B), supporting that the *alx* RBS is dynamic. These dynamics may serve a functional role; for instance, the RBS may be rapidly docking/undocking with another tertiary structural element in *alx*, with the docking rate fine-tuning RBS accessibility. Underscoring the importance of co-transcriptional chemical probing, we found that T7-synthesized, equilibrium refolded *alx* riboswitch adopts a translationally inactive conformation $\pm$ Mn (DMS/1M7-probed, described further in Supplementary Note 10 and SI Fig. 4.20). Taken together, while *alx* and *mntP* differ structurally, they may share common dynamic features that gate RBS accessibility.

4. Discussion

In this work, we performed a suite of chemical probing experiments on the Mn-sensing translational riboswitches in *E. coli* that precede the *mntP* and *alx* genes encoding Mn exporters. While static structural snapshots are available for Mn-sensing aptamers, these structures do not answer the question of how binding of Mn to the aptamer is “communicated” to the expression

platform – the question at the heart of riboswitch function. Because the *mntP* and *alx* aptamers display unique architectures, they are an interesting test case of how structurally distinct aptamers in the same riboswitch family “talk” to the expression platform. We demonstrated that co-transcriptional folding with the native *E. coli* RNAP is essential for Mn-coupled structural rearrangements in the *mntP* and *alx* riboswitches *in vivo*, which prompted us to perform co-transcriptional chemical probing experiments.

The co-transcriptional chemical probing revealed that 5'-proximal halves in both *alx* and *mntP* aptamers traverse low- T_m RNA hairpin intermediates during folding, presumably as a safeguard against premature folding of RNA into stable but non-functional folds. Surprisingly, we found that both riboswitches show Mn-dependent reactivity changes *before* the full aptamer domain has folded. The Mn-dependent structural changes in aptamer folding intermediates map to nucleotides in the *alx* and *mntP* metal-sensing cores, indicating that a partially formed metal-sensing core can interrogate the environment for Mn. Curiously, these Mn-dependent differences emerge at much earlier transcript lengths for *alx* aptamer folding, which may permit a rapid response to the subtle increase in [Mn] prompted by cytosol alkalinization(19).

Notably, other riboswitches that have been investigated with co-transcriptional SHAPE-seq or SHAPE-MaP (ZTP, ppGpp, and fluoride) did not show ligand-dependent structural changes for folding intermediates(27). Ligands such as ZTP and ppGpp are larger than an Mn ion and require a network of interactions to bind their cognate aptamers(48, 49). The resulting structural complexity of these aptamer-ligand interactions likely prevents their stable interaction with an aptamer folding intermediate. In the case of fluoride riboswitch, binding of fluoride anion to the aptamer is mediated by a network of Mg^{2+} ions to overcome steric repulsion by the phosphate backbone(50), making fluoride less likely to bind an aptamer folding intermediate. Additionally, it will be interesting to see whether other divalent-sensing riboswitches (e.g., Mg^{2+} - and Ni^{2+}/Co^{2+} -sensing riboswitches) can begin sampling for ligand before the complete aptamer has folded.

To get at the molecular basis of the divergent pH response of the *alx* and *mntP* riboswitches, we performed co-transcriptional DMS probing of RNA at both neutral and alkaline pH $\pm$ Mn. We found that pH uniquely affects the structure of the *alx* and *mntP* metal-sensing cores, providing the first clues to how pH differentially affects Mn sensing. The *alx* and *mntP* aptamers differ in their global architecture, forming a 3WJ and a 4WJ, respectively. Future study should investigate how nucleobase protonation states in the *alx* and *mntP* metal-sensing cores are coupled to junction dynamics, which controls the docking of the aptamer's helical legs. Prior smFRET study on the *X. oryzae* aptamer, a 4WJ aptamer like *mntP*, illustrated how Mn binding promotes docking of the two helical legs that in turn stabilizes the P1.1 switch helix(12). The dynamics of *alx*-like aptamers has not yet been studied but may play a role in the differential pH effect on *alx* vs. *mntP* Mn sensing and consequent riboswitching.

The stark differences in the flexibility, architecture, and pH/Mn response of the *mntP* vs *alx* riboswitches may be coupled to the cellular function of their respective Mn exporters. We showed that alkaline pH alone does not rearrange the *alx* riboswitch into a translationally active conformer; rather, the *alx* riboswitch senses the increased cytosolic [Mn] brought about by cytosol alkalinization. We hypothesize that alkaline pH lowers the K_D of the *alx* aptamer for Mn, explaining why an alkaline pH shift and consequent increase of [Mn] from 28 to 42 μ M (free and bound) activates expression of *alx* but not *mntP*(19) (Fig. 4.9). The physiological functions of such pH-enhanced expression of Mn exporters could be multi-fold. For example, Mn(II) is oxidized at alkaline pH(51, 52), deeming it unusable by Mn-dependent enzymes. Further, oxidized Mn(IV) is an extremely reactive agent able to oxidize carbon-carbon double bonds in nucleobases, the capability exploited in permanganate footprinting of protein-nucleic acid complexes(53, 54). Removal of Mn by exporters to minimize the pool of oxidized Mn could thus be a protective measure for the cell experiencing alkaline stress.

Limitations of this study. There are multiple limitations to the experimental approaches used here. First, co-transcriptional chemical probing experiments are an endpoint measurement and do not capture the true structural intermediates traversed by an RNA during transcription. Further, the DMS probing reaction is 5 min., meaning that the co-transcriptionally displayed RNA can equilibrate to some degree. As a control, we performed DMS probing experiments on equilibrium refolded intermediates, which showed structural differences. However, we cannot rule out the possibility that the RNA folds observed with co-transcriptional DMS probing are not always representative of a true folding intermediate. Lastly, since the DMS probing reactions were performed *in vitro*, the absence of ribosomal proteins may have altered the observed structural ensemble, which should be addressed in future work.

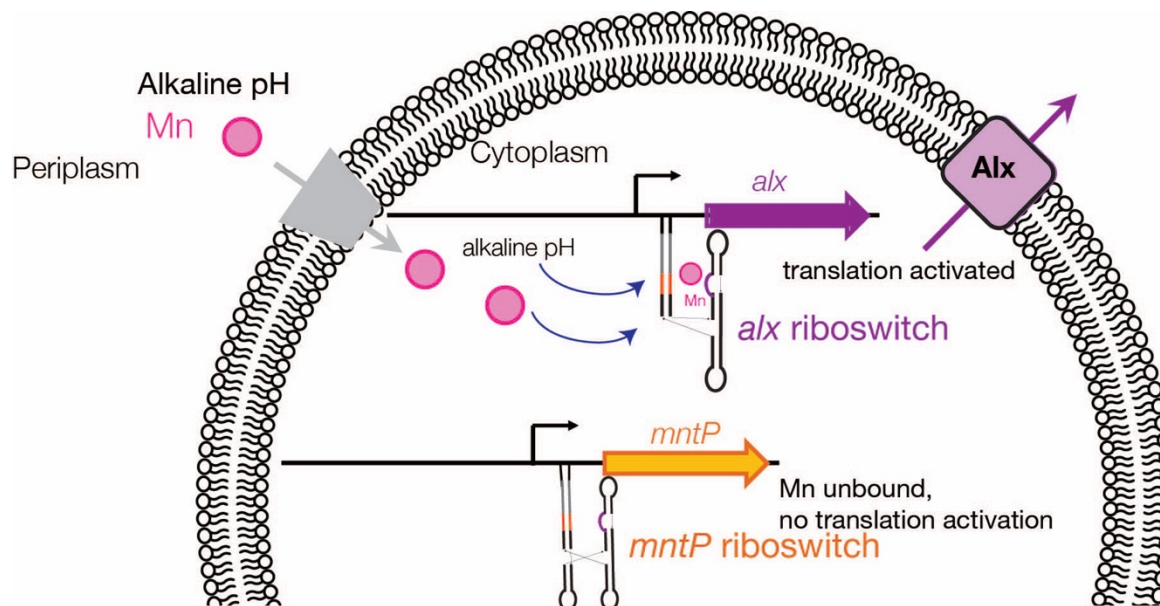


Figure 4.9 Model for *alx* and *mntP* riboregulation in *E. coli*.

Cartoon model of *alx* and *mntP* riboregulation of their respective genes in *E. coli*. At alkaline pH, the cytosolic [Mn] increases, resulting in the *alx* riboswitch turning “on” expression of Alx to prevent toxic accumulation of Mn in the cytosol. The modest increase in cytosolic [Mn] brought about by alkaline pH is insufficient to shift the *mntP* riboswitch towards its translationally active conformation, leaving *mntP* expression “off”.

Data availability

The raw sequencing data generated in this study will be publicly available in the Sequencing Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) with the BioProject accession code PRJNA1197522. Information for individual BioSamples is provided in Supplementary Tables 4.7-4.8. All processed reactivity data will be deposited in the RNA Mapping Database (<https://rmdb.stanford.edu/>) and the individual accession codes will be provided in a Supplementary Table.

Supplementary data

Supplementary Data are available at NAR online.

Author contributions

C.S. and T.V.M, conceptualization; C.S. and D.P., methodology; C.S., D.P., and C.B., investigation; C.S. and D.P., formal analysis; C.S. and D.P., visualization; C.S. and D.P., writing-initial draft; C.S., D.P., and T.V.M, writing-review and editing; T.V.M, supervision; T.V.M., funding acquisition.

Acknowledgements

We thank Dr. Eric Strobel and Dr. Courtney Szyjka for protocols and helpful discussions regarding TECprobe-VL experiments. We also thank members of our research group for their valuable feedback during the course of this study.

Funding

This work was supported by the National Institutes of Health [NIGMS ESI grant R35GM142785]; Burroughs Wellcome Fund CASI [1016945] to T.V.M.; UCSD institutional support to T.V.M; Yinan Wang Memorial Chancellor's Endowed Junior Faculty Fellowship to T.V.M; and National Science Foundation Graduate Research Fellowship Program [DGE-2038238] to D.P. Any opinions, findings, and conclusions or recommendations expressed in this

material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Funding for open access charge: National Institutes of Health.

Conflict of interest

The authors declare no competing interests.

5. References

1. Imlay, J.A. (2014) The Mismetallation of Enzymes during Oxidative Stress. *J. Biol. Chem.*, **289**, 28121–28128.
2. Imlay, J.A. (2013) The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. *Nat. Rev. Microbiol.*, **11**, 443–454.
3. Chandrangu, P., Rensing, C. and Helmann, J.D. (2017) Metal homeostasis and resistance in bacteria. *Nat. Rev. Microbiol.*, **15**, 338–350.
4. Lee, J.-W. and Helmann, J.D. (2007) Functional specialization within the Fur family of metalloregulators. *BioMetals*, **20**, 485–499.
5. Glasfeld, A., Guedon, E., Helmann, J.D. and Brennan, R.G. (2003) Structure of the manganese-bound manganese transport regulator of *Bacillus subtilis*. *Nat. Struct. Mol. Biol.*, **10**, 652–657.
6. Price, I.R., Gaballa, A., Ding, F., Helmann, J.D. and Ke, A. (2015) Mn²⁺-Sensing Mechanisms of *yypP-ykoY* Orphan Riboswitches. *Mol. Cell*, **57**, 1110–1123.
7. Dann, C.E., Wakeman, C.A., Sieling, C.L., Baker, S.C., Irnov, I. and Winkler, W.C. (2007) Structure and Mechanism of a Metal-Sensing Regulatory RNA. *Cell*, **130**, 878–892.
8. Furukawa, K., Ramesh, A., Zhou, Z., Weinberg, Z., Vallery, T., Winkler, W.C. and Breaker, R.R. (2015) Bacterial Riboswitches Cooperatively Bind Ni²⁺ or Co²⁺ Ions and Control Expression of Heavy Metal Transporters. *Mol. Cell*, **57**, 1088–1098.
9. Serganov, A. and Nudler, E. (2013) A Decade of Riboswitches. *Cell*, **152**, 17–24.
10. Dambach, M., Sandoval, M., Updegrove, T.B., Anantharaman, V., Aravind, L., Waters, L.S. and Storz, G. (2015) The Ubiquitous *yypP-ykoY* Riboswitch Is a Manganese-Responsive Regulatory Element. *Mol. Cell*, **57**, 1099–1109.
11. Kavita, K. and Breaker, R.R. (2023) Discovering riboswitches: the past and the future. *Trends Biochem. Sci.*, **48**, 119–141.
12. Suddala, K.C., Price, I.R., Dandpat, S.S., Janeček, M., Kührová, P., Šponer, J., Banáš, P., Ke, A. and Walter, N.G. (2019) Local-to-global signal transduction at the core of a Mn²⁺ sensing riboswitch. *Nat. Commun.*, **10**, 4304.

13. Sung,H.-L. and Nesbitt,D.J. (2019) Single-Molecule FRET Kinetics of the Mn²⁺ Riboswitch: Evidence for Allosteric Mg²⁺ Control of “Induced-Fit” vs “Conformational Selection” Folding Pathways. *J. Phys. Chem. B*, **123**, 2005–2015.
14. Wickiser,J.K., Winkler,W.C., Breaker,R.R. and Crothers,D.M. (2005) The Speed of RNA Transcription and Metabolite Binding Kinetics Operate an FMN Riboswitch. *Mol. Cell*, **18**, 49–60.
15. Chauvier,A., Picard-Jean,F., Berger-Dancause,J.-C., Bastet,L., Naghdi,M.R., Dubé,A., Turcotte,P., Perreault,J. and Lafontaine,D.A. (2017) Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation. *Nat. Commun.*, **8**, 13892.
16. Strobel,E.J., Cheng,L., Berman,K.E., Carlson,P.D. and Lucks,J.B. (2019) A ligand-gated strand displacement mechanism for ZTP riboswitch transcription control. *Nat. Chem. Biol.*, **15**, 1067–1076.
17. Chauvier,A., Dandpat,S.S., Romero,R. and Walter,N.G. (2024) A nascent riboswitch helix orchestrates robust transcriptional regulation through signal integration. *Nat. Commun.*, **15**, 3955.
18. Nechooshtan,G., Elgrably-Weiss,M., Sheaffer,A., Westhof,E. and Altuvia,S. (2009) A pH-responsive riboregulator. *Genes Dev.*, **23**, 2650–2662.
19. Sharma,R. and Mishanina,T.V. (2024) A riboswitch-controlled TerC family transporter Alx tunes intracellular manganese concentration in *Escherichia coli* at alkaline pH. *J. Bacteriol.*, **206**, e00168-24.
20. Zeinert,R., Martinez,E., Schmitz,J., Senn,K., Usman,B., Anantharaman,V., Aravind,L. and Waters,L.S. (2018) Structure–function analysis of manganese exporter proteins across bacteria. *J. Biol. Chem.*, **293**, 5715–5730.
21. Bachas,S.T. and Ferré-D’Amaré,A.R. (2018) Convergent Use of Heptacoordination for Cation Selectivity by RNA and Protein Metalloregulators. *Cell Chem. Biol.*, **25**, 962-973.e5.
22. Adeolu,M., Alnajar,S., Naushad,S. and S. Gupta,R. (2016) Genome-based phylogeny and taxonomy of the ‘*Enterobacteriales*’: proposal for *Enterobacterales* ord. nov. divided into the families *Enterobacteriaceae*, *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov., *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., *Morganellaceae* fam. nov., and *Budviciaceae* fam. nov. *Int. J. Syst. Evol. Microbiol.*, **66**, 5575–5599.
23. Padan,E., Bibi,E., Ito,M. and Krulwich,T.A. (2005) Alkaline pH homeostasis in bacteria: New insights. *Biochim. Biophys. Acta BBA - Biomembr.*, **1717**, 67–88.
24. Hartmann,F.S.F., Weiß,T., Shen,J., Smahajcsik,D., Savickas,S. and Seibold,G.M. (2022) Visualizing the pH in *Escherichia coli* Colonies via the Sensor Protein mCherryEA Allows High-Throughput Screening of Mutant Libraries. *mSystems*, **7**, e00219-22.

25. Smola,M.J., Rice,G.M., Busan,S., Siegfried,N.A. and Weeks,K.M. (2015) Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat. Protoc.*, **10**, 1643–1669.
26. Szyjka,C.E. and Strobel,E.J. (2022) Cotranscriptional RNA Chemical Probing. *Riboregulator Design and Analysis*, Methods in Molecular Biology. Springer US, New York, NY, Vol. 2518, pp. 291–330.
27. Szyjka,C.E. and Strobel,E.J. (2023) Observation of coordinated RNA folding events by systematic cotranscriptional RNA structure probing. *Nat. Commun.*, **14**, 7839.
28. Jolivet, P. and Foley,J. (2020) SPRI bead mix. *protocols.io*, <https://doi.org/10.17504/protocols.io.bnz4mf8w>.
29. Reuter,J.S. and Mathews,D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics*, **11**, 129.
30. Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.
31. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
32. Miller, JH (1992) A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
33. Waterhouse,A.M., Procter,J.B., Martin,D.M.A., Clamp,M. and Barton,G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.
34. Stephen,C. and Mishanina,T.V. (2022) Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the *Escherichia coli* pH-responsive riboswitch. *J. Biol. Chem.*, **298**, 102302.
35. Kaur,G., Sengupta,S., Kumar,V., Kumari,A., Ghosh,A., Parrack,P. and Dutta,D. (2014) Novel MntR-Independent Mechanism of Manganese Homeostasis in *Escherichia coli* by the Ribosome-Associated Protein HflX. *J. Bacteriol.*, **196**, 2587–2597.
36. Kaur,G., Kumar,V., Arora,A., Tomar,A., Ashish, Sur,R. and Dutta,D. (2017) Affected energy metabolism under manganese stress governs cellular toxicity. *Sci. Rep.*, **7**, 11645.
37. Martin,J.E., Waters,L.S., Storz,G. and Imlay,J.A. (2015) The *Escherichia coli* Small Protein MntS and Exporter MntP Optimize the Intracellular Concentration of Manganese. *PLOS Genet.*, **11**, e1004977.
38. Siegfried,N.A., Busan,S., Rice,G.M., Nelson,J.A.E. and Weeks,K.M. (2014) RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods*, **11**, 959–965.

39. Papp-Wallace,K.M. and Maguire,M.E. (2008) Magnesium Transport and Magnesium Homeostasis. *EcoSal Plus*, **3**, 10.1128/ecosalplus.5.4.4.2.
40. Anand,V.S. and Patel,S.S. (2006) Transient State Kinetics of Transcription Elongation by T7 RNA Polymerase. *J. Biol. Chem.*, **281**, 35677–35685.
41. Wang,M.D., Schnitzer,M.J., Yin,H., Landick,R., Gelles,J. and Block,S.M. (1998) Force and Velocity Measured for Single Molecules of RNA Polymerase. *Science*, **282**, 902–907.
42. Balint,E. and Unk,I. (2023) For the Better or for the Worse? The Effect of Manganese on the Activity of Eukaryotic DNA Polymerases. *Int. J. Mol. Sci.*, **25**, 363.
43. Zubradt,M., Gupta,P., Persad,S., Lambowitz,A.M., Weissman,J.S. and Rouskin,S. (2017) DMS-MaPseq for genome-wide or targeted RNA structure probing *in vivo*. *Nat. Methods*, **14**, 75–82.
44. Mustoe,A.M., Lama,N.N., Irving,P.S., Olson,S.W. and Weeks,K.M. (2019) RNA base-pairing complexity in living cells visualized by correlated chemical probing. *Proc. Natl. Acad. Sci.*, **116**, 24574–24582.
45. Nechooshtan,G., Elgrably-Weiss,M. and Altuvia,S. (2014) Changes in transcriptional pausing modify the folding dynamics of the pH-responsive RNA element. *Nucleic Acids Res.*, **42**, 622–630.
46. Saenger,W. (1984) Principles of Nucleic Acid Structure Springer New York, New York, NY.
47. Liberman,J.A., Suddala,K.C., Aytenfisu,A., Chan,D., Belashov,I.A., Salim,M., Mathews,D.H., Spitale,R.C., Walter,N.G. and Wedekind,J.E. (2015) Structural analysis of a class III preQ₁ riboswitch reveals an aptamer distant from a ribosome-binding site regulated by fast dynamics. *Proc. Natl. Acad. Sci.*, **112**.
48. Jones,C.P. and Ferré-D'Amaré,A.R. (2015) Recognition of the bacterial alarmone ZMP through long-distance association of two RNA subdomains. *Nat. Struct. Mol. Biol.*, **22**, 679–685.
49. Peselis,A. and Serganov,A. (2018) ykkC riboswitches employ an add-on helix to adjust specificity for polyanionic ligands. *Nat. Chem. Biol.*, **14**, 887–894.
50. Ren,A., Rajashankar,K.R. and Patel,D.J. (2012) Fluoride ion encapsulation by Mg²⁺ ions and phosphates in a fluoride riboswitch. *Nature*, **486**, 85–89.
51. Morgan,J.J. (2005) Kinetics of reaction between O₂ and Mn(II) species in aqueous solutions. *Geochim. Cosmochim. Acta*, **69**, 35–48.
52. Hu,S., Zheng,L., Zhang,H., Chen,G., Yang,Y., Ouyang,Z., Chen,S., Gao,K., Liu,C., Wang,Q. and Liu, T. (2023) Hematite-mediated Mn(II) abiotic oxidation under oxic conditions: pH effect and mineralization. *J. Colloid Interface Sci.*, **636**, 267–278.

53. Spicuglia,S., Kumar,S., Chasson,L., Payet-Bornet,D. and Ferrier,P. (2004) Potassium permanganate as a probe to map DNA–protein interactions in vivo. *J. Biochem. Biophys. Methods*, **59**, 189–194.
54. Davis,C.A., Bingman,C.A., Landick,R., Record,M.T. and Saecker,R.M. (2007) Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci.*, **104**, 7833–7838.

6. Supplementary Information for: Structurally distinct manganese-sensing riboswitch aptamers regulate diverse expression platform architectures

Authors:

Christine N. Stephen¹, Danea E. Palmer¹, Clarisa Bautista, and Tatiana V. Mishanina^{1*}

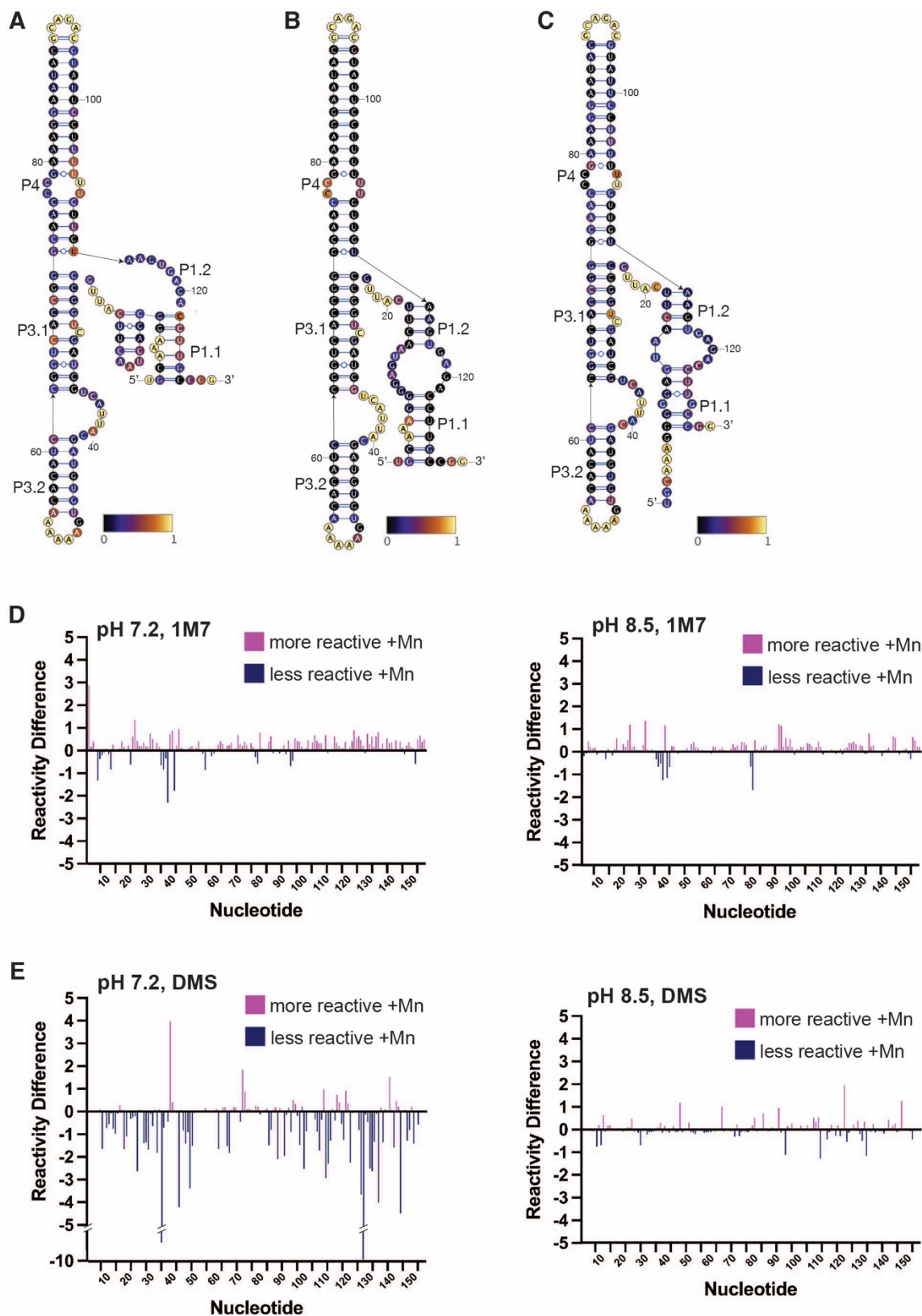
¹Department of Chemistry and Biochemistry, University of California San Diego, 9500 Gilman Dr, La Jolla, CA 92093

* To whom correspondence should be addressed. Tel: +1 858 246-5883; Email: tmishanina@ucsd.edu

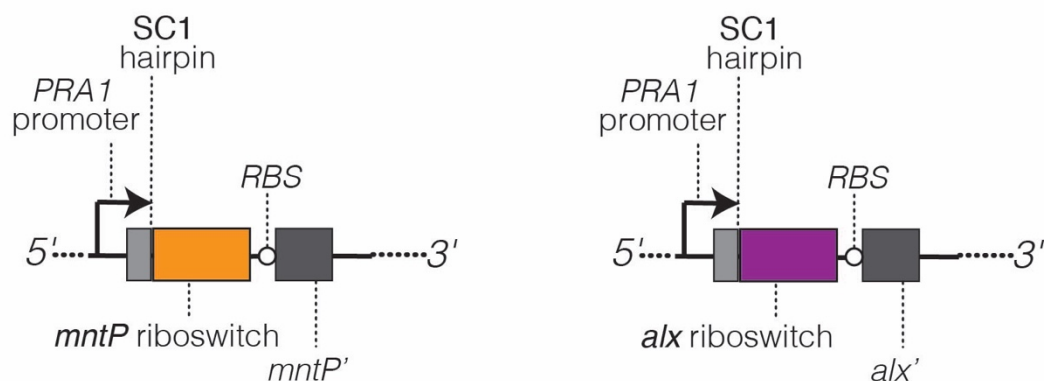
This pdf file includes: Supplementary Figures 4.1-4.32, Supplementary Notes 1-10, and Supplemental Methods.

Supplemental Figure 4.1 Structure probing of equilibrium-refolded *a/x* aptamer RNA.

Secondary structure of equilibrium-refolded *a/x* aptamer RNA modeled based on reactivity with 1M7 at **A.** pH 7.2 +Mn. **B.** pH 8.5 -Mn. **C.** pH 8.5 +Mn. **D.** Individual nucleotide reactivity differences when modified with 1M7 to compare +Mn vs. -Mn, measured at both pH 7.2 and 8.5. **E.** Individual nucleotide reactivity differences when modified with DMS to compare +Mn vs. -Mn, measured at both pH 7.2 and 8.5.



TECprobe-VL construct designs, *alx* and *mntP* riboswitches

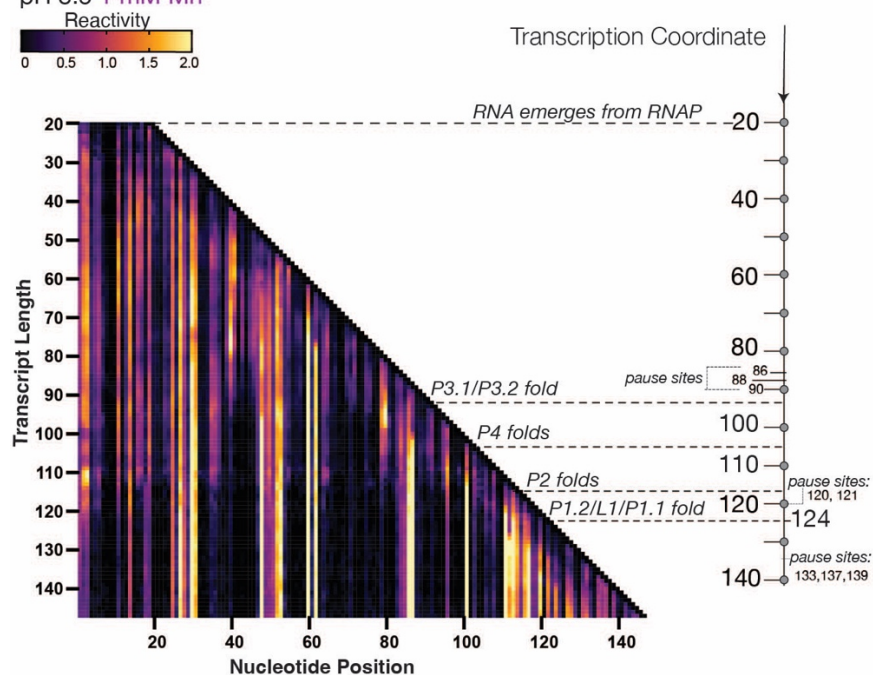


Supplemental Figure 4.2 TECprobe-VL construct designs for the *alx* and *mntP* riboswitches.

Transcription template schematics annotated with key sequence elements. All TECprobe-VL transcription templates contained the PRA1 promoter, 5' SC1 hairpin, full-length *alx* or *mntP* riboswitch sequence, the RBS, and a portion of the downstream gene sequence (termed *alx'* - nt 1-100 *alx* coding sequence, and *mntP'* - nt 1-45 *mntP* coding sequence).

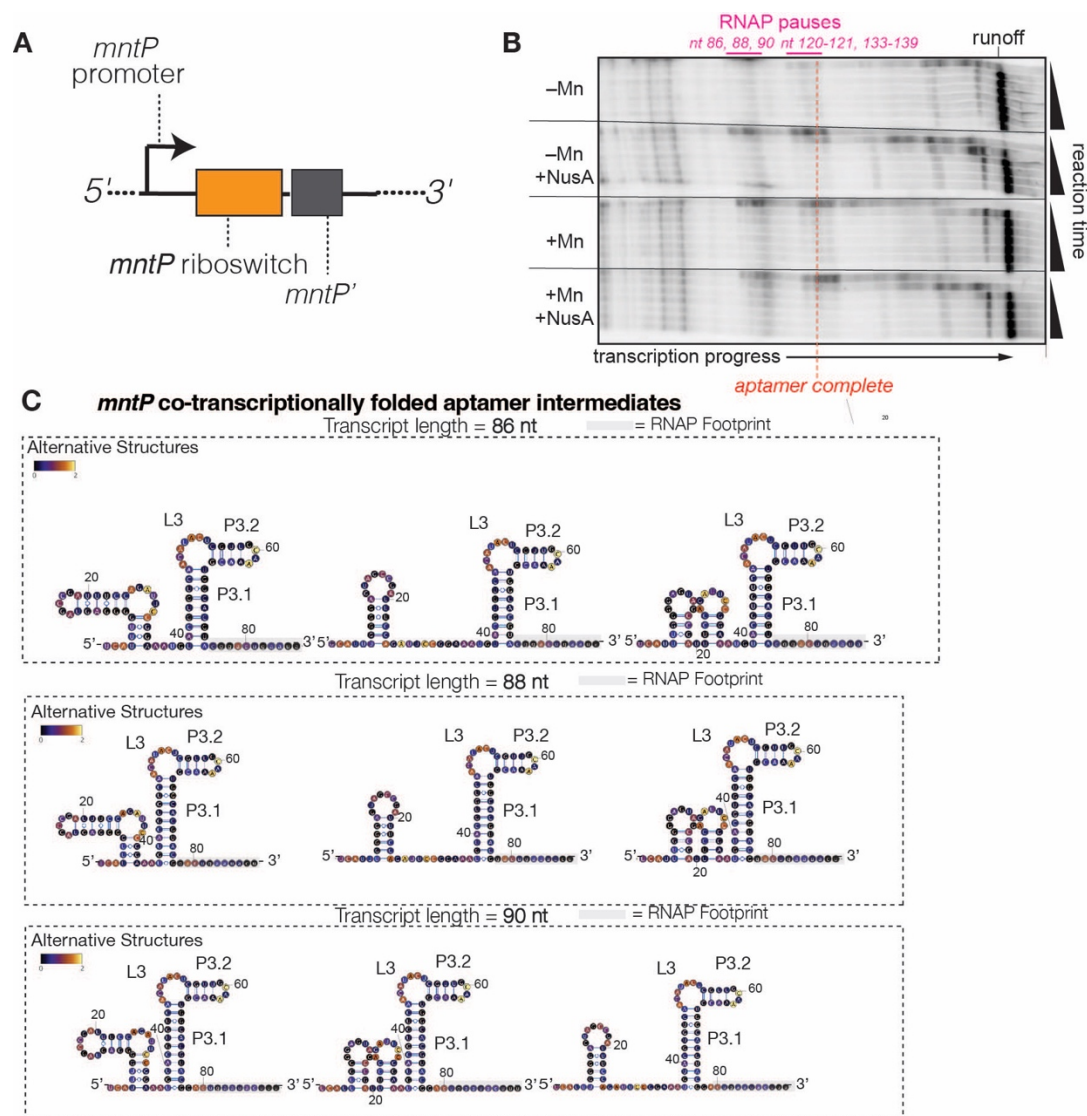
E. coli mntP riboswitch: folding of the aptamer
cotranscriptionally probed, DMS

pH 8.5 1 mM Mn



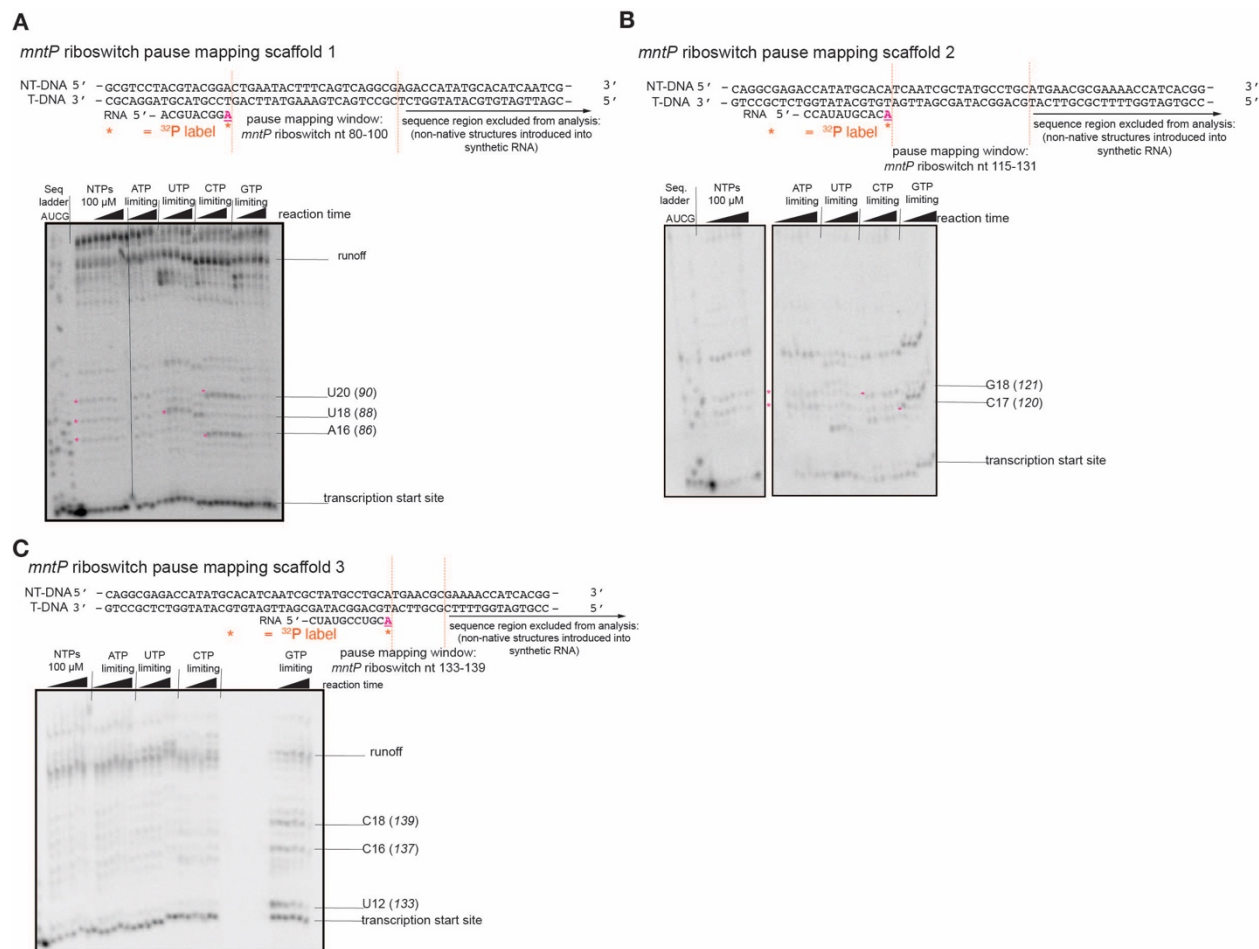
Supplemental Figure 4.3 Co-transcriptional DMS probing of the *mntP* riboswitch aptamer at alkaline pH +Mn.

TECprobe-VL DMS reactivity matrix for the *mntP* aptamer at pH 8.5 +1 mM Mn. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together.



Supplemental Figure 4.4 Synchronized, single-round *in vitro* transcription initiated from the *mntP* promoter and RNA structure models for pause-adjacent aptamer folding intermediates.

A. Transcription template schematic with key sequence features highlighted, including the native *mntP* promoter, full-length *mntP* riboswitch, and a portion of the downstream *mntP* gene (termed *mntP'*). **B.** *mntP* riboswitch transcripts were radiolabeled via incorporation of [α - 32 P]ATP and visualized on a 6% denaturing polyacrylamide gel. RNA species in the vicinity of pause sites are labeled with transcript lengths. Transcription elongation assays were performed $\pm$ Mn and $\pm$ NusA. **C.** Secondary structures of *mntP* aptamer folding intermediates, which were colored by the DMS reactivity of the indicated transcript at pH 8.5 -Mn.

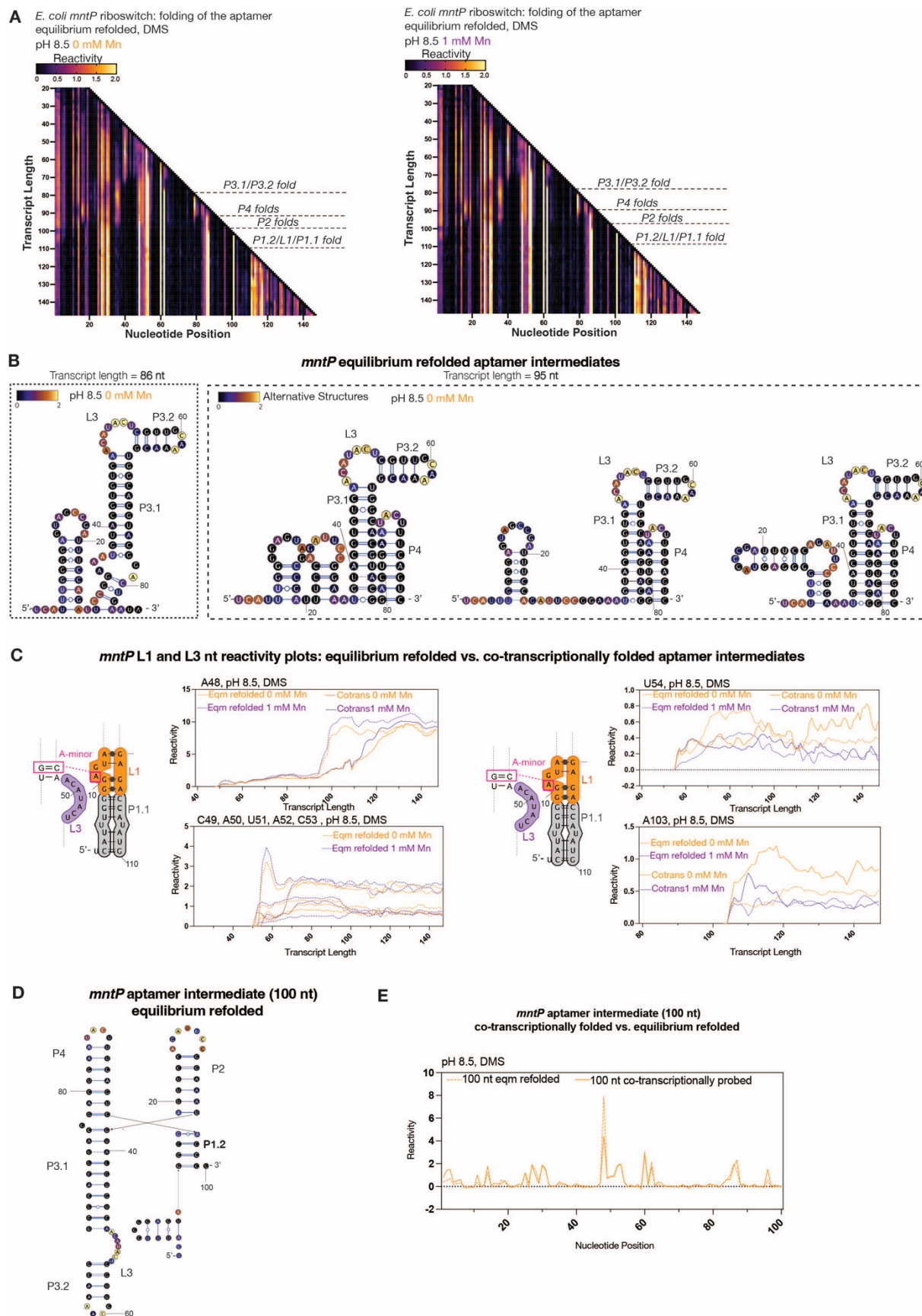


Supplemental Figure 4.5 Mapping *mntP* pause locations during aptamer and early expression platform folding.

A-C. Experiments performed on synthetic nucleic acid scaffolds to map pause sites spanning transcript lengths 80-139 nt encoded by the *mntP* riboswitch sequence. Elongation complexes were positioned upstream of the designated pause mapping window and radiolabeled by incorporation of [α - 32 P]ATP. Transcription was initiated by addition of all four NTPs with one NTP present at a limiting concentration to increase pause duration. Radiolabeled RNAs were visualized on a 15% denaturing polyacrylamide gel. All pause sites are labeled next to gel images by the 3' RNA nucleotide in the RNAP active site and the length of radiolabeled RNA, with the corresponding transcript length in full-length *mntP* riboswitch denoted in parentheses. Pause sites were located based on the sequencing ladder and by comparing results from limiting NTP time courses.

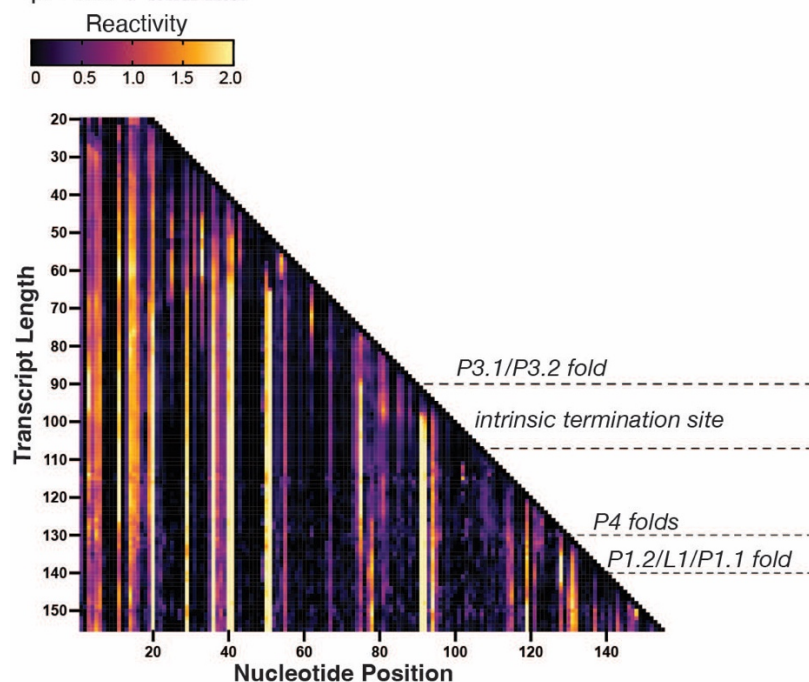
Supplemental Figure 4.6 Equilibrium DMS probing of *mntP* aptamer folding intermediates at alkaline pH $\pm$ Mn.

A. TECprobe-VL DMS reactivity matrices for *mntP* aptamer folding intermediates, which were purified and equilibrium refolded at pH 8.5 $\pm$ 1 mM Mn. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together. **B.** Secondary structures of *mntP* aptamer folding intermediates, which were purified and equilibrium refolded prior to DMS probing. Secondary structure models are colored by the DMS reactivity of the indicated transcript at pH 8.5 -Mn. **C.** Plots showing *mntP* L3 and L1 nt transcript length- and Mn- dependent reactivity changes in equilibrium refolded vs. co-transcriptionally folded aptamer intermediates. DMS data from 0 mM Mn (orange) and 1 mM Mn (purple) conditions from Fig. 4.3A, SI Fig. 4.3, and SI Fig. 4.6A. **D.** Secondary structure model for the 100 nt *mntP* aptamer folding intermediate, which was purified and equilibrium refolded prior to DMS probing. The secondary structure model is colored by the DMS reactivity at pH 8.5 -Mn. **E.** Plot of nt reactivities in the equilibrium refolded vs. co-transcriptionally folded 100 nt *mntP* aptamer intermediate, probed with DMS at pH 8.5 -Mn. The DMS reactivities for the co-transcriptionally folded RNA are from transcript length 115nt, which has 100 nt of structured RNA exposed for chemical probing due to the RNAP footprint. Data from SI Fig. 4.6A and Fig. 4.3A.



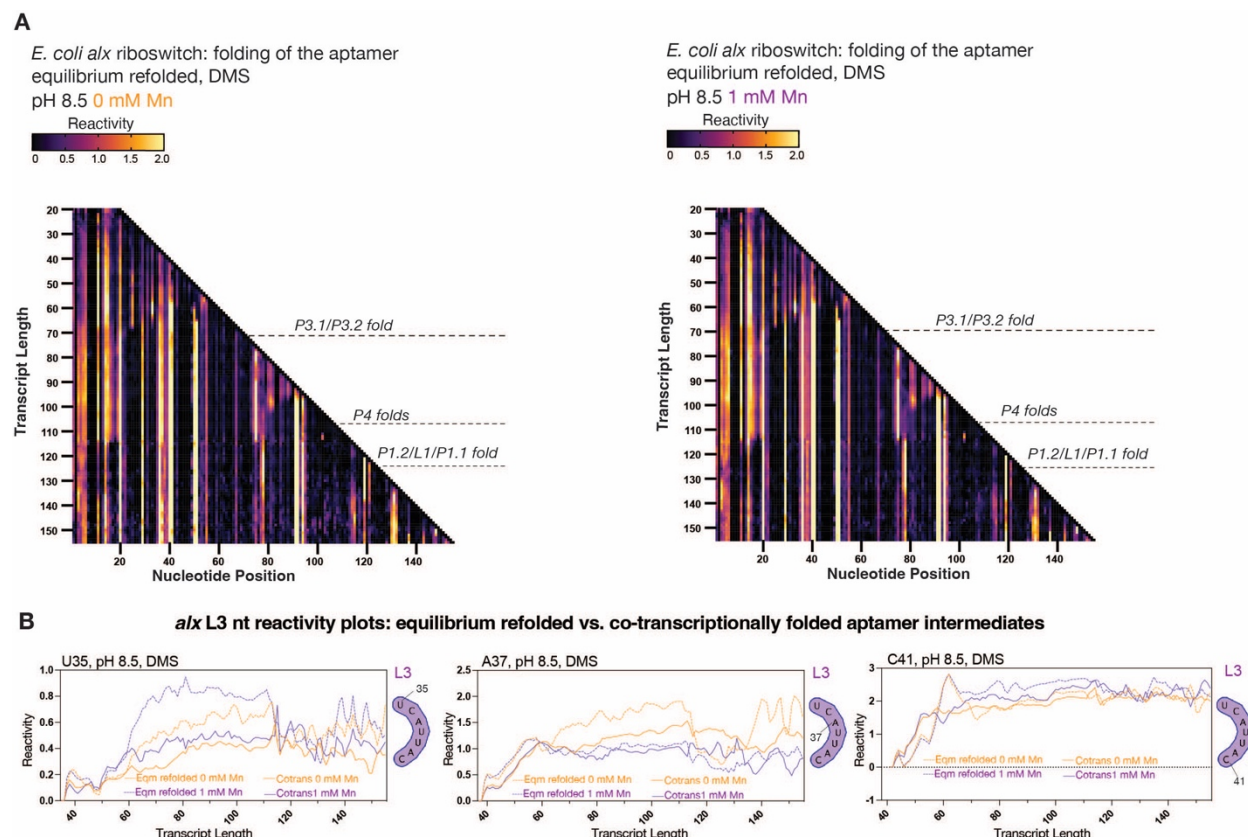
E. coli *alx* riboswitch: folding of the aptamer
cotranscriptionally probed, DMS

pH 8.5 1 mM Mn



Supplemental Figure 4.7 Co-transcriptional DMS probing of the *alx* riboswitch aptamer at alkaline pH +Mn.

TECprobe-VL reactivity matrix for the DMS-probed *alx* aptamer at pH 8.5 + 1 mM Mn. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from five independent replicates that were concatenated and analyzed together.

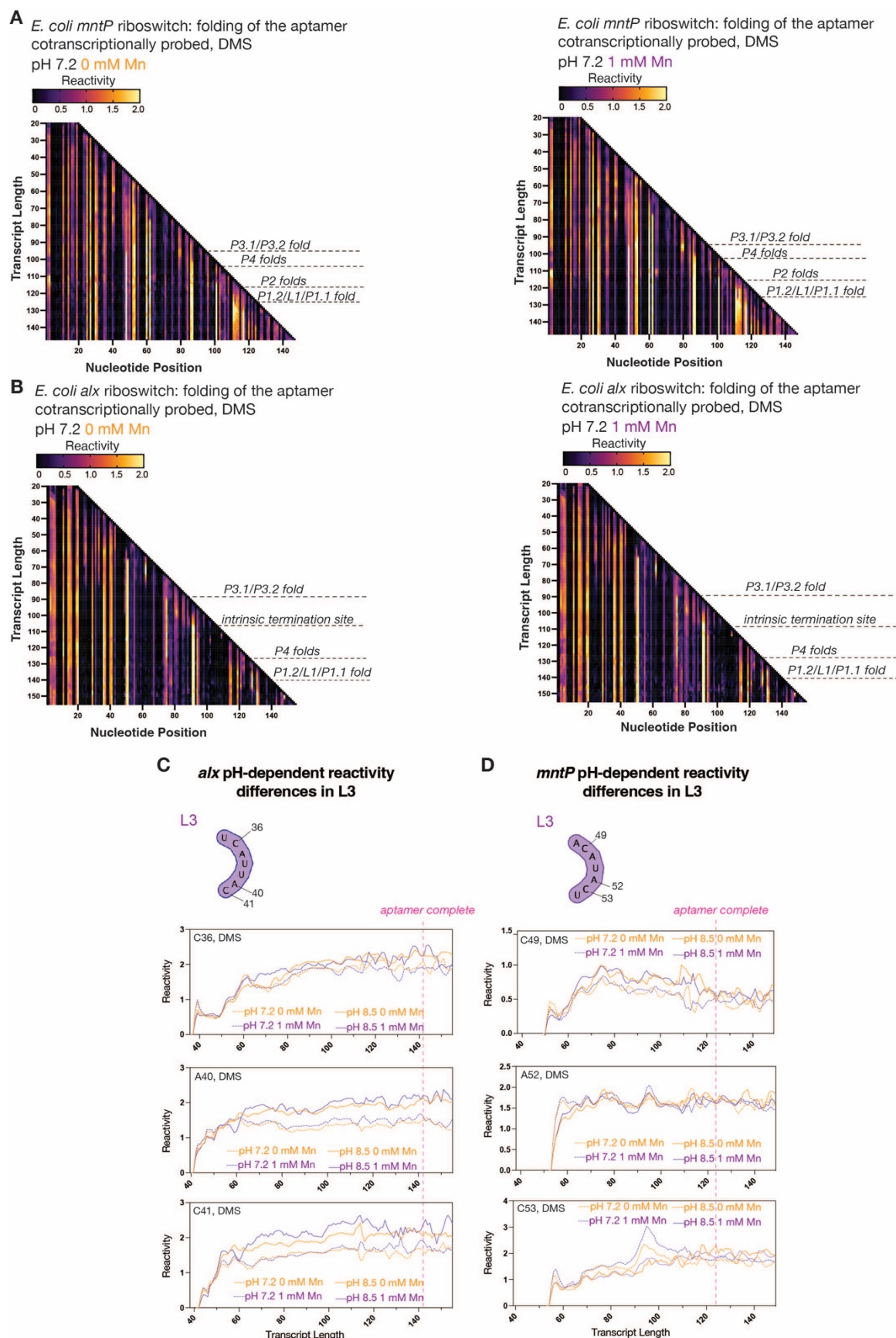


Supplemental Figure 4.8 Equilibrium DMS probing of *alx* aptamer folding intermediates at alkaline pH $\pm$ Mn.

A. TECprobe-VL DMS reactivity matrices for *alx* aptamer folding intermediates, which were purified and equilibrium refolded prior to DMS probing at pH 8.5 $\pm$ 1 mM Mn. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together. **B.** Plots showing *alx* L3 nt transcript length- and Mn-dependent reactivity changes in equilibrium refolded vs. co-transcriptionally folded aptamer intermediates at pH 8.5 $\pm$ 1 mM Mn. DMS data from 0 mM Mn (orange) and 1 mM Mn (purple) conditions are from Fig. 4.4A, Fig. 4.7, and SI Fig. 4.8A.

Supplemental Figure 4.9 Co-transcriptional DMS probing of the *alx* and *mntP* aptamers at neutral pH $\pm$ Mn and pH-dependent reactivity changes in L3.

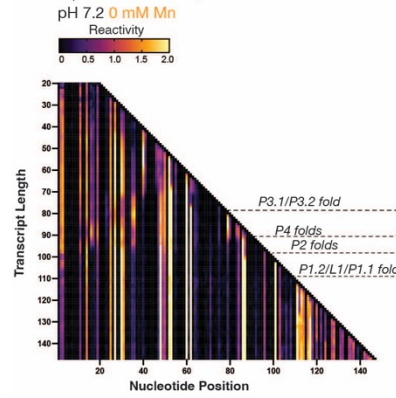
A-B. TECprobe-VL DMS reactivity matrices for the *mntP* (**A**) and *alx* (**B**) aptamers at pH 7.2 ± 1 mM Mn. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two (*mntP*, **A**) or five (*alx*, **B**) independent replicates that were concatenated and analyzed together. **C-D.** Plots showing *alx* (**C**) and *mntP* (**D**) L3 nt transcript length-, pH-, and Mn-dependent reactivity changes in co-transcriptional aptamer folding intermediates at pH 7.2 or 8.5 ± 1 mM Mn. DMS data from 0 mM Mn (orange) and 1 mM Mn (purple) are from Fig. 4.4A, Fig. 4.7, and SI Fig. 4.9B for *alx* and Fig. 4.3A, SI Fig. 4.3, and SI Fig. 4.9A for *mntP*. Vertical dotted line marks completion of aptamer folding.



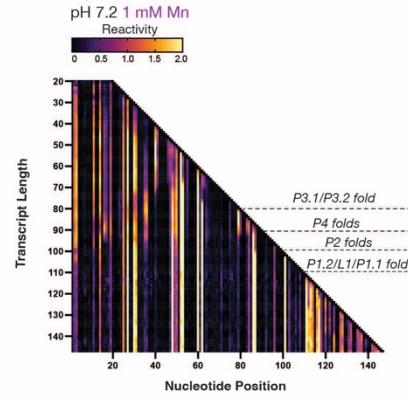
Supplemental Figure 4.10 Equilibrium DMS probing of *alx* and *mntP* aptamer folding intermediates at neutral pH $\pm$ Mn and pH-dependent reactivity changes in L3.

A-B. TECprobe-VL DMS reactivity matrices for the *mntP* (**A**) and *alx* (**B**) aptamer intermediates, which were purified and equilibrium refolded at pH 7.2 $\pm$ 1 mM Mn. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together. **C-D.** Plots showing *alx* (**C**) and *mntP* (**D**) L3 transcript length-, pH-, and Mn-dependent reactivity changes in equilibrium refolded aptamer intermediates at pH 7.2 or 8.5 $\pm$ 1 mM Mn. DMS data from 0 mM Mn (orange) and 1 mM Mn (purple) are from SI Fig. 4.8A and SI Fig. 4.10B for *alx* and SI Fig. 4.6A and SI Fig. 4.10A for *mntP*. Vertical dotted line marks completion of aptamer folding.

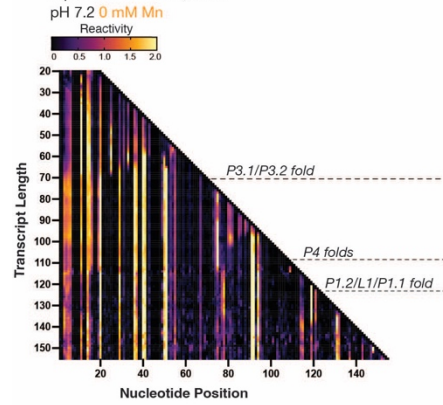
A *E. coli mntP* riboswitch: folding of the aptamer equilibrium refolded, DMS



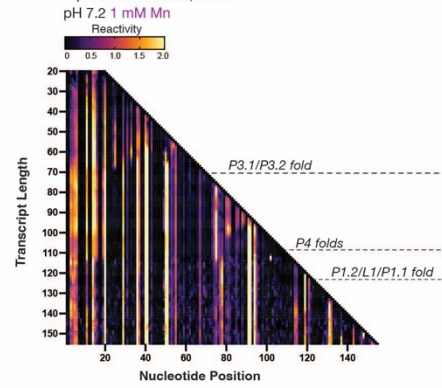
E. coli mntP riboswitch: folding of the aptamer equilibrium refolded, DMS



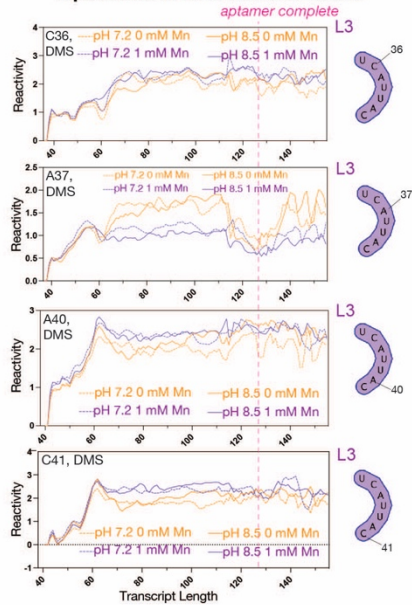
B *E. coli alx* riboswitch: folding of the aptamer equilibrium refolded, DMS



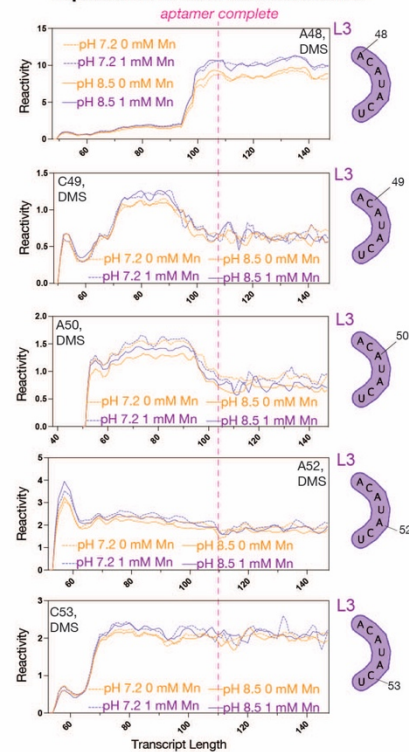
E. coli alx riboswitch: folding of the aptamer equilibrium refolded, DMS

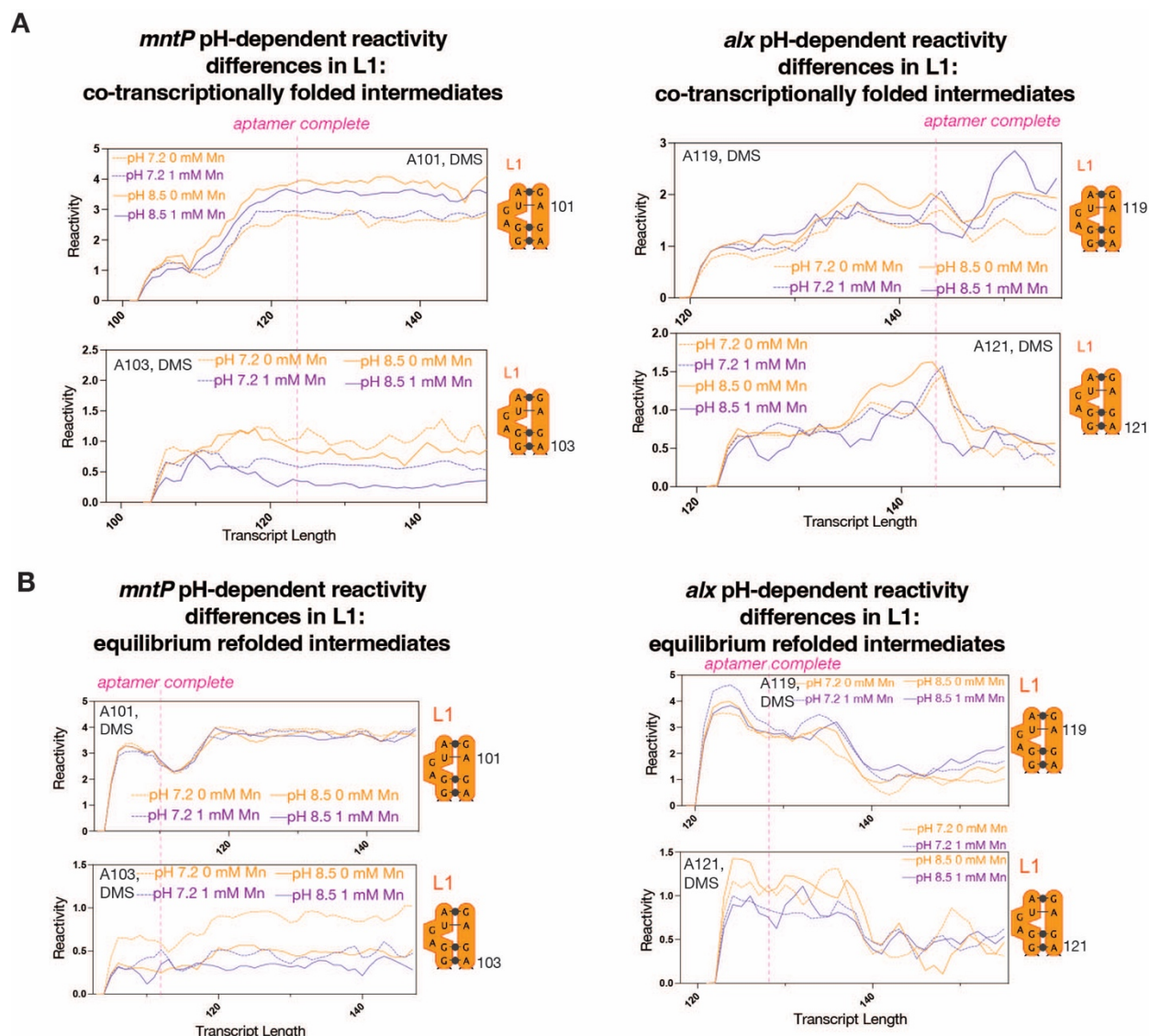


C *alx* pH-dependent reactivity differences in L3: equilibrium refolded intermediates



D *mntP* pH-dependent reactivity differences in L3: equilibrium refolded intermediates



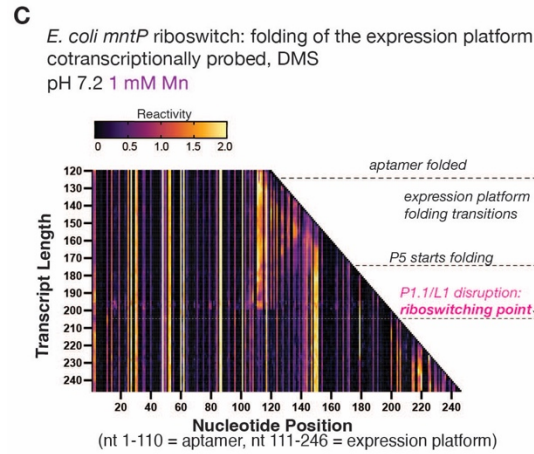
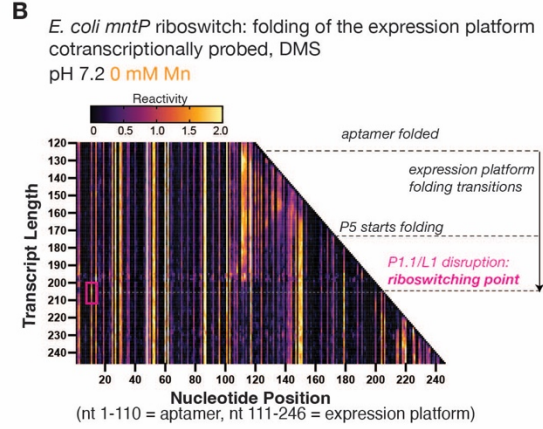
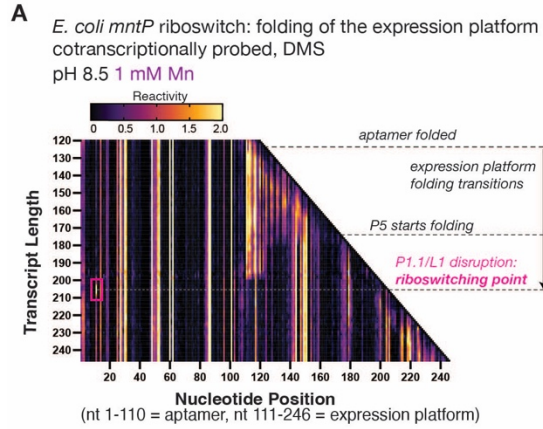


Supplemental Figure 4.11 pH-dependent reactivity changes in *alx* and *mntP* L1 in co-transcriptionally folded and equilibrium refolded aptamer intermediates.

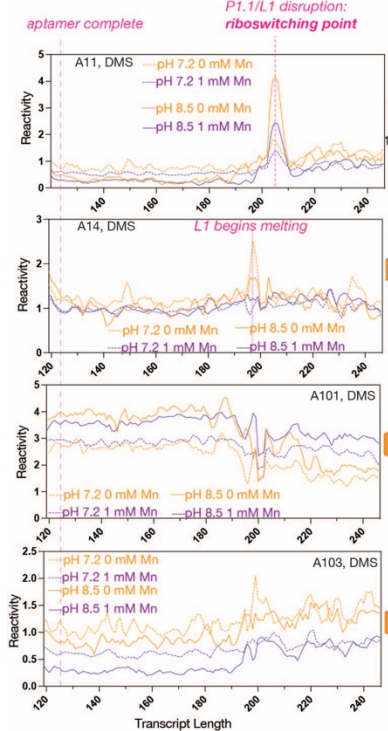
A. Plots showing *alx* and *mntP* L1 nt transcript length-, pH-, and Mn-dependent reactivity changes in co-transcriptionally folded aptamer intermediates at pH 7.2 or 8.5 $\pm$ 1 mM Mn. DMS data from 0 mM (orange) and 1 mM Mn (purple) conditions are from Fig. 4.4A, SI Fig. 4.7, and SI Fig. 4.9B for *alx* and Fig. 4.3A, SI Fig. 4.3, and SI Fig. 4.9A for *mntP*. Vertical dotted line marks completion of aptamer folding. **B.** Plots showing *alx* and *mntP* L1 nt transcript length-, pH-, and Mn-dependent reactivity changes in equilibrium refolded aptamer intermediates at pH 7.2 or 8.5 $\pm$ 1 mM Mn. DMS data from 0 mM (orange) and 1 mM Mn (purple) conditions from SI Fig. 4.8A and SI Fig. 4.10B for *alx* and SI Fig. 4.6A and SI Fig. 4.10A for *mntP*. Vertical dotted line marks completion of aptamer folding.

Supplemental Figure 4.12 Co-transcriptional DMS probing of *mntP* riboswitch during expression platform synthesis at neutral and alkaline pH.

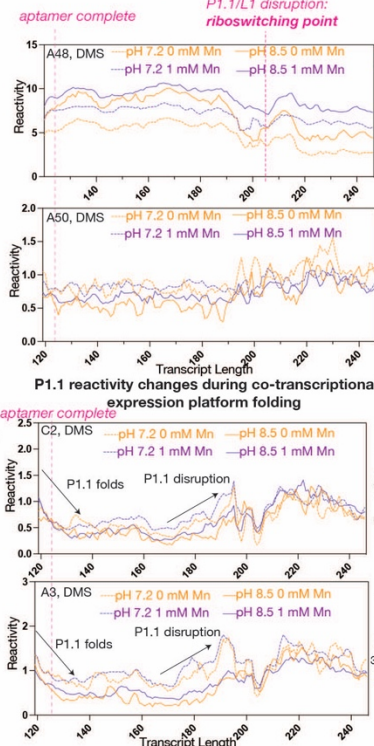
A-C. TECprobe-VL DMS reactivity matrices for *mntP* folding intermediates at pH 8.5 +1 mM Mn (**A**), pH 7.2 (**B**), and pH 7.2 + 1 mM Mn (**C**). Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together. **D.** Plots showing *mntP* L1, L3, and P1.1 nt transcript length-, pH-, and Mn-dependent reactivity changes in co-transcriptionally folded intermediates at pH 7.2 or 8.5 ± 1 mM Mn. DMS data from 0 mM Mn (orange) and 1 mM Mn (purple) conditions from Fig. 4.6A-B, and SI Fig. 4.12A-C. Vertical pink dotted lines mark key co-transcriptional RNA folding transitions.



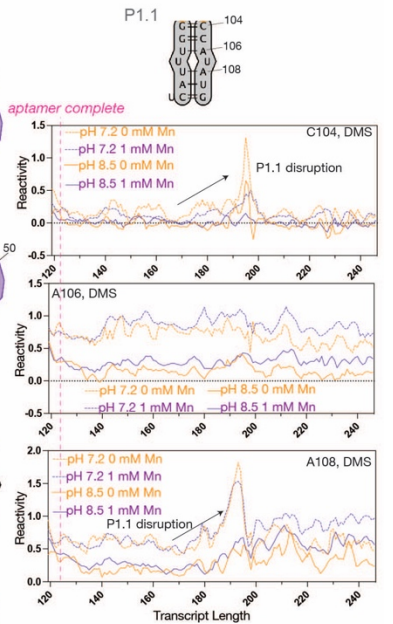
D L1 reactivity changes during co-transcriptional expression platform folding



L3 reactivity changes during co-transcriptional expression platform folding



P1.1 reactivity changes during co-transcriptional expression platform folding



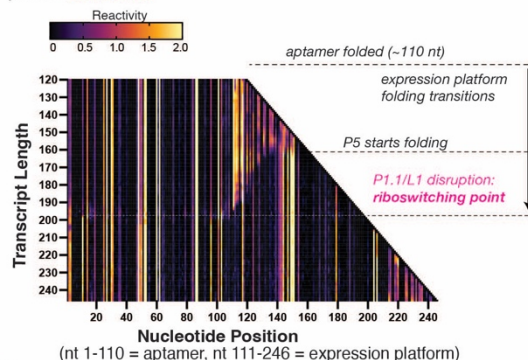
Supplemental Figure 4.13 Equilibrium DMS probing on *mntP* folding intermediates during expression platform synthesis.

A-C. TECprobe-VL DMS reactivity matrices for *mntP* folding intermediates, which were purified and equilibrium refolded at pH 7.2 ± 1 mM Mn (**A**) or pH 8.5 ± 1 mM Mn (**B**). Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together. **C.** Plot showing *mntP* A11 transcript length-, pH-, and Mn-dependent reactivity changes in equilibrium refolded intermediates at pH 7.2 or 8.5 ± 1 mM Mn. Data for 0 mM Mn (orange) and 1 mM Mn (purple) from SI Figs. 4.13A-B. Vertical dotted lines mark indicated RNA structural transitions. **D.** Plots showing *mntP* L3 nt transcript length- and Mn-dependent reactivity changes in co-transcriptionally folded vs. equilibrium refolded intermediates, which were DMS-probed at the designated pH value. DMS data for 0 mM Mn (orange) and 1 mM Mn (purple) conditions from Fig. 4.6A-B, SI Fig. 4.12A-C, and SI Figs. 4.13A-B. Vertical dotted lines mark indicated co-transcriptional RNA folding transitions.

A

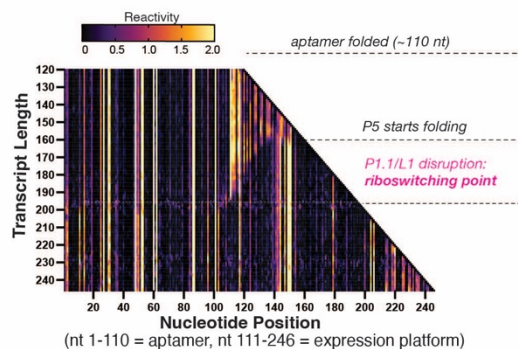
E. coli mntP riboswitch: folding of the expression platform equilibrium refolded, DMS

pH 7.2 0 mM Mn



E. coli mntP riboswitch: folding of the expression platform equilibrium refolded, DMS

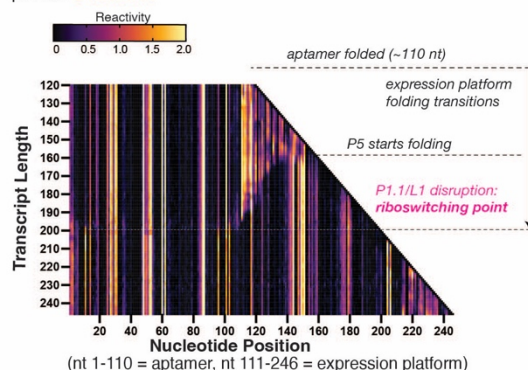
pH 7.2 1 mM Mn



B

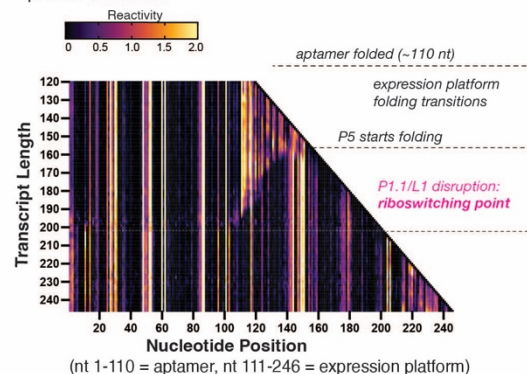
E. coli mntP riboswitch: folding of the expression platform equilibrium refolded, DMS

pH 8.5 0 mM Mn



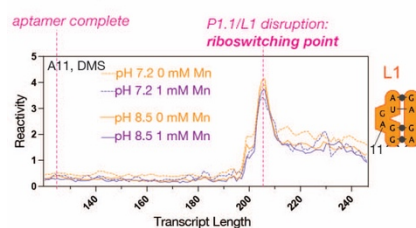
E. coli mntP riboswitch: folding of the expression platform equilibrium refolded, DMS

pH 8.5 1 mM Mn



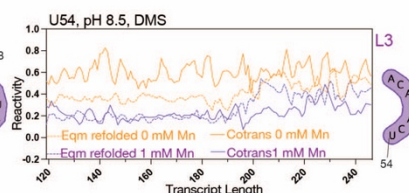
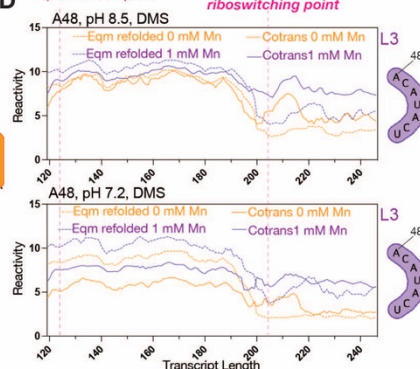
C

A11 Mn-dependent reactivity spike abolished for equilibrium refolded *mntP* intermediates



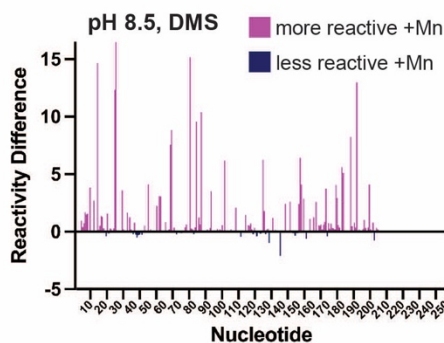
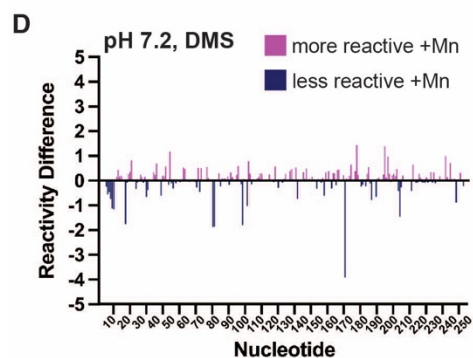
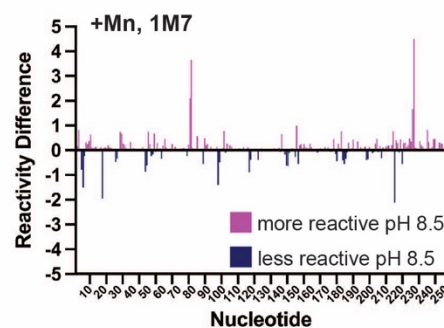
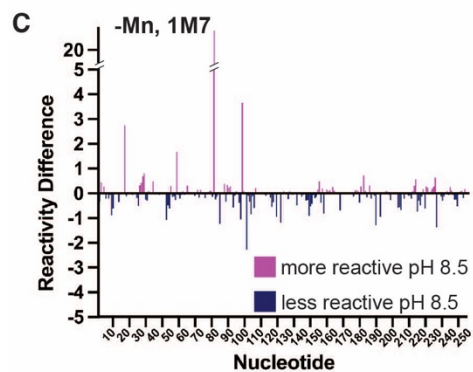
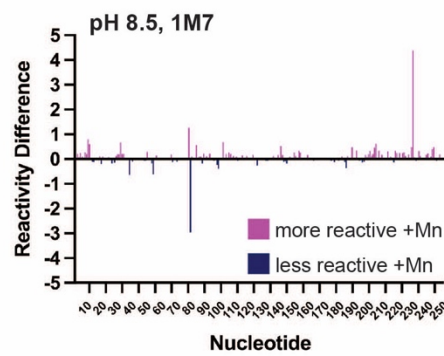
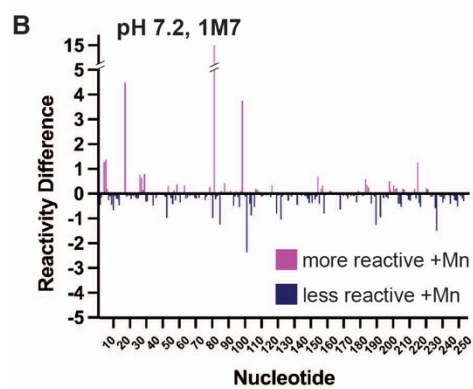
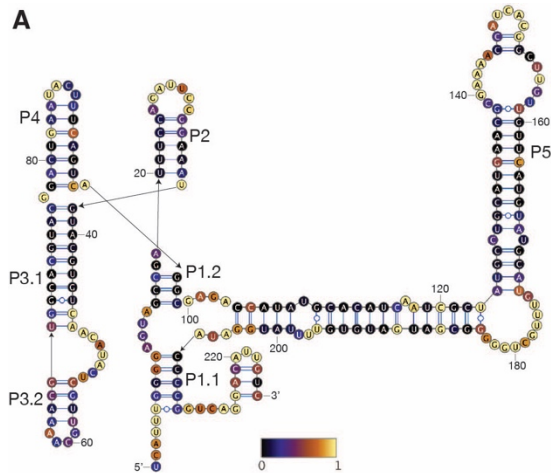
D

aptamer complete P1.1/L1 disruption: riboswitching point



Supplemental Figure 4.14 Structure probing of equilibrium-refolded *mntP* full-length riboswitch RNA.

A. Secondary structure of equilibrium-refolded *mntP* full-length RNA modeled based on reactivity with 1M7 at pH 7.2 -Mn. **B.** Individual nucleotide reactivity differences when modified with 1M7 to compare +Mn vs. -Mn, measured at both pH 7.2 and 8.5. **C.** Individual nucleotide reactivity differences at neutral pH (7.2) vs. alkaline pH (8.5), measured both in the presence and absence of Mn. **D.** Individual nucleotide reactivity differences when modified with DMS to compare +Mn vs. -Mn, measured at both pH 7.2 and 8.5.

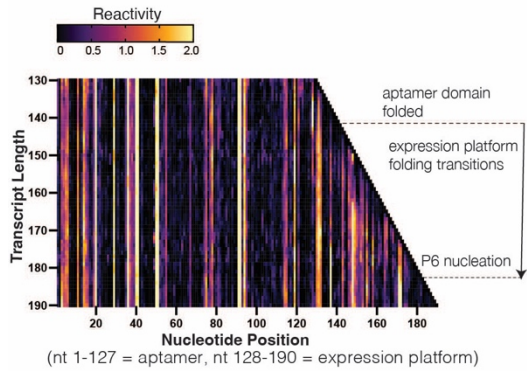


Supplemental Figure 4.15 Co-transcriptional DMS probing of the *a/x* riboswitch during expression platform synthesis at alkaline pH.

A. TECprobe-VL DMS reactivity matrix for *a/x* folding intermediates at pH 8.5 +1 mM Mn. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from five independent replicates that were concatenated and analyzed together. **B-D.** Secondary structure models for *a/x* expression platform folding intermediates, which were co-transcriptionally folded prior to DMS probing. The secondary structure models are colored by DMS reactivity at pH 8.5 at the designated Mn concentration. To simplify the structure modeling, we executed the RNAstructure “Fold” command constrained by DMS reactivities for the expression platform sequence only, with the L1/P1.1 falling stem either included or excluded.

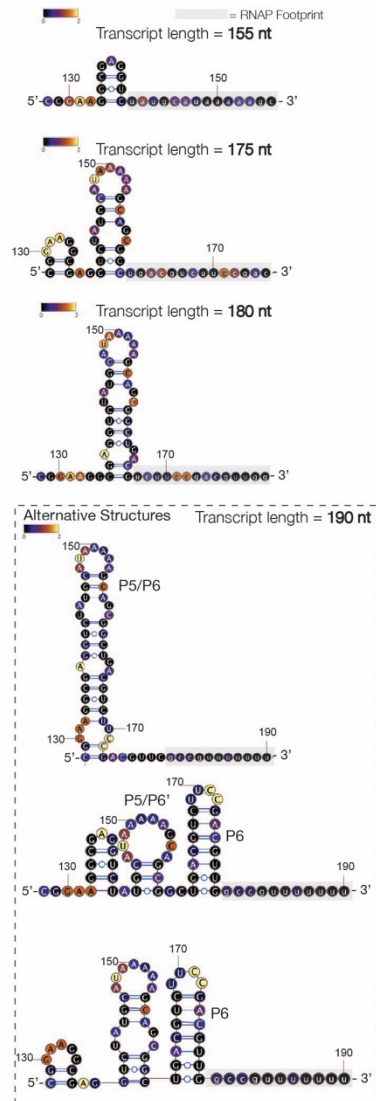
A

E. coli alx riboswitch: folding of the expression platform
cotranscriptionally probed, DMS
pH 8.5 1 mM Mn



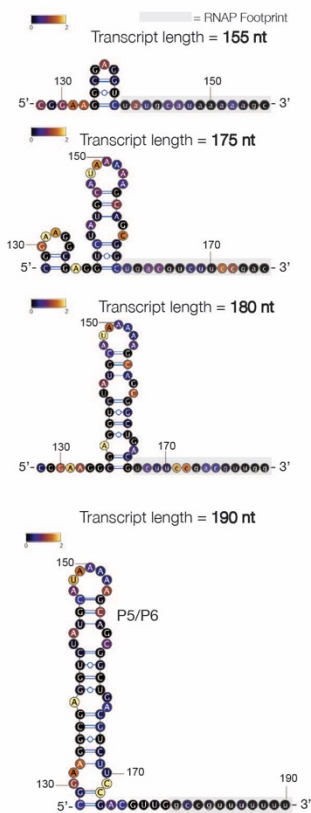
B

expression platform folding intermediates:
pH 8.5 0 mM Mn; P1.1/L1 falling stem sequence
not included



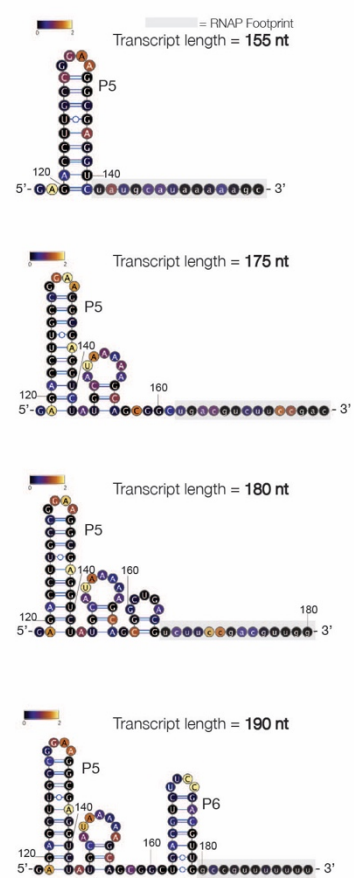
C

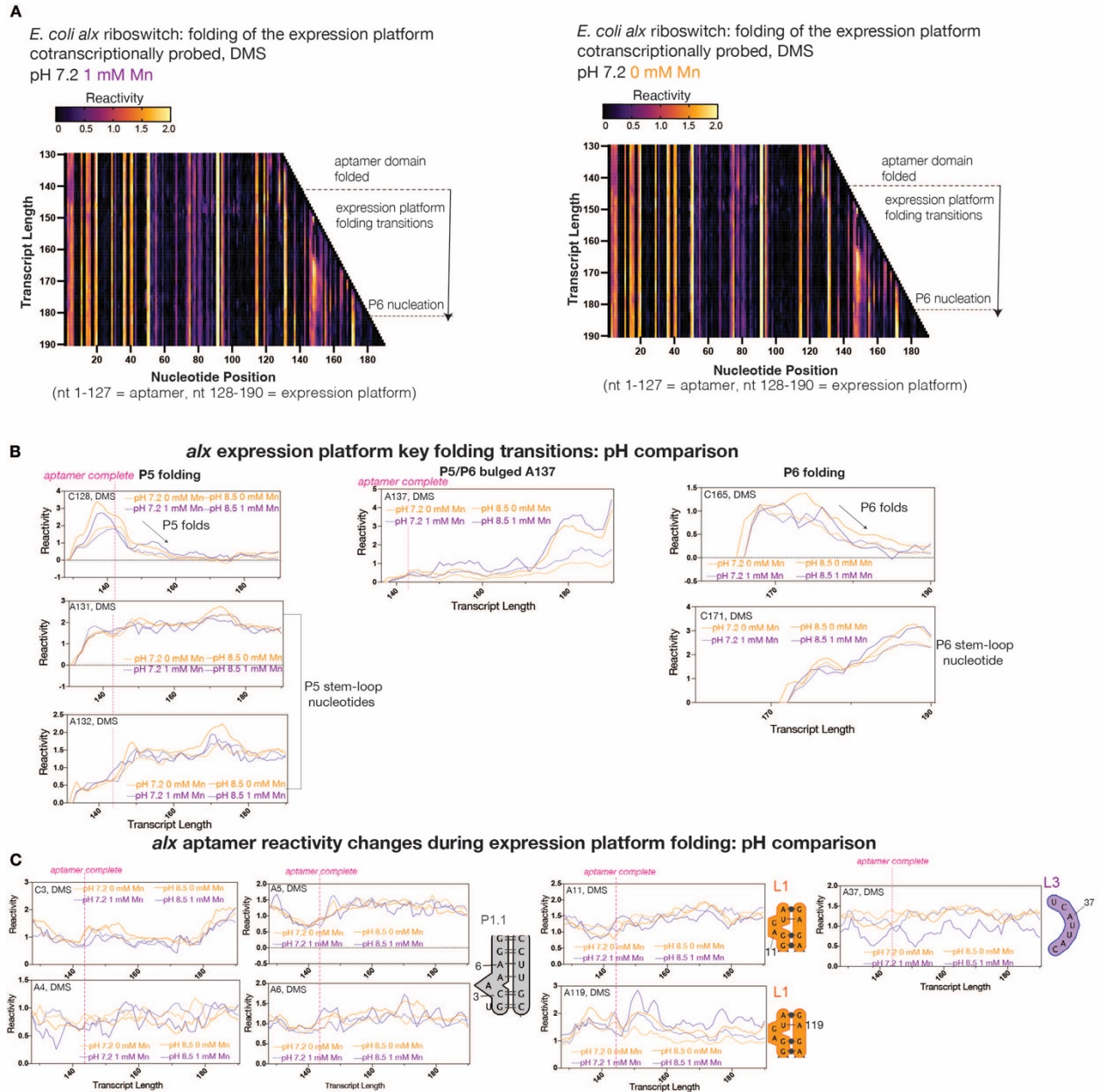
expression platform folding intermediates:
pH 8.5 1 mM Mn; P1.1/L1 falling stem sequence
not included



D

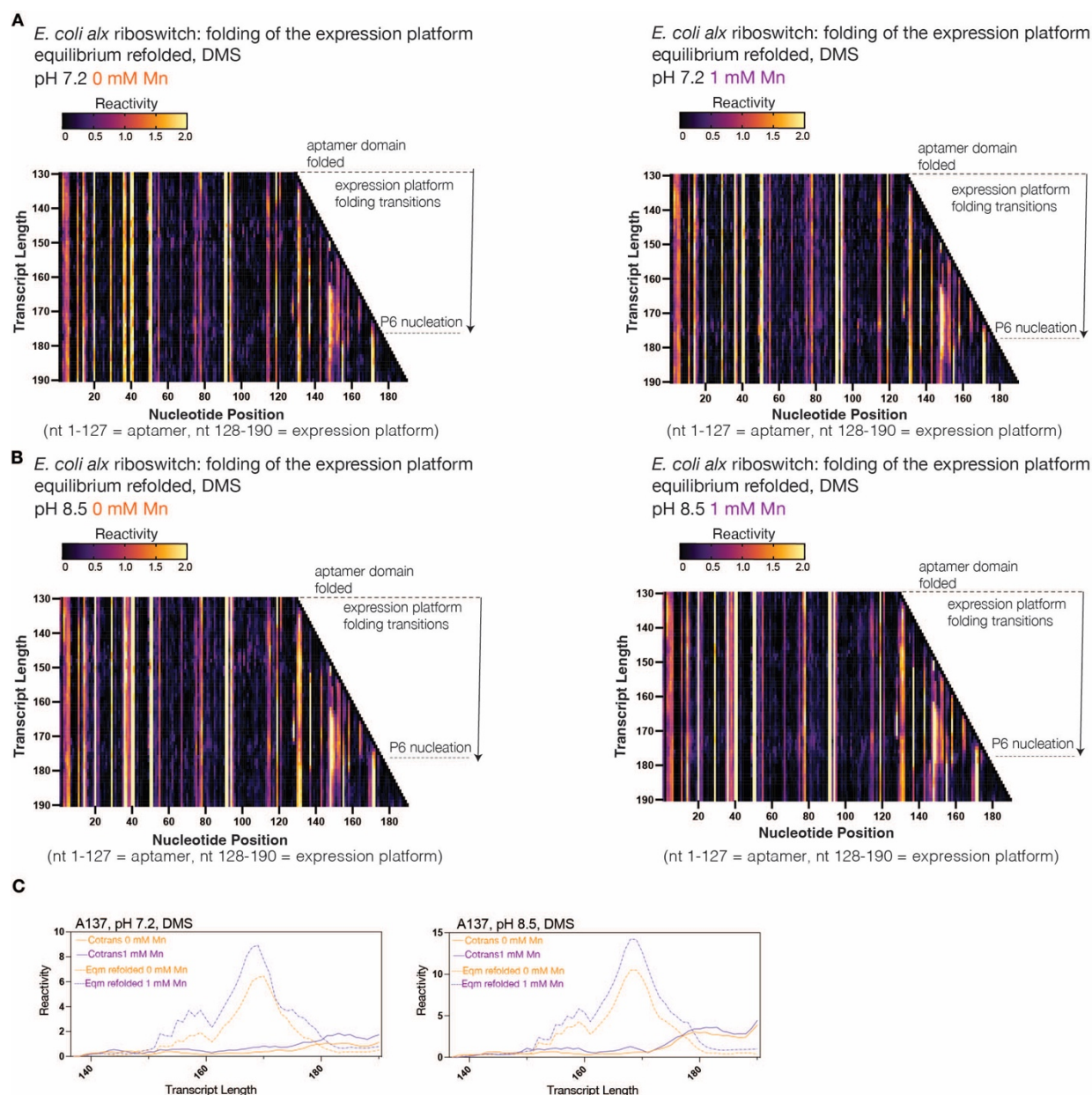
expression platform folding intermediates:
pH 8.5 1 mM Mn; P1.1/L1 falling stem sequence
included





Supplemental Figure 4.16 Co-transcriptional DMS probing of the *alx* riboswitch during expression platform synthesis at neutral pH and pH-dependent reactivity differences in the metal-sensing core.

A. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates at pH 7.2 ± 1 mM Mn. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from five independent replicates that were concatenated and analyzed together. **B.** Plots showing *alx* expression platform nt transcript length-, pH-, and Mn-dependent reactivity changes in co-transcriptionally folded intermediates at pH 7.2 or 8.5 ± 1 mM Mn. DMS data for 0 mM Mn (orange) and 1 mM Mn (purple) from Fig. 4.7A SI Fig. 4.15A, and SI Fig. 4.16A. Vertical dotted lines mark indicated RNA folding transitions. **C.** Plots showing *alx* aptamer nt transcript length-, pH-, and Mn-dependent reactivity changes in co-transcriptionally folded intermediates at pH 7.2 or 8.5 ± 1 mM Mn. DMS Data for 0 mM Mn (orange) and 1 mM Mn (purple) from Fig. 4.8A, SI Fig. 4.15A, and SI Fig. 4.16A. Vertical dotted lines mark indicated RNA folding transitions.

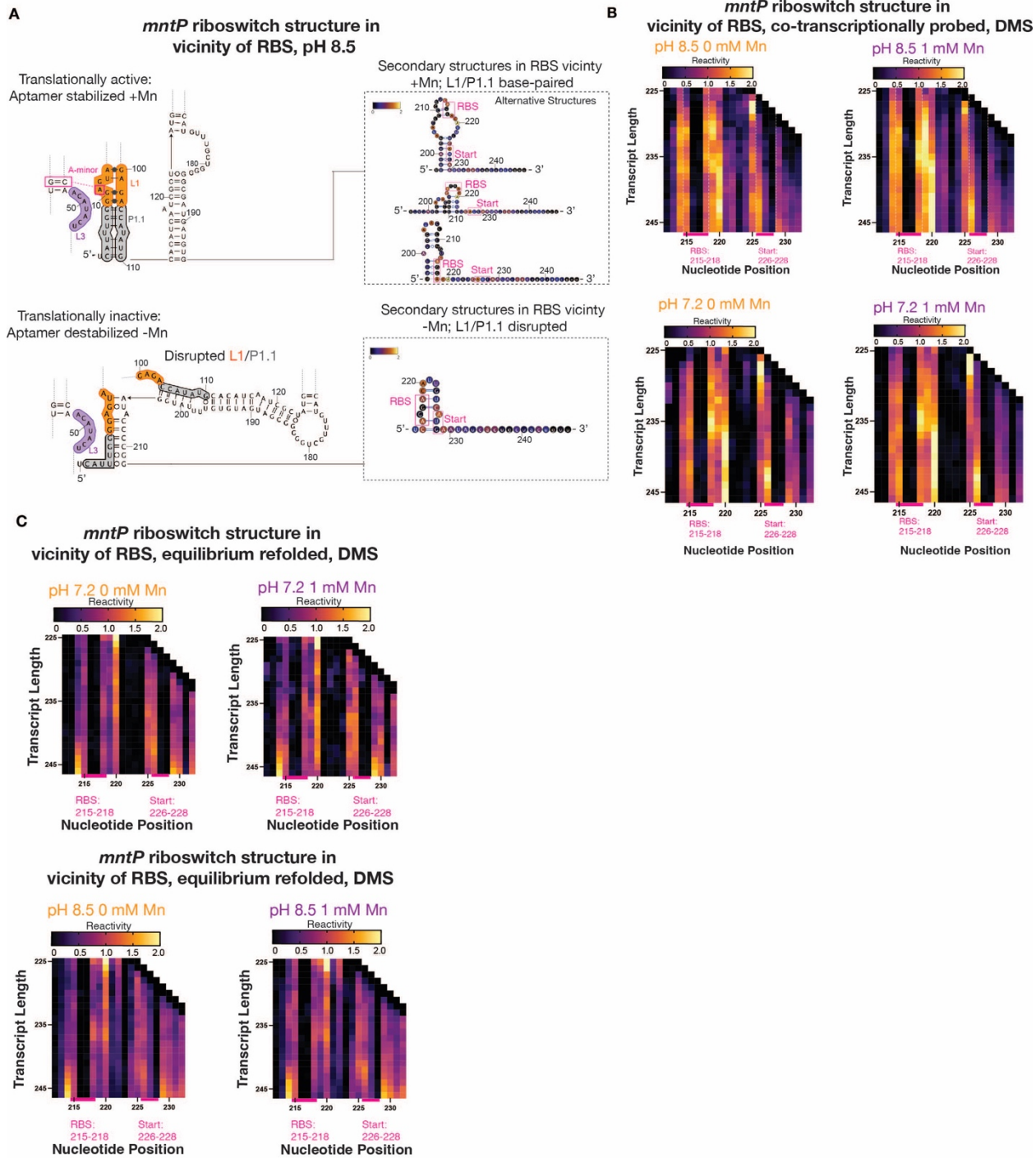


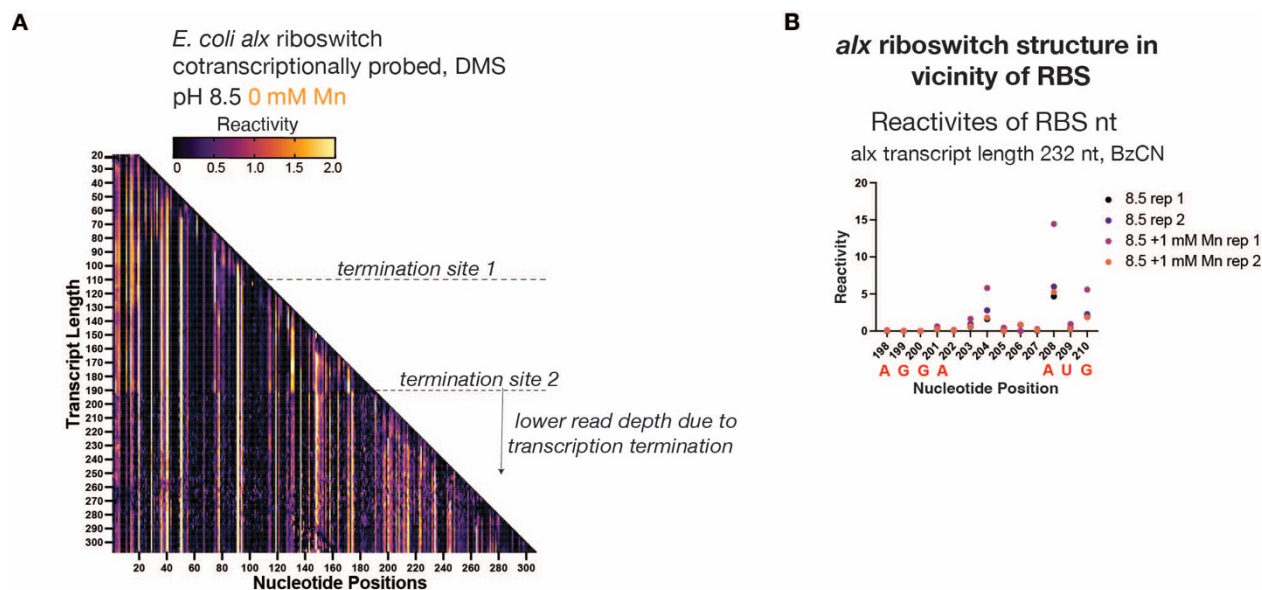
Supplemental Figure 4.17 Equilibrium DMS probing of the *alx* riboswitch during expression platform synthesis at neutral vs. alkaline pH and Mn-dependent reactivity differences in the expression platform.

A-B. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates, which were purified and equilibrium refolded at pH 7.2 $\pm$ 1 mM Mn (**A**) or pH 8.5 $\pm$ 1 mM Mn (**B**). Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together. **C.** Plots showing comparison of A137 transcript length- and Mn-dependent reactivity changes in co-transcriptionally folded vs. equilibrium refolded intermediates at the designated pH $\pm$ 1 mM Mn. DMS data for 0 mM Mn (orange) and 1 mM Mn (purple) from Fig. 4.7A, SI Fig. 4.15A, SI Fig. 4.16A, and SI Fig. 4.17A-B.

Supplemental Figure 4.18 DMS reactivities of *mntP* RBS nt in co-transcriptionally folded and equilibrium refolded intermediates.

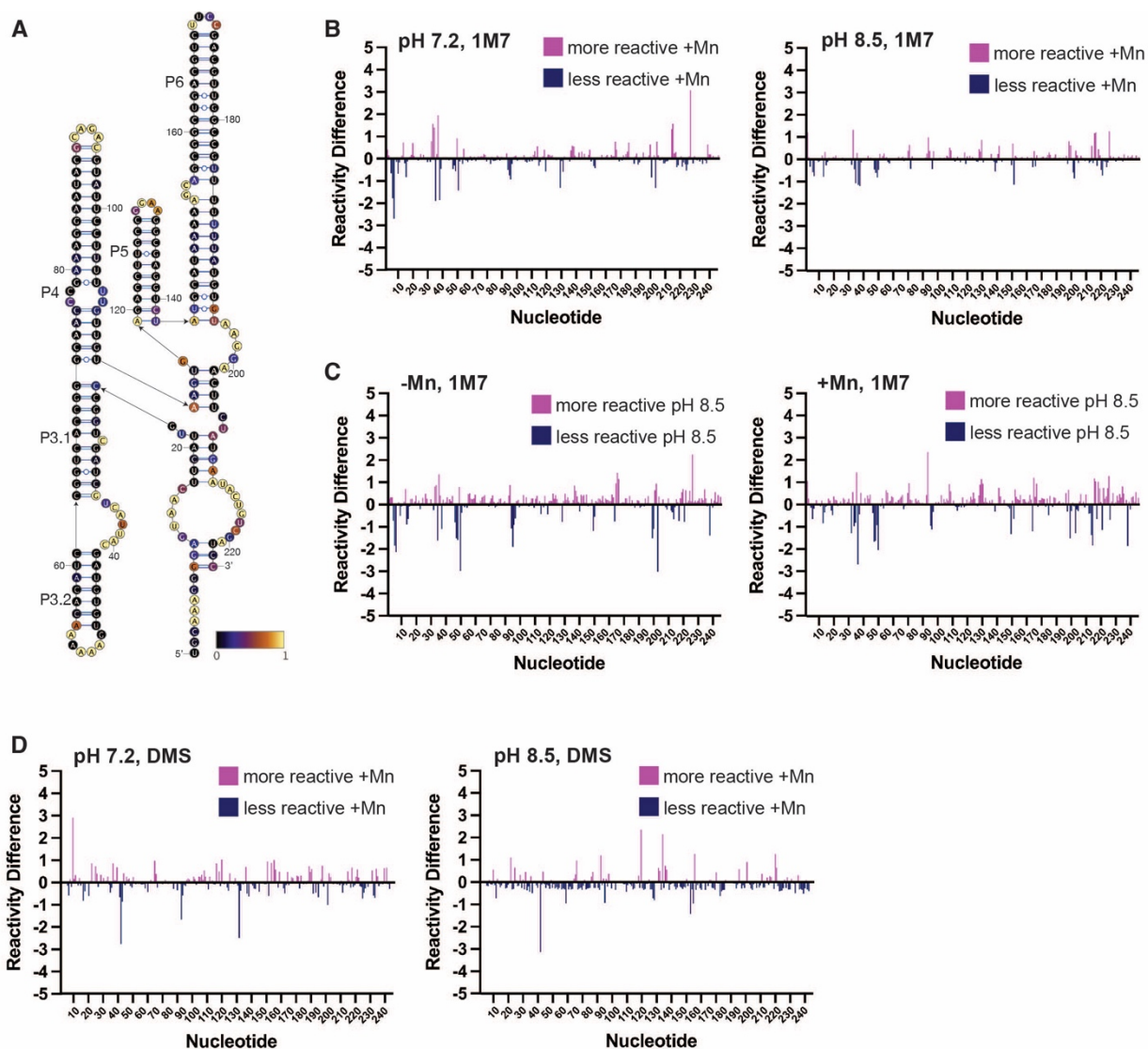
A. (Top panel) Truncated secondary structure prediction of full-length, Mn-bound *mntP* riboswitch, with L1 and P1.1 stabilized. Adjacent are secondary structure models for *mntP* expression platform structure at transcript length 246 nt, which are colored by DMS reactivity at pH 8.5 +Mn. (Bottom panel) Truncated secondary structure prediction of full-length, Mn-unbound *mntP* riboswitch, with L1 and P1.1 disrupted. Adjacent is a secondary structure model for *mntP* expression platform structure at transcript length 246 nt, which is colored by DMS reactivity at pH 8.5 -Mn. **B.** TECprobe-VL DMS reactivity matrix for co-transcriptionally folded *mntP* intermediates at pH 7.2 and 8.5 ± 1 mM Mn., windowing over transcript lengths containing the RBS and start codon. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together. **C.** TECprobe-VL DMS reactivity matrix for equilibrium refolded *mntP* intermediates at pH 7.2 and 8.5 ± 1 mM Mn., windowing over transcripts lengths containing the RBS and start codon nt. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from two independent replicates that were concatenated and analyzed together.





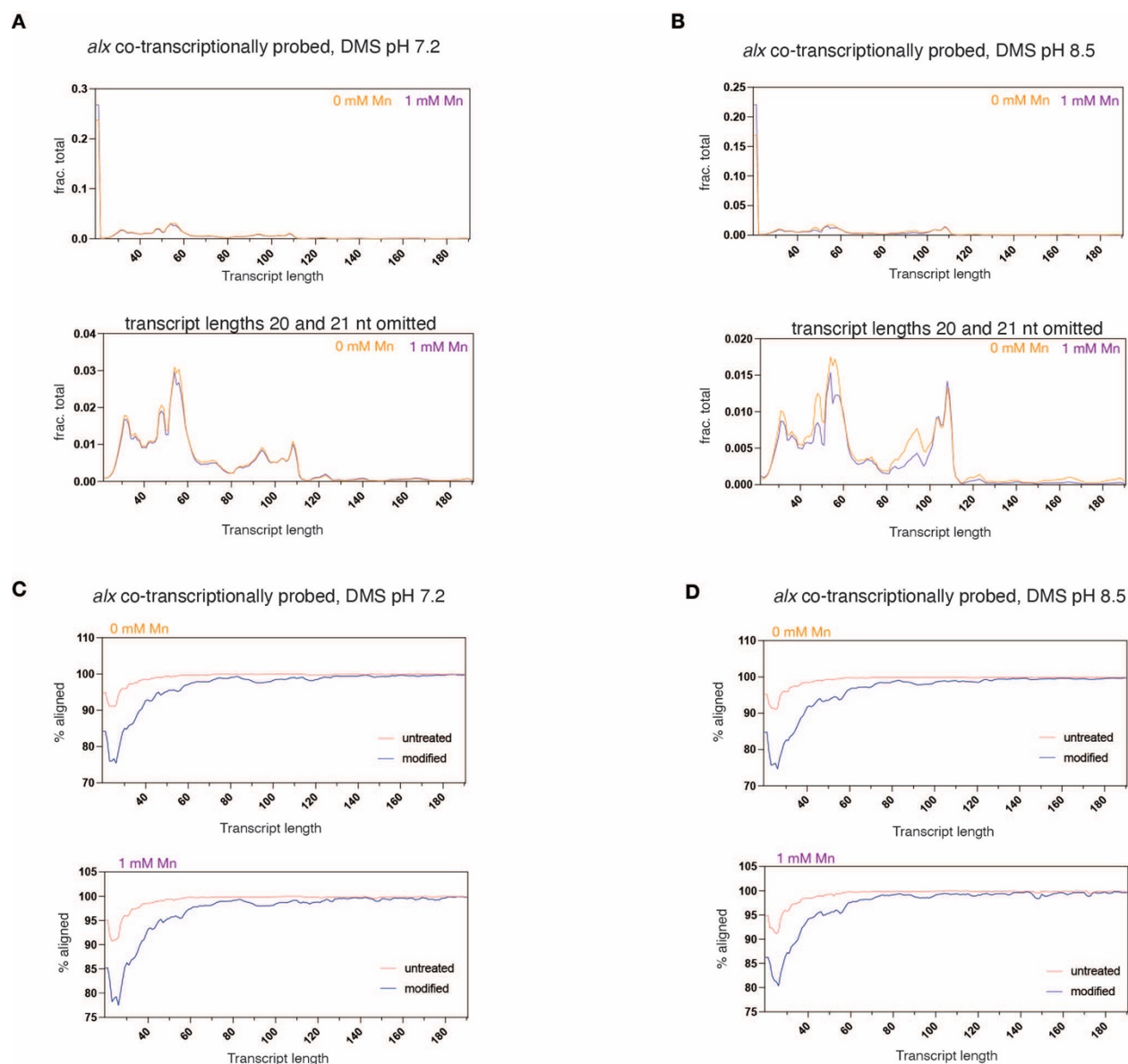
Supplemental Figure 4.19 Dropoff in sequencing depth following second *alx* intrinsic termination site and co-transcriptional SHAPE probing of full-length *alx* to capture RBS reactivities.

A. TECprobe-VL DMS reactivity matrix for *alx* co-transcriptionally folded intermediates at pH 8.5. Intrinsic transcription termination sites are annotated adjacent to the heatmap, showing a clear drop in reactivity data quality following the second intrinsic termination site. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. Data are from seven independent replicates that were concatenated and analyzed together. **B.** Plot showing BzCN reactivities of *alx* RBS and start codon nt in the full-length, co-transcriptionally folded riboswitch at pH 8.5 $\pm$ 1 mM Mn, corresponding to transcript length 232 nt. Two independent replicates were performed for the $\pm$ Mn conditions.



Supplemental Figure 4.20 Structure probing of equilibrium-refolded *a/x* full-length riboswitch RNA.

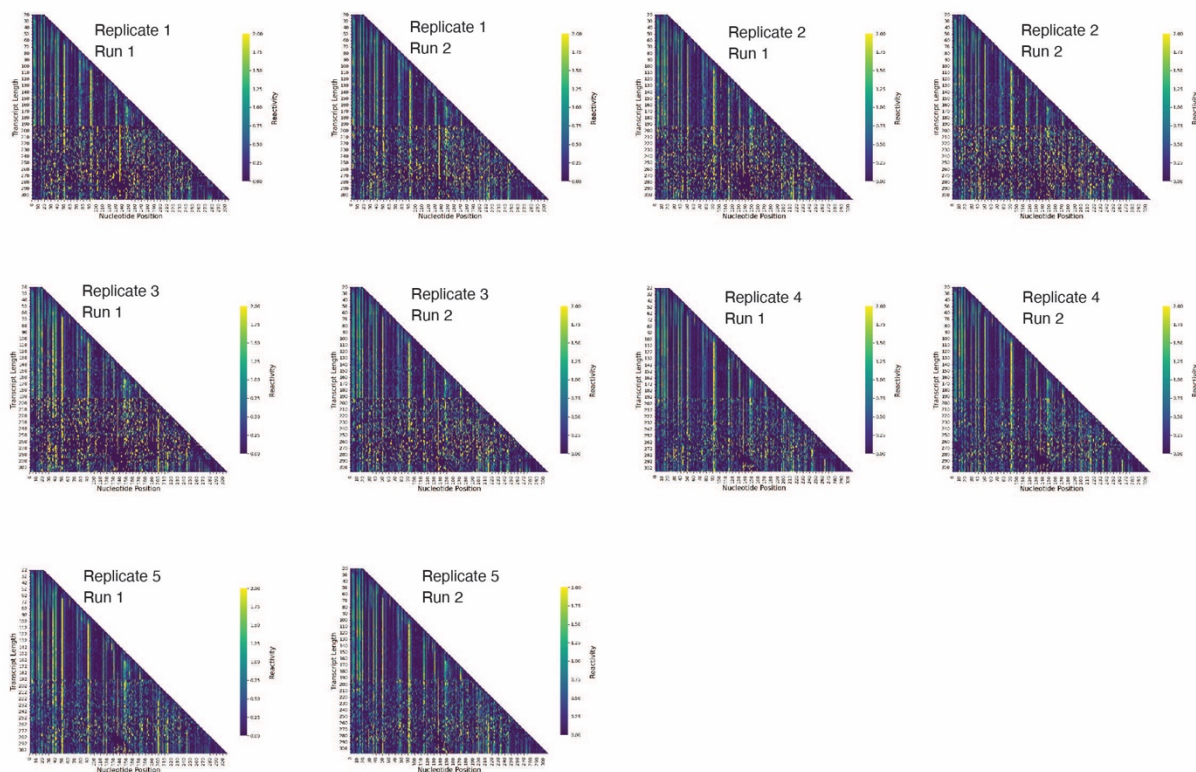
A. Secondary structure of equilibrium-refolded *a/x* full-length RNA modeled based on reactivity with 1M7 at pH 7.2 -Mn. **B.** Individual nucleotide reactivity differences when modified with 1M7 to compare +Mn vs. -Mn, measured at both pH 7.2 and 8.5. **C.** Individual nucleotide reactivity differences at neutral pH (7.2) vs. alkaline pH (8.5), measured both in the presence and absence of Mn. **D.** Individual nucleotide reactivity differences when modified with DMS to compare +Mn vs. -Mn, measured at both pH 7.2 and 8.5.



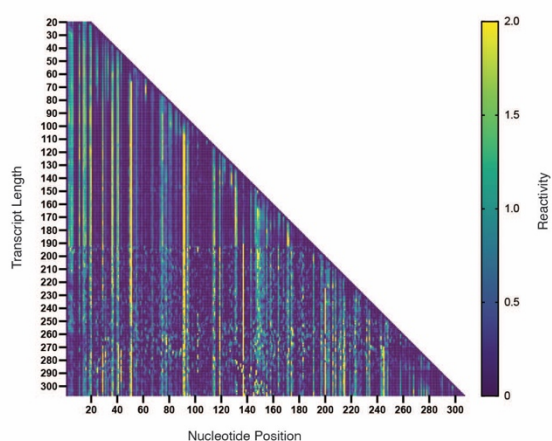
Supplemental Figure 4.21 Aligned reads and alignment rates for co-transcriptional DMS probing of the *alx* riboswitch.

A-B. (Top panel) Plot showing fraction of aligned reads mapped to each transcript length. (Bottom panel) Plot showing fraction of aligned reads mapped to each transcript length, with transcript lengths 20 and 21 nt omitted to better visualize the fraction of aligned reads for transcript lengths 22-190 nt. **C-D.** Percentage of reads retained after splitting those that aligned to each intermediate transcript length.

A *alx* co-transcriptionally probed, DMS pH 7.2 individual replicate heat maps



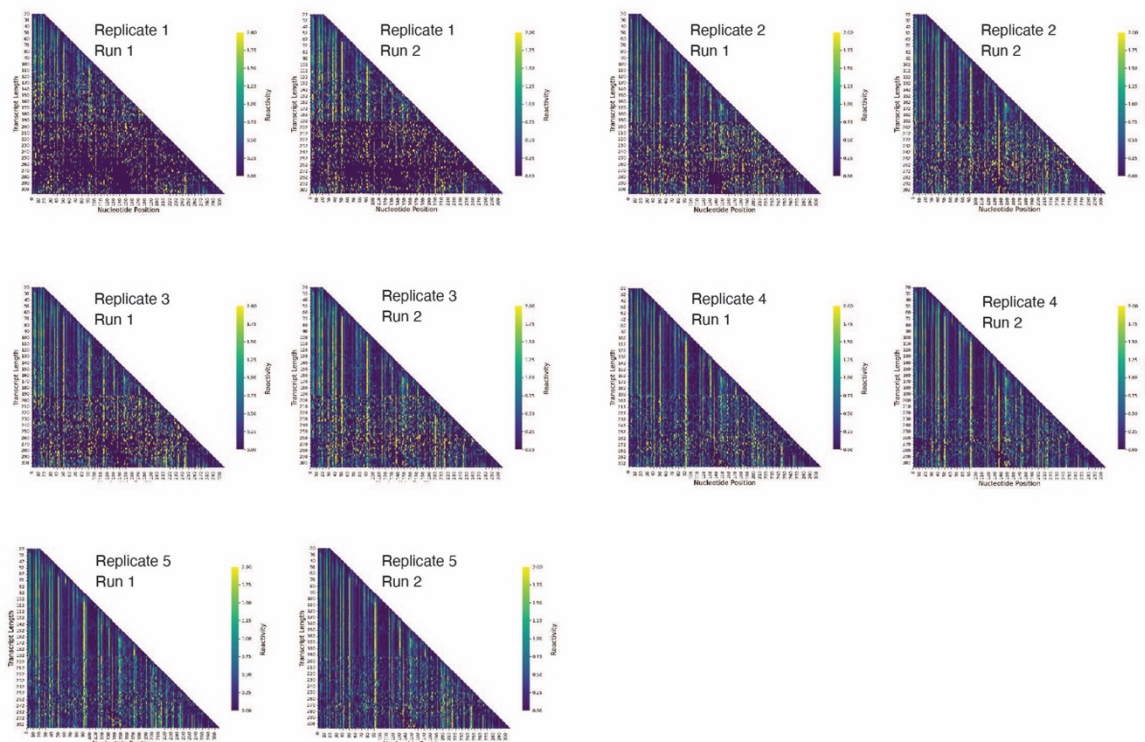
B *alx* co-transcriptionally probed, DMS pH 7.2 full concatenated heat map



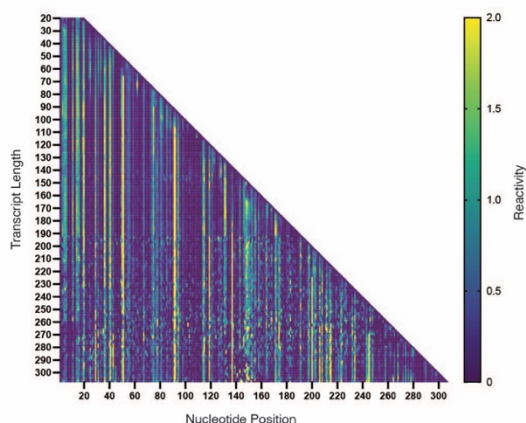
Supplemental Figure 4.22 Reactivity matrices for individual and concatenated replicates for co-transcriptional DMS probing of the *alx* riboswitch at pH 7.2.

A. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates from five independent replicates at pH 7.2. “Run 1” and “Run 2” denote sequencing runs on independent PCR amplifications of the same cDNA library. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. **B.** TECprobe-VL DMS reactivity matrix for *alx* co-transcriptionally folded intermediates at pH 7.2. Reads from the five independent replicates shown in panel A, which were concatenated and analyzed together.

A *alx* co-transcriptionally probed, DMS pH 7.2 1mM Mn individual heat maps



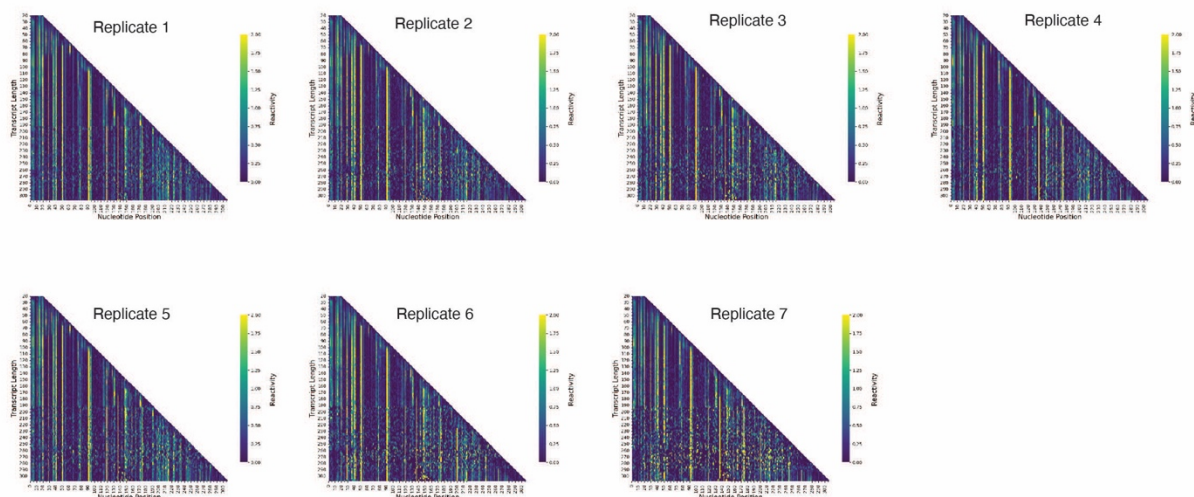
B *alx* co-transcriptionally probed, DMS pH 7.2 1 mM Mn full concatenated heat map



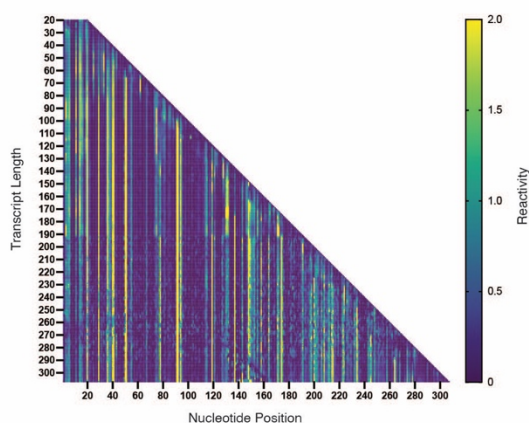
Supplemental Figure 4.23 Reactivity matrices for individual and concatenated replicates for co-transcriptional DMS probing of the *alx* riboswitch at pH 7.2 +Mn.

A. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates from five independent replicates at pH 7.2 + 1 mM Mn. “Run 1” and “Run 2” denote sequencing runs on independent PCR amplifications of the same cDNA library. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. **B.** TECprobe-VL DMS reactivity matrix for *alx* co-transcriptionally folded intermediates at pH 7.2 + 1 mM Mn. Reads from the five independent replicates shown in panel A, which were concatenated and analyzed together.

A *alx* co-transcriptionally probed, DMS pH 8.5 individual heat maps



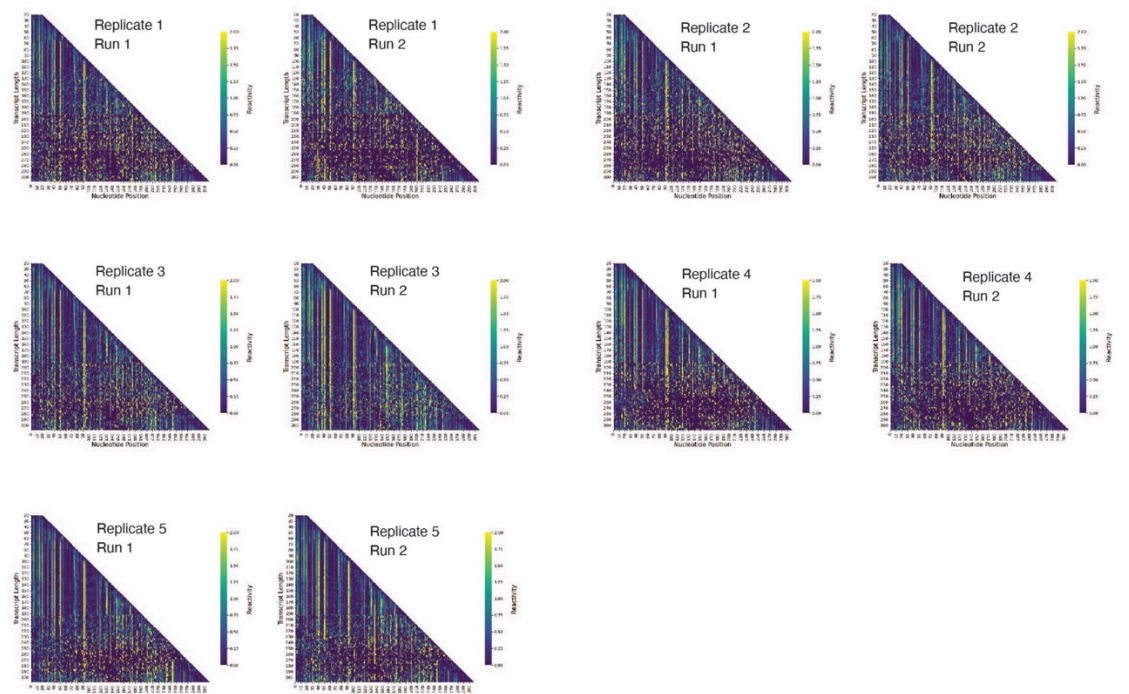
B *alx* co-transcriptionally probed, DMS pH 8.5 full concatenated heat map



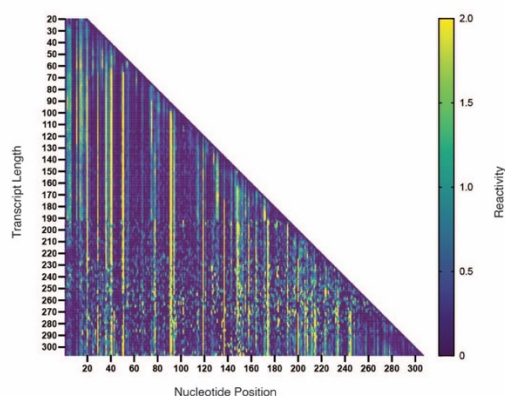
Supplemental Figure 4.24 Reactivity matrices for individual and concatenated replicates for co-transcriptional DMS probing of the *alx* riboswitch at pH 8.5.

A. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates from seven independent replicates at pH 8.5. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. **B.** TECprobe-VL DMS reactivity matrix for *alx* co-transcriptionally folded intermediates at pH 8.5. Reads from the seven independent replicates shown in panel A, which were concatenated and analyzed together.

A *alx* co-transcriptionally probed, DMS pH 8.5 1 mM Mn individual heat maps

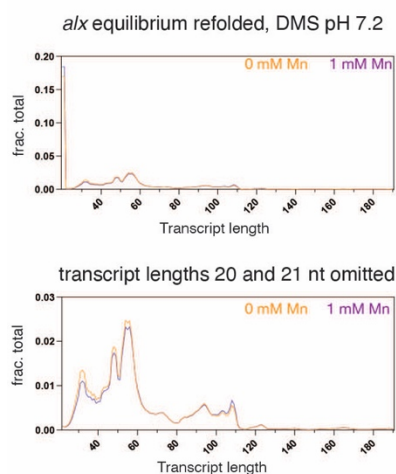
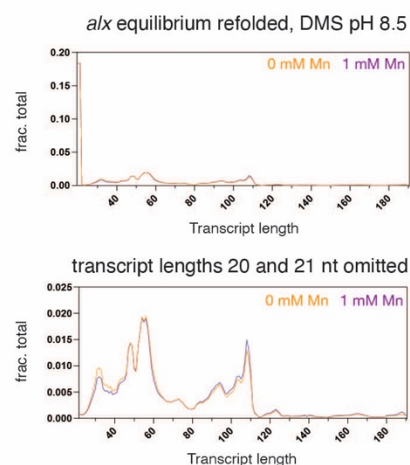
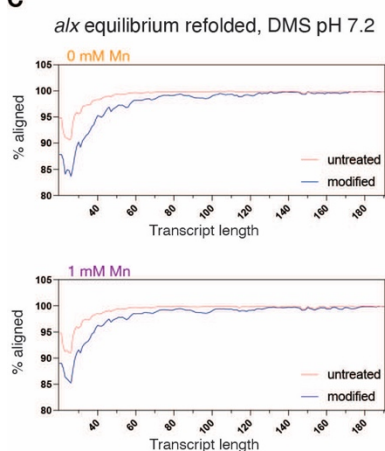
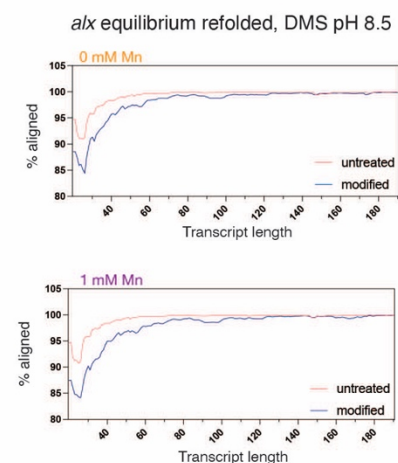


B *alx* co-transcriptionally probed, DMS pH 8.5 1 mM Mn full concatenated heat map



Supplemental Figure 4.25 Reactivity matrices for individual and concatenated replicates for co-transcriptional DMS probing of the *alx* riboswitch at pH 8.5 +Mn.

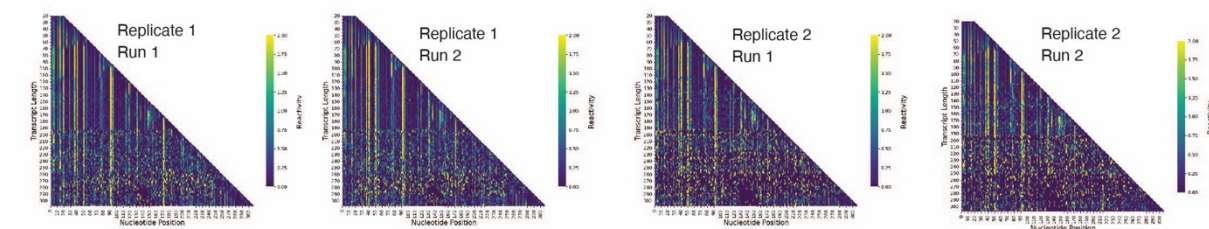
A. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates from five independent replicates at pH 8.5 + 1 mM Mn. “Run 1” and “Run 2” denote sequencing runs on independent PCR amplifications of the same cDNA library. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. **B.** TECprobe-VL DMS reactivity matrix for *alx* co-transcriptionally folded intermediates at pH 8.5 + 1 mM Mn. Reads from the five independent replicates shown in panel A, which were concatenated and analyzed together.

A**B****C****D**

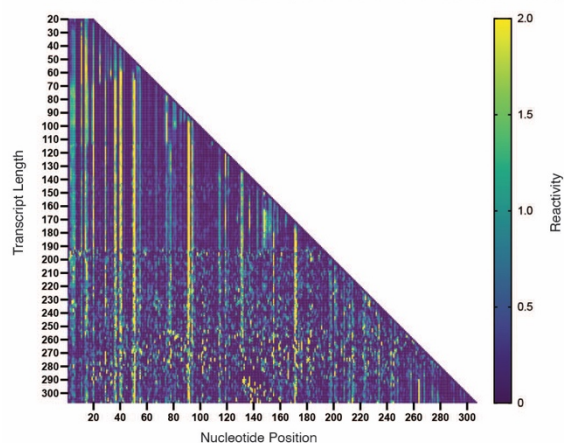
Supplemental Figure 4.26 Aligned reads and alignment rates for DMS probing of *alx* equilibrium refolded intermediates.

A-B. (Top panel) Plot showing fraction of aligned reads mapped to each transcript length. (Bottom panel) Plot showing fraction of aligned reads mapped to each transcript length, with transcript lengths 20 and 21 nt omitted to better visualize the fraction of aligned reads for transcript lengths 22-190 nt. **C-D.** Percentage of reads retained after splitting those that aligned to each intermediate transcript length.

A *alx* equilibrium refolded, DMS pH 7.2 individual heat maps



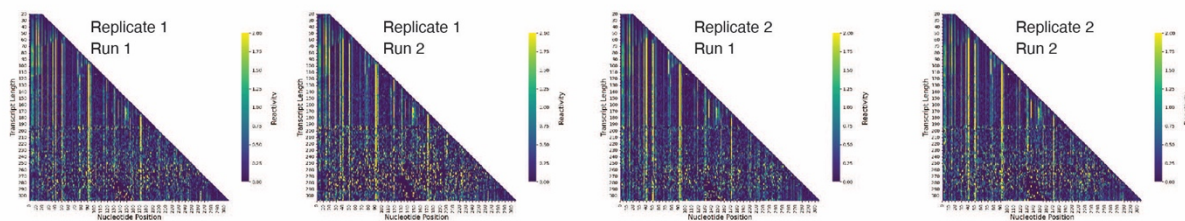
B *alx* equilibrium refolded, DMS pH 7.2 full concatenated heat map



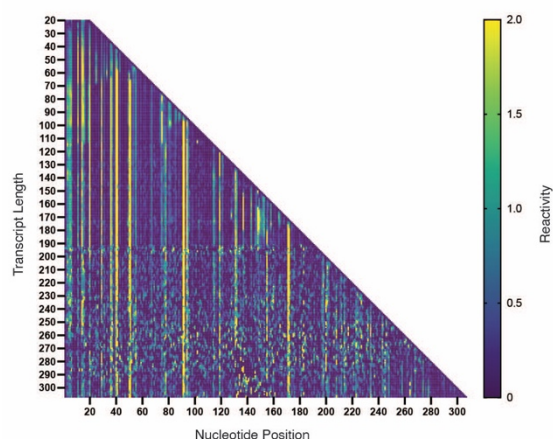
Supplemental Figure 4.27 Reactivity matrices for individual and concatenated replicates for equilibrium DMS probing of the *alx* riboswitch at pH 7.2.

A. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates from two independent replicates at pH 7.2. “Run 1” and “Run 2” denote sequencing runs on independent PCR amplifications of the same cDNA library. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. **B.** TECprobe-VL DMS reactivity matrix for *alx* equilibrium refolded intermediates at pH 7.2. Reads from the two independent replicates shown in panel A, which were concatenated and analyzed together.

A *alx* equilibrium refolded, DMS pH 7.2 1mM Mn individual heat maps



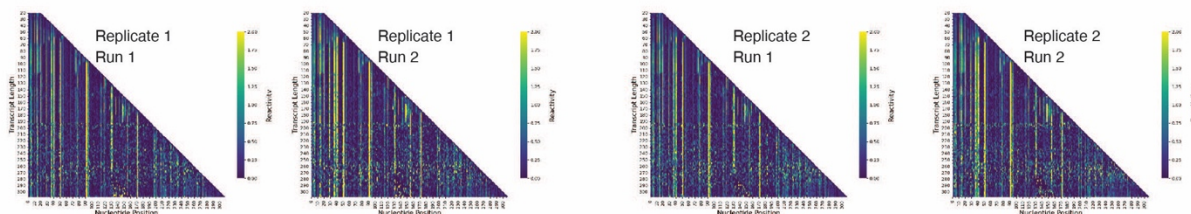
B *alx* equilibrium refolded, DMS pH 7.2 1mM Mn full concatenated heat map



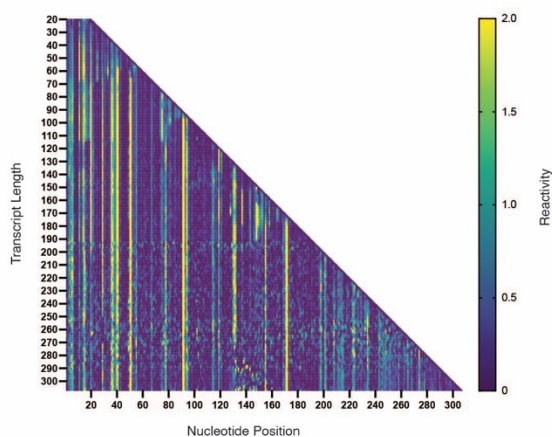
Supplemental Figure 4.28 Reactivity matrices for individual and concatenated replicates for equilibrium DMS probing of the *alx* riboswitch at pH 7.2 +Mn.

A. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates from two independent replicates at pH 7.2 + 1 mM Mn. “Run 1” and “Run 2” denote sequencing runs on independent PCR amplifications of the same cDNA library. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. **B.** TECprobe-VL DMS reactivity matrix for *alx* equilibrium refolded intermediates at pH 7.2 + 1 mM Mn. Reads from the two independent replicates shown in panel A, which were concatenated and analyzed together.

A *alx* equilibrium refolded, DMS pH 8.5 individual heat maps



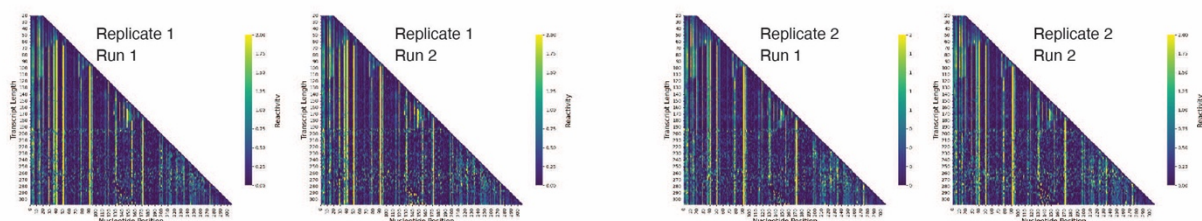
B *alx* equilibrium refolded, DMS pH 8.5 full concatenated heat map



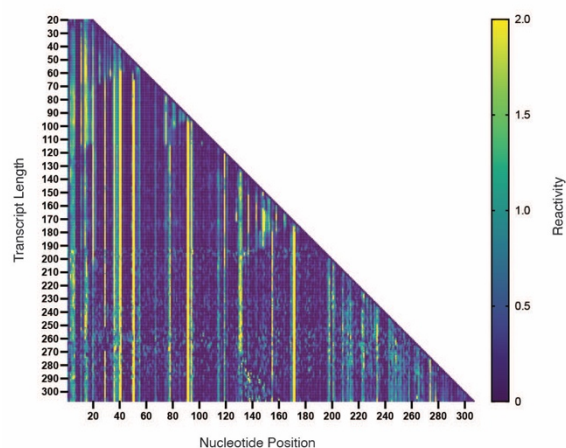
Supplemental Figure 4.29 Reactivity matrices for individual and concatenated replicates for equilibrium DMS probing of the *alx* riboswitch at pH 8.5.

A. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates from two independent replicates at pH 8.5. “Run 1” and “Run 2” denote sequencing runs on independent PCR amplifications of the same cDNA library. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. **B.** TECprobe-VL DMS reactivity matrix for *alx* equilibrium refolded intermediates at pH 8.5. Reads from the two independent replicates shown in panel A, which were concatenated and analyzed together.

A *alx* equilibrium refolded, DMS pH 8.5 1 mM Mn individual heat maps

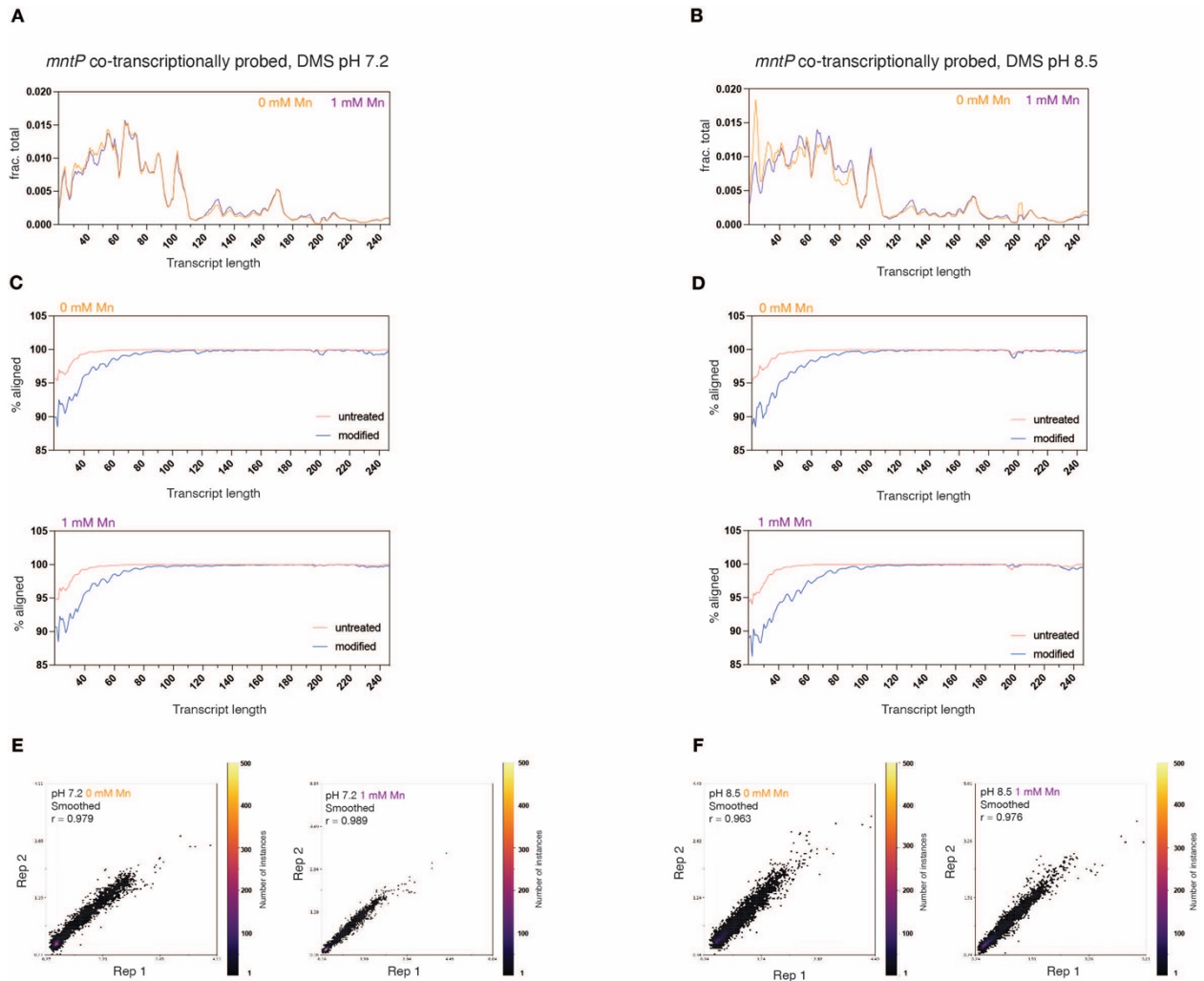


B *alx* equilibrium refolded, DMS pH 8.5 1 mM Mn full concatenated heat map



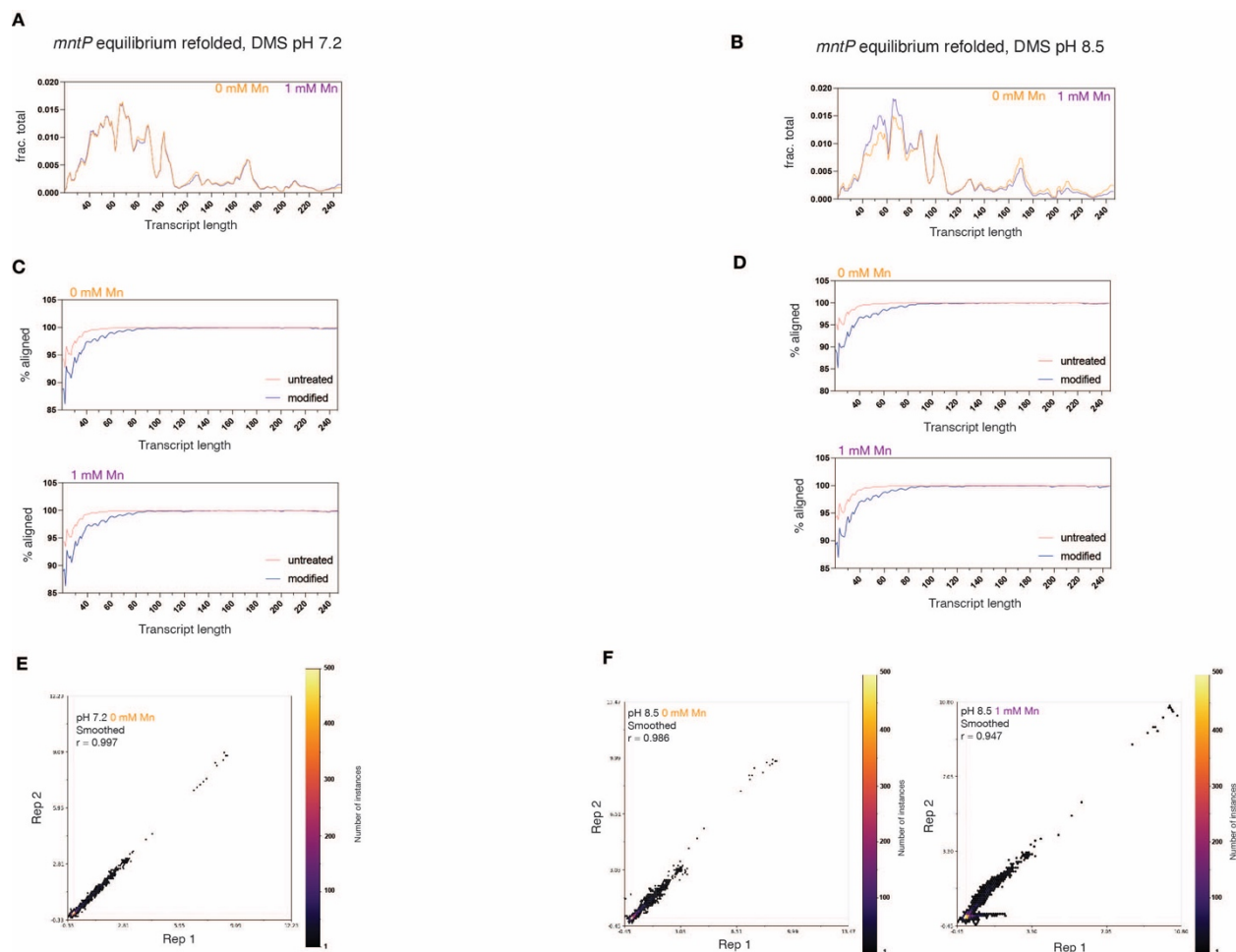
Supplemental Figure 4.30 Reactivity matrices for individual and concatenated replicates for equilibrium DMS probing of the *alx* riboswitch at pH 8.5 +Mn.

A. TECprobe-VL DMS reactivity matrices for *alx* folding intermediates from two independent replicates at pH 8.5 + 1 mM Mn. “Run 1” and “Run 2” denote sequencing runs on independent PCR amplifications of the same cDNA library. Reactivities shown were normalized with a single, whole-dataset calculated normalization factor. **B.** TECprobe-VL DMS reactivity matrix for *alx* equilibrium refolded intermediates at pH 8.5 + 1 mM Mn. Reads from the two independent replicates shown in panel A, which were concatenated and analyzed together.



Supplemental Figure 4.31 Aligned reads and alignment rates for co-transcriptional DMS probing of the *mntP* riboswitch.

A-B. Plot showing fraction of aligned reads mapped to each transcript length. **C-D.** Percentage of reads retained after splitting those that aligned to each intermediate transcript length. **E-F.** Hexbin plots comparing the DMS reactivity of replicates.



Supplemental Figure 4.32 Aligned reads and alignment rates for equilibrium DMS probing of the *mntP* riboswitch.

A-B. Plot showing fraction of aligned reads mapped to each transcript length. **C-D.** Percentage of reads retained after splitting those that aligned to each intermediate transcript length. **E-F.** Hexbin plots comparing the DMS reactivity of replicates. No hexbin plot is shown for the pH 7.2 + 1 mM Mn condition because only one replicate was analyzed for this condition.

Supplemental Notes

Supplementary Note 1

pH effect on the refolded alx aptamer structure. Comparison of the nucleotide (nt) reactivities as a function of refolding pH revealed clear flexibility differences in L3 (Fig. 4.2C). The strictly conserved adenosine (A37) that coordinates Mn is less accessible for modification by both DMS and 1M7 at alkaline pH, despite the absence of Mn (Fig. 4.2C and SI Fig. 4.1E). Similarly, most L3 nt exhibit lower reactivities at alkaline vs neutral pH when modified with 1M7; however, the C36 and U38 that neighbor A37 are both more reactive at alkaline pH (Fig. 4.2C). Also, the L1 conserved adenosine (A11) that forms the cross-helix A-minor interaction becomes less reactive towards 1M7 at alkaline pH -Mn (Fig. 4.2C). Interestingly, this A11 pH-dependent reactivity difference is abolished when RNA is modified with DMS (SI Fig. 4.1E). Together, these data indicate that alkaline pH alters the reactivity of key nucleotides in the *alx* metal-sensing core, which could be due to a local pH effect on L1/L3 structure or a global pH-dependent trend in aptamer dynamics.

Mn effect on the refolded alx aptamer structure. Reactivity analysis of 1M7-modified RNA revealed that L3 generally becomes less flexible +Mn, consistent with Mn binding, except for U38, which remains more flexible +Mn regardless of pH (SI Fig. 4.1D). U39 is also more reactive towards 1M7 +Mn at neutral pH (SI Fig. 4.1D). For DMS-modified aptamer RNA, most L3 nt do not show significant Mn-dependent reactivity differences; however, consistent with 1M7 reactivity analysis, U38 and U39 are both more flexible at neutral pH +Mn (Fig. 4.1E). The crystal structure of a Cd²⁺-bound *L. lactis* Mn-sensing aptamer mutated to have the *alx* L3 sequence shows that metal ion binding flips U38 and U39 out of the nucleobase stack formed by the A-minor motif, explaining the observed reactivity increase in reactivity when Mn is present(1).

Consistent with the role of A37 in directly coordinating Mn, no pH-dependent change in reactivity is observed +Mn when RNA is modified with either 1M7 or DMS (SI Fig. 4.1D-E). Further, in the presence of Mn, the pH-dependent change in reactivity towards 1M7 for A11, which forms the cross-helix A-minor interaction, is abolished (Fig. 4.2C). This finding agrees with published smFRET studies on the related *X. oryzae* Mn-sensing aptamer, which demonstrated that Mn increases docking of the two helical legs(2). Since A11 is a key player in this docking via the A-minor interaction, it is expected to be more frequently buried in the metal-sensing core, and thus less reactive towards a chemical probe +Mn.

Besides A11, L1 nt in 1M7-modified RNA are either more reactive or show minor reactivity differences +Mn at neutral pH; however, upon shifting to alkaline pH, L1 nt become less reactive +Mn (SI Fig. 4.1D). When aptamer RNA was DMS-probed at neutral pH, most L1 nucleotides are less reactive +Mn; however, upon a shift to alkaline pH, Mn-dependent reactivity differences are lessened, with G10 and G120 showing increased reactivity (SI Fig. 4.1E). Taken together, these data show that Mn alone reduces the flexibility of critical Mn-sensing nucleotides in L1 and L3 of the *alx* aptamer and are consistent with prior structural study of Mn-bound aptamers(1–3).

Supplementary Note 2

Mn-enhanced RNA polymerase (RNAP) pause longevity during mntP aptamer synthesis.

As a further support for Mn interacting with RNA folding intermediates, we mapped multiple RNAP pauses during *mntP* aptamer synthesis (nt 86-90) that are enhanced by Mn (SI Fig. 4.4A-B). Similar enhancement of pauses within the *mntP* aptamer was achieved by addition of NusA in the absence of Mn (SI Fig. 4.4B), pointing to RNA hairpin stabilization of paused RNAP conformation at these sites. Thus, these pauses are likely prolonged via allosteric effect of Mn-RNA on the RNAP conformation akin to the allosteric effects of nascent RNA hairpins(4, 5). The presence of these hairpin-stabilized pauses is further supported by RNA structure models constrained by DMS reactivities (stabilized by the P3.1 hairpin based on pause locations, SI Fig. 4.4C). All pause sites

described above were mapped by scaffold-based transcription assays, which provided single-nt precision for mapping pause locations (SI Fig. 4.5A-C). Taken together, these data suggest that the hairpin-stabilized pauses during aptamer synthesis mediate sampling of divalents in the environment by the *mntP* RNA folding intermediates. Recent smFRET study on the *L. lactis* Mn-sensing riboswitch also identified a pause site enhanced either in the presence of Mn or NusA, with the pause in this case positioned at the end of aptamer folding to stabilize P1.1 switch helix(6).

Supplementary Note 3

Some but not all co-transcriptional folding features are preserved in unfolded-refolded RNA intermediates. To identify the intermediates whose folding is truly controlled by the transcription kinetics vs. thermodynamic stability of the folds, we applied a modified TECprobe-VL workflow wherein intermediate RNA transcripts are extracted from the stalled transcription elongation complexes and equilibrium refolded before DMS probing (SI Fig. 4.6A). As with co-transcriptionally folded intermediates, we observed transient, low- T_m stem-loops that form and resolve once the full aptamer sequence is present (SI Fig. 4.6B). We observed similar Mn-dependent differences in reactivity in equilibrium refolded as the co-transcriptionally folded intermediates for A48, U54, and A103 (SI Fig. 4.6C). Despite the L3 being single-stranded throughout the aptamer folding in both the equilibrium refolded and co-transcriptional intermediates, the reactivity of A48 shows a marked increase ~10 nt before the aptamer folds (SI Fig. 4.6C). This suggests that once the first 100 nt of the aptamer are available for base-pairing, a structural rearrangement occurs to reposition the L3, resulting in increased reactivity of A48 overall. A key folding transition that likely prompts this structural rearrangement is the pairing of P1.2 in the right leg, where nt 15-18 at the 5' end pair with nt 96-99. Once P1.2 folds, the DMS reactivity profile for transcript length 100 nt largely resembles that of the folded aptamer, with only L1 and P1.1 absent due to lack of their pairing partners (SI Fig. 4.6D-E). Thus, formation of P1.2

initiates folding of the right leg, which finishes folding once the 5' end of the aptamer transcript (nt 2-18) pairs with the nt that emerge later in transcription (nt 96-110, Fig. 4.3B). We hypothesize that folding of P1.2 positions nt 2-14 in closer proximity to the partially folded right leg to facilitate pairing of nt 2-14 with nt 100-110 once they emerge from RNAP, permitting efficient completion of the aptamer and metal-sensing core folding, and thus a rapid response to ligand. Further, repositioning of nt 1-14 likely brings these nucleotides into an even closer proximity to L3, resulting in nucleation of the cross-helix A-minor with A11 stacked atop A48 and explaining the Mn-dependent changes in A48 reactivity at this point.

For U54, Mn-dependent differences in reactivity emerge before the transcript is 100 nt, indicating that some positions in the L3 loop can begin sampling for metal even before the right leg in the aptamer begins folding (SI Fig. 4.6C). Interestingly, the magnitude of the U54 reactivity difference $\pm$ Mn is higher for the early equilibrium refolded intermediates (transcript lengths ~60-100 nt) compared to the co-transcriptionally folded intermediates (SI Fig. 4.6C). However, past transcript length ~100 nt, the magnitude of the U54 reactivity difference $\pm$ Mn is higher for the co-transcriptionally folded intermediates, with U54 being more reactive in the absence of ligand (SI Fig. 4.6C). This indicates that the L3 is more flexible in co-transcriptionally folded intermediates once the metal-sensing core begins folding, resulting in higher reactivity at U54. Additionally, past transcript lengths ~110-120 nt, A103 demonstrates higher reactivity $-$ Mn, with the magnitude of the reactivity difference larger for the co-transcriptionally folded intermediates (SI Fig. 4.6C).

Supplementary Note 4

alx L3 nt show distinct Mn-dependent reactivity trends in equilibrium refolded intermediates. Similar to *mntP*, we observed Mn-dependent reactivity differences for equilibrium refolded *alx* intermediates (SI Fig. 4.8A). First, U35 reactivity trends higher $+$ Mn for equilibrium refolded intermediates starting at transcript length ~50 nt. Curiously, the magnitude of the reactivity difference $\pm$ Mn is greater for the early equilibrium refolded intermediates (transcript

lengths ~60-115 nt, SI Fig. 4.8B) than for later intermediates, indicating a potential structural change (e.g., a loss of a tertiary RNA contact) induced by equilibrium refolding that affects L3 dynamics in these early intermediates. Additionally, A37 in the *alx* aptamer demonstrates higher reactivity $-Mn$ for equilibrium refolded intermediates, mimicking the trend observed for co-transcriptionally folded intermediates. However, the magnitude of the Mn-dependent reactivity difference for A37 is greater for equilibrium refolded intermediates across transcript lengths ~60-115 nt and ~135-155 nt (SI Fig. 4.8B). Lastly, C41 demonstrates higher reactivity $+Mn$ for equilibrium refolded intermediates, with a larger Mn-dependent reactivity difference for equilibrium refolded intermediates spanning transcript lengths ~70 nt-110 nt (SI Fig. 4.8B). Taken together, these data indicate a potential structural change induced by equilibrium refolding that affects L3 dynamics, resulting in differing reactivity patterns compared to co-transcriptionally folded intermediates. Further, these data provide additional pieces of evidence that aptamer intermediates undergo Mn-dependent structural changes in L3.

Supplementary Note 5

*pH-dependent reactivity trends in *alx* and *mntP* L3 nt in equilibrium refolded intermediates.*

We next assessed pH differences in L3 nucleotide reactivities for *alx* and *mntP* equilibrium refolded aptamer intermediates, after verifying that the global folding transitions in the aptamers are unaffected by pH (SI Fig. 4.10A-B). First, *alx* A37 shows a clear Mn-dependent difference in reactivity at both pH from transcript lengths ~60-115 nt, with lower reactivity in general $+Mn$ (SI Fig. 4.10C). This trend is different from that seen for A37 in co-transcriptionally folded intermediates, where the lowest reactivity overall was observed for the pH 8.5 $+Mn$ condition throughout aptamer folding (Fig. 4.5B). Additionally, positions C36, A40 and C41 in *alx* equilibrium refolded intermediates do not show the same increased reactivity at pH 8.5 $\pm Mn$ as described for co-transcriptionally folded intermediates (SI Fig. 4.9C and SI Fig. 4.10C). These data support that

an alkaline pH shift uniquely affects the structure of co-transcriptionally folded *alx* aptamer intermediates.

The *mntP* L3 also demonstrates different reactivity trends in equilibrium refolded aptamer intermediates. A48 shows no clear pH difference in reactivity (SI Fig. 4.10D); this is in stark contrast to co-transcriptionally folded aptamer intermediates, where A48 is more reactive at alkaline pH (Fig. 4.5C). This result indicates that equilibrium refolding of *mntP* aptamer intermediates alters the orientation of A48, the A11 stacking partner, such that pH no longer alters the accessibility of its Watson-Crick face for DMS methylation. The absolutely conserved adenosine in *mntP* L3 (A50) is more reactive at neutral pH $\pm$ Mn for the duration of aptamer folding (SI Fig. 4.10D), which was not observed for co-transcriptionally folded intermediates (Fig. 4.5C). These data support that equilibrium refolding of *mntP* aptamer intermediates repositions the L3 loop and changes how pH affects L3 structure.

pH-dependent reactivity trends in alx and mntP L1 nt in co-transcriptionally folded vs. equilibrium refolded intermediates. Beyond L3, both the *alx* and *mntP* aptamers also exhibit curious pH-dependent reactivity differences in the absolutely conserved L1. A101 in the *mntP* aptamer is more reactive at alkaline pH whether or not Mn is present (SI Fig. 4.11A). A103 in *mntP* is generally more reactive $-$ Mn at both pH, with the absolute reactivity values greater at pH 7.2 $\pm$ Mn (SI Fig. 4.11A). At the analogous positions in the *alx* aptamer (A119 and A121) we observed different reactivity trends. In the absence of Mn, A119 in *alx* is more reactive at alkaline pH during aptamer folding. In the presence of Mn, however, there is a clear pH-dependent divergence in the *alx* A119 reactivity at the end of aptamer folding, with a peak in reactivity observed at pH 7.2 and a drop in reactivity at pH 8.5 (SI Fig. 4.11A). A121 exhibited similar trends to those described above for A119, with a stark divergence in reactivity between the pH 7.2 $+$ Mn

and pH 8.5 +Mn conditions towards the end of aptamer folding (SI Fig. 4.11A). These data solidify that a shift to alkaline pH exerts distinct structural effects on the *alx* and *mntP* aptamers.

Lastly, we assessed single nucleotide reactivity trajectories for *alx* and *mntP* L1 nucleotides in equilibrium refolded intermediates and compared them to their co-transcriptionally folded counterparts. For *mntP* A101, there is no clear pH-dependent reactivity trend (SI Fig. 4.11B), which differs from co-transcriptionally folded intermediates where A101 was more reactive at pH 8.5 in general (SI Fig. 4.11A). A119 and A121 in the *alx* L1 showed no clear pH trends for equilibrium refolded aptamer intermediates. A119 reactivity in general was higher towards the end of aptamer folding in equilibrium refolded transcripts for all conditions tested compared to co-transcriptionally folded transcripts (SI Figs 4.11A-B), indicating that equilibrium refolding for *alx* aptamer intermediates results in a more dynamic L1. Taken together, these data support that equilibrium refolding alters the structure and dynamics of the metal-sensing core.

Supplementary Note 6

pH- and Mn-dependent changes in mntP aptamer structure during co-transcriptional expression platform folding. We observed similar reactivity trends in *mntP* expression platform folding at pH 7.2 as pH 8.5 (SI Fig. 4.12B-C). For example, at pH 7.2, we also observed the transient spike in A11 reactivity at transcript length ~205 nt; however, both the amplitude of this spike and the difference $\pm$ Mn is smaller compared to pH 8.5 (SI Fig. 4.12D). Other curious ligand- and pH-dependent trends stood out for *mntP* nucleotides in L3, L1, and the switch helix P1.1 during expression platform folding. A48 in L3 drops in reactivity at transcript length ~200 nt for all conditions followed by an increase in reactivity at pH 7.2 +Mn and pH 8.5 +Mn compared to the respective –Mn conditions (SI Fig. 4.12D). This indicates that a global, Mn-dependent structural rearrangement occurs in the aptamer at this point, consistent with the reactivity trend described for A11. Since A11 and A48 are stacking partners (Fig. 4.6C), it makes sense that their reactivity changes are coupled upon disruption of the cross-helix A-minor interaction. A50 in L3, which

coordinates Mn, also exhibits ligand-dependent reactivity differences for the duration of expression platform folding (SI Fig. 4.12D).

L1 nucleotides besides A11 undergo ligand-dependent structural changes at the riboswitching point. Similar to A11, A14 reactivity spikes towards the end of expression platform folding –Mn (SI Fig. 4.12D); however, the increase in reactivity occurs ~5-10 nt before the spike in A11 reactivity, indicating that the A14-G100 base-pair in L1 is broken before the cross-helix A-minor is disrupted. A101 in L1 remains generally more reactive at pH 8.5 ±Mn vs. pH 7.2 ±Mn from transcript lengths ~120-190 nt and demonstrates high reactivity in general despite its predicted base-pairing with U13 (SI Fig. 4.12D). Curiously, at transcript length ~200 nt, a drop in A101 reactivity is observed for all conditions, with the +Mn condition trending higher for the remainder of expression platform folding (SI Fig. 4.12D). When L1 is disrupted –Mn, A101 is predicted to be single-stranded vs. remaining base-paired with U13 in the Mn-stabilized aptamer. However, past transcript length ~220 nt, A101 is generally more reactive +Mn (SI Fig. 4.12D), indicating that the Watson-Crick face of A101 is transiently exposed +Mn at the end of expression platform folding, perhaps due to differences in the expression platform tertiary structure. A103 in L1 remains generally more reactive –Mn at pH 7.2 and 8.5 for the entirety of expression platform folding (SI Fig. 4.12D). At transcript length ~200 nt, A103 reactivity increases across all conditions, with the absolute reactivity values being higher –Mn (SI Fig. 4.12D). A103 is predicted to base-pair with G9 in the Mn-stabilized aptamer vs. single-stranded –Mn (Fig. 4.6C); thus, higher A103 reactivity -Mn is consistent with these predictions. Notably, A14, A101, and A103 are reactive towards DMS in general across all conditions tested, indicating that L1 transiently melts during expression platform folding in a pH- and Mn-dependent manner.

Lastly, we analyzed co-transcriptional Mn- and pH-dependent reactivity differences in the P1.1 switch helix, the key structural riboswitching element. C2 and A3 at the base of the P1.1 switch helix undergo a steady increase in reactivity from transcript lengths ~180-195 nt, followed

by a reactivity drop ~200 nt and subsequent increase across all conditions tested (SI Fig. 4.12D). There is no clear Mn-dependent trend observed for C2 and A3, suggesting that a disrupted P1.1 is transiently sampled even in the presence of Mn. C104 spikes in reactivity at transcript length ~195 nt and the amplitude of this spike is higher –Mn (SI Fig. 4.12D), supporting that the C104-G8 base-pair in the P1.1 helix is disrupted –Mn before C104 alternatively pairs with G203 in the expression platform (Fig. 4.6C). A106 exhibits high reactivity at pH 7.2 in general for the entire duration of expression platform folding, despite its base-pairing U6 in P1.1 (SI Fig. 4.12D). Further, A106 reactivity is higher at pH 7.2 +Mn and pH 8.5 +Mn compared to their respective –Mn conditions (SI Fig. 4.12D). This suggests that in the presence of Mn, the Watson-crick face of A106 is more often accessible for DMS methylation. Lastly, A108 is generally more reactive at pH 7.2 during expression platform folding and demonstrates a clear pH-dependent spike in reactivity at transcript length ~190 nt, with the amplitude of the spike being higher at pH 7.2 ±Mn (SI Fig. 4.12D). Taken together, our observations of *mntP* L3, L1, and P1.1 reactivities during expression platform folding highlight that a global structural change occurs in the expression platform at transcript lengths ~200-205 nt, which is likely prompted by the start of strand invasion at P1.1 in the aptamer.

Supplementary Note 7

pH- and Mn-dependent reactivity changes in mntP equilibrium refolded intermediates. We expanded our analysis of equilibrium refolded *mntP* intermediates to other nucleotides in the metal-sensing core, A48 and U54. At pH 8.5, A48, the A11 stacking partner, markedly drops in reactivity ±Mn at transcript length ~200 nt, with its reactivity trending higher +Mn (SI Fig. 4.13D). While a similar trend is observed for co-transcriptionally folded intermediates, the magnitude of the Mn-dependent reactivity difference is smaller in the equilibrium refolded context. This result indicates that the *mntP* L3 in the aptamer can still perform divalent sampling even with the fully folded expression platform. A similar trend was observed at pH 7.2, illustrating that pH affects L3

structure when the expression platform is present (SI Fig. 4.13D). U54 in L3 also shows a clear difference in reactivity between co-transcriptionally folded vs. equilibrium refolded intermediates. Starting at transcript length ~115 nt, U54 is more reactive +Mn at pH 8.5 and this trend is maintained for the entirety of expression platform folding (SI Fig. 4.13D). Although a similar trend is observed for equilibrium refolded intermediates, the magnitude of the reactivity difference is decreased, indicating that equilibrium refolding impacts L3 structure and consequently its sampling for Mn. Taken together with A48, these data support that L3 structure is altered in the equilibrium refolded context.

Supplementary Note 8

We observed that, regardless of pH or presence of Mn, the T7-synthesized, equilibrium refolded *mntP* riboswitch assumes the same translationally inactive conformation, supported by DMS- and 1M7-constrained secondary structures (SI Fig. 4.14A). Nevertheless, interesting trends were extracted from the refolded full-length riboswitch chemical probing.

pH- and Mn-dependent trends in L3 reactivity. We observed interesting reactivity trends in the L3 loop for *mntP*. First, the absolutely conserved A50 shows less reactivity +Mn vs. –Mn when the RNA was refolded at pH 7.2, likely due to its direct coordination of Mn (SI Fig. 4.14B). However, there are no significant reactivity changes for A50 when RNA is refolded at pH 8.5 ±Mn, demonstrating that the flexibility of this nucleotide is limited by alkaline pH alone (SI Fig. 4.14B-C). In DMS-modified *mntP* RNA, A50 is less reactive +Mn at both pH (SI Fig. 4.14D), further supporting its direct coordination with Mn. Lastly, A48 in *mntP* L3 is less reactive towards 1M7 in general at alkaline pH but does not show a ligand-dependent reactivity difference (SI Fig. 4.14B-C). In contrast to 1M7 data, we observed that A48 was more reactive towards DMS +Mn at either pH (SI Fig. 4.14D). This observation supports that Mn binding specifically orients the Watson-Crick face of L3 such that A48 is more susceptible to DMS methylation. Together, these data

support that L3 retains its ability to sample for divalents even when the RNA is refolded into an inactive conformer with a disrupted metal-sensing core.

pH- and Mn-dependent 1M7 reactivity trends in the mntP expression platform. In the *mntP* expression platform, several pH- and ligand-dependent reactivity differences are observed when RNA is modified with 1M7. Upon a shift to alkaline pH, A117 and A118 in P5 become less reactive at both $\pm$ Mn (SI Fig. 4.14C). At alkaline pH, other P5 nucleotides G140-A144 are less reactive - Mn. Interestingly, this effect becomes overall less pronounced when Mn is introduced and C139 shows a marked increase in reactivity at pH 8.5 +Mn (SI Fig. 4.14B-C). Finally, when Mn is present, the P5 loop from G174-G183 shows a variety of reactivity difference upon a shift to alkaline pH (Fig. 4.14C). Notably, G174 and C179 are both more reactive at alkaline pH compared to neutral pH when Mn is present while U176, U180, and G181 are less reactive (SI Fig. 4.14C). RBS nucleotides are reactive with 1M7 at all conditions, but show some interesting reactivity differences. At pH 7.2, A215 is more reactive +Mn than -Mn while G216 and G217 are less reactive +Mn (SI Fig. 4.14B). In contrast, A215 is also less reactive $\pm$ Mn at pH 8.5 (SI Fig 4.14C). Taken together, these findings show that shifts to alkaline pH $\pm$ Mn affect the reactivity of many single stranded nucleotides in the *mntP* expression platform.

Supplementary Note 9

pH- and Mn-dependent reactivity changes in alx equilibrium refolded intermediates. Comparison between co-transcriptionally folded and equilibrium refolded *alx* intermediates revealed pH- and Mn-dependent reactivity trends for multiple nt in the *alx* expression platform that were unique to equilibrium refolded intermediates (see SI Figs. 4.17A-B for *alx* expression platform reactivity heatmaps, pH 7.2 and 8.5 $\pm$ Mn). A137 demonstrates a stark increase in reactivity $\pm$ Mn starting at transcript length ~150 nt, which reaches an apex at transcript length ~170 nt and then decreases until transcript length ~180 nt (SI Fig. 4.17C). The peak in reactivity

observed for A137 $\pm$ Mn in equilibrium refolded intermediates is markedly higher than its reactivity in co-transcriptionally folded intermediates at both pH 7.2 and 8.5 (SI Fig. 4.17C), pointing to a distinct structural transition induced by equilibrium refolding spanning transcript lengths $\sim$ 150-180 nt. Further, the amplitude of the A137 peak for equilibrium refolded intermediates is Mn-dependent with A137 being more reactive in the presence of Mn (SI Fig. 4.17C), indicating that Mn can influence the expression platform structure post-transcriptionally.

Supplementary Note 10

Like with *mntP*, the equilibrium refolded *alx* riboswitch assumes the same translationally inactive conformation regardless of pH or presence of Mn, as supported by DMS- and 1M7-constrained secondary structures (SI Fig. 4.20A). The following trends of interest were observed.

pH- and Mn-dependent trends in L3 reactivity. For *alx*, we saw that the strictly conserved A37 in L3 is always less reactive +Mn regardless of pH when modified with both 1M7 and DMS (SI Fig. 4.20B-D). Similar to *mntP*, this provides additional evidence that L3 can sample for divalents even when the RNA is largely refolded into the inactive conformation. When the *alx* aptamer-only was refolded, we observed that L3 becomes less flexible at pH 7.2 +Mn (SI Fig. 4.1D), with the exception of U38 and U39 that were shown to be flipped out of the A-minor stack in a published crystal structure(1). In contrast to the 1M7-probed aptamer at pH 7.2 +Mn, the full-length *alx* riboswitch RNA shows increased reactivity for U35 and C36 (SI Fig. 4.20B), likely due to a difference in L3 positioning when RNA is refolded with the expression platform.

*pH- and Mn-dependent 1M7 reactivity trends in *alx* expression platform.* In the *alx* expression platform, the loop nt G129-A132 in P5 are more reactive +Mn_at alkaline vs. neutral pH (SI Fig. 4.20C). In P6, A153 and G154 are less reactive when refolded -Mn at pH 8.5 than at

pH 7.2 (SI Fig. 4.20C). In the presence of Mn, A153 and C155 are less reactive when refolded at alkaline pH than at neutral pH while A156 is more reactive (SI Fig. 4.20C). Interestingly, U169 in the P6 loop of the inactive conformer is less reactive upon a shift to pH 8.5 $\pm$ Mn while even though U170-C172 in the same loop are more reactive upon such shift, demonstrating clear pH reactivity differences between proximal nt within the same structural element (SI Fig. 4.20C). C171 and C172 are also more reactive when Mn is present at neutral pH (SI Fig. 4.20C) We observe that RBS nucleotides are overall less reactive +Mn, where A198 and A201 are less reactive at pH 7.2 and G199 and G200 are less reactive at pH 8.5 (SI Fig. 4.20B-C). Surprisingly, G199 and G200 are more reactive at pH 8.5 -Mn than pH 7.2 -Mn, but adding Mn abolishes this reactivity difference (SI Fig. 4.20C). Taken together, these RBS reactivity differences demonstrate condition-dependent fluctuations in the flexibility of these nucleotides when the RNA has been equilibrium refolded.

Supplemental Methods

Proteins

Q5 High-Fidelity DNA Polymerase, Vent (exo-) DNA polymerase, Mth RNA Ligase (as part of the 5' DNA Adenylation kit), T4 RNA Ligase 2 truncated KQ, ET SSB, RNase H, RNase I, T4 DNA Ligase, Sall, HindIII, EcoRI, PstI, KpnI, BamHI, XbaI, DpnI, TIPP, NEBNext rRNA depletion kit, and HiScribe T7 High Yield RNA Kit were purchased from New England Biolabs. TURBO DNase, Superscript RNase Inhibitor, and SuperScript II were purchased from Thermo Fisher Scientific. SAvPhire Monomeric Streptavidin was purchased from Sigma-Aldrich. Catalog numbers for all commercial proteins are tabulated in Supplementary Table 4.1. Preparation of *E. coli* RNAP, *E. coli* σ^{70} , and *E. coli* NusA were performed as described previously(7).

Bacterial strains

The strains used in this study are derivatives of *E. coli* K-12 and are tabulated in Supplementary Table 4.4.

The $\Delta alx::Kan$ mutations were sourced from strains in the Keio collection(8) and introduced into MC4100 *E. coli* strain via P1 transduction. Chromosomally encoded T7 RNAP was introduced to this MC4100 Δalx strain using P1 transduction for the *alx* and *mntP* reporter fusions and was expressed from a *lac* promoter.

Plasmids and cloning procedures

The plasmids used in this study are tabulated in Supplementary Table 4.5.

The plasmid pMIS309 carrying the template for promoter-initiated transcription on the wild-type *mntP* riboswitch sequence (Supplementary table 4.5) was constructed by PCR-amplifying from a plasmid template, pRA57, (Supplementary table 4.5) from 486 nt upstream and 42 nt downstream of the *mntP* translation start site using oligos 0257 and 0258. The PCR product was then cloned into the BamHI and EcoRI sites of pUC19. The plasmid was assembled *via* restriction-ligation using T4 DNA Ligase (New England Biolabs).

The plasmid pMIS313 carrying the template for TECprobe-VL experiments on the *alx* riboswitch (Supplementary table 4.5) was constructed using a two-step PCR strategy. The PCRs employed oligos with 5' overhangs that introduced the PRA1 promoter and SC1 hairpin upstream of the *alx* riboswitch and additional downstream *alx* sequence. The first PCR used pMIS301 (Supplementary table 4.5) as a template and oligos 0155 and 0187. The second PCR used the purified first step PCR product as a template and oligos 0156 and 0188. The second-step PCR product was cloned into the HindIII and EcoRI sites of pUC19 *via* Gibson assembly using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs).

The plasmid pMIS315 carrying the template for TECprobe-VL experiments on the *mntP* riboswitch (Supplementary table 4.5) was constructed using a two-step PCR strategy. The PCRs employed oligos with overhangs that introduced the PRA1 promoter and SC1 hairpin upstream of the *mntP* riboswitch. The first PCR used pRA57 (Supplementary table 4.5) as a template and oligos 0258 and 0259. The second PCR used the purified first step PCR product as a template and oligos 0264 and 0265. The second step PCR product was cloned into the BamHI and EcoRI sites of pUC19 *via* restriction-ligation using T4 DNA Ligase (New England Biolabs).

The plasmid pMIS349 carrying the *alx* translational reporter with the A37T mutation in the sequence encoding the *alx* riboswitch was constructed by overlap extension PCR using published methods(9) with oligos 0229,0232,0458, and 0459.

The plasmid pMIS508 carrying the template for the *alx* aptamer under the control of a T7 promoter was constructed by cloning the template sequence (Supplementary Table 4.5) as a gene block from Integrated DNA Technologies into the BamHI site of pUC57. The plasmid was assembled *via* restriction-ligation using T4 DNA Ligase (New England Biolabs).

The plasmid pMIS509 carrying the template for the *mntP* aptamer under the control of a T7 promoter was constructed by cloning the template sequence (Supplementary Table 4.5) as a

gene block from Integrated DNA Technologies into the BamHI site of pUC57. The plasmid was assembled via restriction-ligation using T4 DNA Ligase (New England Biolabs).

The plasmid pMIS511 carrying the template for *in vitro* SHAPE/DMS-MaP of the *alx* riboswitch was constructed by cloning the template sequence (Supplementary Table 4.5) as a gene block from Integrated DNA Technologies into the PstI and XbaI sites of pUC57. The plasmid was assembled via restriction-ligation using T4 DNA Ligase (New England Biolabs).

The plasmid pMIS513 carrying the template for *in vitro* SHAPE/DMS-MaP of the *alx* aptamer was constructed by cloning the template sequence (Supplementary Table 4.5) as a gene block from Integrated DNA Technologies into the PstI and XbaI sites of pUC57. The plasmid was assembled via restriction-ligation using T4 DNA Ligase (New England Biolabs).

The plasmid pMIS518 carrying the template for *in vivo* DMS-MaP of the *alx* riboswitch was constructed by PCR amplifying the template sequence (Supplementary Table 4.5) from pMIS447 (BioBasic) using oligos 0005 and 0026 and cloning it into the XbaI and PstI sites of pUC19. The plasmid was assembled via restriction-ligation using T4 DNA Ligase (New England Biolabs).

The plasmids with *alx* and *mntP* reporter gene fusions under the control of a native *E. coli* promoter are pRA54 (*alx* translational reporter) and pRA57 (*mntP* translational reporter) and were constructed in prior work(10) and used as controls in this work (Supplementary Table 4.5). Reporter gene fusions under the control of a T7 promoter (*alx*: pMIS501 and *mntP*: pMIS517, Supplementary Table 4.5) were constructed via sequential PCR amplification. For *alx*, the first amplification generated 208 nucleotides of the *alx* 5'-UTR under control of a T7 promoter. This was amplified from pMIS303 using oligos 0202 and 0203, adding a PstI restriction site. For the second amplification, *alx-lacZ* was amplified from pRA54 for the translational fusion using oligos 0204 and RAV23, giving a 97 nucleotide overlap with the first PCR amplification. A third PCR step using 0202 and RAV23 combined the first two PCRs into the full length *alx* translational fusion under the control of a T7 promoter. The translational fusion (pMIS501) was cloned into the PstI

and BamHI sites of a single copy plasmid pMU2386. The plasmids were assembled via restriction-ligation using T4 DNA Ligase (New England Biolabs). For *mntP*, the first amplification generated 228 nucleotides of the *mntP* 5'-UTR under the control of a T7 promoter. This was amplified from pMIS310 using oligos 0202 and 0454, adding a PstI restriction site. For the second amplification, *mntP-lacZ* was amplified from pRA57 for the translational fusion using oligos 0455 and RAV126, giving a 95 nucleotide overlap with the first PCR amplification. A third PCR step using 0202 and RAV126 combined the first two PCRs into the full length *mntP* translational fusion under the control of a T7 promoter. The translational fusion (pMIS517) was cloned into the PstI and SalI sites of a single copy plasmid pMU2386. The plasmids were assembled via restriction-ligation using T4 DNA Ligase (New England Biolabs).

***In vitro* transcription template preparation**

Supplementary Table 4.3 tabulates details for the linear double-stranded DNA templates prepared for this study, including oligonucleotides used for PCR amplification and cleanup method used. Supplementary Table 4.6 provides all DNA template sequences used in this work.

Unmodified or terminally biotinylated DNA templates for promoter-initiated *in vitro* transcription were prepared by PCR with high-fidelity Q5 DNA polymerase (New England Biolabs), treated with DpnI (New England Biolabs) if a plasmid was the PCR template, and purified by either phenol extraction followed by ethanol precipitation or Monarch PCR & DNA cleanup kit (New England Biolabs).

DNA templates for TECprobe-VL were PCR amplified from a linear DNA template containing a terminal biotin modification (*alx*: Template 2, *mntP*: Template 5) using Vent (exo-) DNA polymerase (New England Biolabs). The PCR contained 1X ThermoPol Buffer (New England Biolabs), 200 μ M dNTP Solution Mix, 250 nM forward primer, 250 nM reverse primer, 0.02 nM template DNA, 0.02 U/ μ l Vent (exo-) DNA polymerase, and a concentration of biotin-11-dNTPs (AAT BioQuest and Biotium) that promoted incorporation of ~3 biotin modifications in the

transcribed region of the DNA template(11). The PCR was performed using the following thermal cycler protocol: 95 °C for 3 min [95 °C for 20 s, 58 °C for 30 s, 72 °C for 25 s] x 30 cycles, 72 °C for 5 min, hold at 12 °C. The PCR products were purified by phenol extraction followed by ethanol precipitation and quantified using the Quant-iT dsDNA Assay Kit (HS) (Invitrogen) with a BioTek Synergy H1 microplate reader (Agilent).

Linear DNA templates were prepared for transcription by T7 RNAP via PCR amplification of a plasmid using Q5 DNA Polymerase. For the *alx* riboswitch SHAPE/DMS-MaP template DNA, oligos 0026 and 0277 were used to amplify the template DNA from plasmid pMIS511 and DpnI was used to digest original plasmid DNA (Template 8). For the *alx* aptamer SHAPE/DMS-MaP template DNA, oligos 0026 and 0277 were used to amplify the template DNA from plasmid pMIS513 and DpnI was used to digest original plasmid DNA (Template 9). For the *mntP* riboswitch SHAPE/DMS-MaP template DNA, oligos 0297 and 0298 were used to amplify the template DNA from plasmid pMIS314 and DpnI was used to digest original plasmid DNA (Template 10).

Synchronized single-round *in vitro* transcription from a promoter

To monitor transcription on the *mntP* riboswitch sequence when initiated from its native promoter, we performed synchronized, single-round *in vitro* transcription assays. The template contained the *mntP* promoter (261 nt upstream of the transcription start site), the wild-type *mntP* riboswitch (nt 1-225), and nt 1-45 of *mntP* coding sequence (Template 4, Supplementary Table 4.3). To prepare the RNAP holoenzyme (holoRNAP), *E. coli* RNAP (2 μ M) and σ^{70} (4 μ M) were combined in 1x transcription buffer (20 mM Tris-Cl, pH 8.5, 150 mM KCl, 10 mM MgCl₂, and 5 mM DTT) and incubated for 30 min at 37 °C. Following holoRNAP assembly, *in vitro* transcription reactions were initiated with an ApU dinucleotide and synchronized, taking advantage of the natural lack of cytosines between positions 3 and 15 of the *mntP* riboswitch. The initiation mix for *in vitro* transcription thus lacked CTP and contained 0.1 μ M of the holoRNAP, 200 nM of the DNA template, 1 μ M ATP, 2.5 μ M CTP, 2.5 μ M GTP, 150 μ M ApU (Trilink), [α -³²P]ATP (Revvity), and

transcription buffer. After allowing initiation to proceed for 15 min at 37 °C, rifampicin (Fisher) was added to a final concentration of 10 μ g/ml, to inhibit reinitiation. Completion of this step yielded ECs stalled at the G15 position on the *mntP* riboswitch.

Next, transcription elongation reactions were performed by combining 9 μ l of ECs with 1 μ l of an NTP mix containing 1 mM each ATP, UTP, CTP, and GTP (100 μ M final, each NTP). All elongation reactions were performed at 37 °C. MnCl₂ was introduced concurrently with NTPs to a final concentration of 100 μ M to test the effect of Mn on transcription elongation and pausing. The transcription factor NusA was introduced concurrently with NTPs to a final concentration of 250 nM to test its effect on transcription elongation and pausing. After beginning elongation, the reaction was quenched at desired time points with an equal volume of 2x stop buffer (90 mM Tris-borate buffer, 8 M urea, 50 mM EDTA, and 0.02% of both xylene cyanol and bromophenol blue). The time points collected were 10 s, 30 s, 60 s, 90 s, 2 min, 3 min, and 4 min. Additionally, a “chase” reaction was performed with each NTP present at a final concentration of 500 μ M. Radiolabeled RNAs in the quenched samples were denatured for 2 min at 95 °C and separated using denaturing PAGE (6% 19:1 acrylamide:bisacrylamide, 8 M urea). Samples were run alongside a 5'-³²P radiolabeled ladder (MspI-digested pBR322 plasmid). The gel was exposed to a PhosphorImager screen, and the screen was scanned using a Typhoon PhosphorImager.

***mntP* riboswitch pause mapping using scaffold-based transcription**

To precisely map pauses in the *mntP* riboswitch, transcription elongation assays using nucleic acid scaffolds were performed. Each scaffold initiated transcription upstream of the sequence window where pauses were expected to occur.

Multiple nucleic acid scaffolds were designed to map pauses on the *mntP* riboswitch sequence towards the end of aptamer domain folding and beginning of expression platform folding. The sequence window for scaffold-based pause mapping experiments was narrowed based on results from promoter-based assays. These experiments are specifically designed to

map elemental pauses, meaning that the influence of local RNA structure on pause duration is omitted as a variable. Prior work has established that the elemental pause is an obligate precursor to longer-lived pauses stabilized by nascent RNA structure(12). Therefore, these minimalistic experiments are relevant for mapping pauses encountered by ECs when the full-length riboswitch RNA is present. A brief description of the scaffolds is presented here and are described pictorially in Supplementary Figure 4.5. The nucleotide numbering corresponds to the numbering in the full-length *mntP* riboswitch sequence. Scaffold 1 employed an 8-mer starting RNA to position RNAP upstream of nt 79-100. Scaffold 2 used a 10-mer RNA to position RNAP upstream of nt 101-113. Two separate experiments with different starting RNAs were performed on scaffold 3. The first 10-mer RNA positioned RNAP upstream of nt 114-131, and the second 10-mer RNA positioned RNAP upstream of nt 132-139. Past these defined sequence windows for pause mapping on scaffolds 1-3, the RNA emerging from RNAP could form non-native secondary structures that may have influenced pause behavior. Therefore, RNA species outside of the defined sequence windows were omitted for pause analysis.

All scaffold-based transcription elongation assays used the following protocol. First, the nucleic acid scaffold was annealed in a buffer (20 mM Tris, pH 7.2, 150 mM KCl, and 5 mM DTT) with 5 μ M RNA and 10 μ M template DNA (T-DNA). Then, ECs were reconstituted by combining RNAP with the nucleic acid scaffold in transcription buffer (20 mM Tris, pH 7.2, 150 mM KCl, and 5 mM DTT) and incubated for 15 min at 37 °C. Next, the non-template DNA (NT-DNA) was added and incubation continued for 10 min at 37 °C to complete EC assembly. The ratio of RNA:T-DNA:NT-DNA:RNAP was 1:2:5:3 corresponding to concentrations of 0.5, 1, 2.5, 1.5 μ M, respectively. Next, the ECs were diluted fivefold, and heparin was added to the ECs at a final concentration of 0.1 mg/ml to sequester unbound RNAP. ECs were incubated with heparin for 3 min at 37 °C. To radiolabel the ECs, [α -³²P]ATP (Revvity) was added to the ECs for 1 min at 37 °C, which extended a fraction of RNAs by 1 nt for visualization. To extend all ECs by 1 nt, ECs

were incubated with unlabeled ATP at a final concentration of 2 μ M for 3 min at 37 °C. Transcription elongation reactions were performed by combining the ECs with an NTP mix containing 1 mM each ATP, UTP, CTP, and GTP (100 μ M final, each NTP). $MgCl_2$ was added to a final concentration of 5 mM alongside NTP introduction. Additionally, to extend pause duration and aid with pause mapping, transcription elongation reactions were performed with three of the NTPs at a final concentration of 100 μ M and the fourth NTP limiting at a final concentration of 10 μ M. Samples were removed at various time points and quenched with an equal volume of 2x stop buffer. Chase reactions were performed with each NTP at a final concentration of 500 μ M. In parallel, RNA sequencing ladders were generated to map pause locations. To make the ladder, the ECs were prepared as described above. The ECs were combined with ATP, UTP, CTP, and GTP to 100 μ M and one 3'-dNTP (Millipore Sigma) to 1 mM. The sequencing reactions proceeded for 20 min at 37 °C and were quenched with an equal volume of 2x stop buffer. Radiolabeled RNAs from the quenched time points were heat-denatured for 2 min at 95 °C and separated using denaturing PAGE (15% 19:1 acrylamide:bisacrylamide, 8 M urea). The gel was exposed to a PhosphorImager screen, and the screen was scanned using a Typhoon PhosphorImager.

Supplemental Table 4.1 Key reagents used in this study.

Proteins and associated solutions	Source	Cat no.
Q5 High-Fidelity DNA Polymerase		
5X Q5 Reaction Buffer	New England Biolabs	M0491
5X Q5 High GC Enhancer		
Vent (exo-) DNA polymerase	New England Biolabs	M0257
10X Thermo Pol Buffer	New England Biolabs	
Mth RNA Ligase		
DNA Adenylation Buffer	New England Biolabs	E2610
1 mM ATP		
T4 RNA Ligase 2 truncated KQ		
10X T4 RNA Ligase Buffer	New England Biolabs	M0373
50% PEG 8000		
ET SSB	New England Biolabs	M2401
RNase H	New England Biolabs	M0297
RNase If	New England Biolabs	M0243
TURBO DNase	Invitrogen (Thermo Fisher Scientific)	AM2239
SuperaseIN RNase Inhibitor	Invitrogen (Thermo Fisher Scientific)	AM2694
SuperScriptII	Invitrogen (Thermo Fisher Scientific)	18064071
DTT, 100 mM		
SAvPhire Monomeric Streptavidin	Sigma-Aldrich	SAE0094
T4 DNA Ligase	New England Biolabs	M0202
DpnI	New England Biolabs	R0176
<i>E. coli</i> RNAP	Purified in-house	
σ^{70}	Purified in-house	
NusA	Purified in-house	
TIPP	New England Biolabs	M0296
Sall	New England Biolabs	R3138
HindIII	New England Biolabs	R3104
EcoRI	New England Biolabs	R3101
PstI	New England Biolabs	R3140
KpnI	New England Biolabs	R3142
BamHI	New England Biolabs	R3136
XbaI	New England Biolabs	R0145
Nucleotides	Source	Cat no.
dNTP Solution Mix	New England Biolabs	N0447S
Biotin-11-dATP	AAT BioQuest	17014
Biotin-11-dGTP	AAT BioQuest	17015
Biotin-11-dUTP	AAT BioQuest	17016
Biotin-11-dCTP	Biotium	40036
High-Purity rNTPs set	Cytiva	27-2025-01
[α - 32 P]ATP	Revvity	BLU003H250UC
ddATP, ddUTP, ddCTP, ddGTP (Sequencing G)	Millipore Sigma	3732738001
ApU dinucleotide	Trilink	0-31004
ATP	Sigma-Aldrich	A26209
UTP	Sigma-Aldrich	94370
CTP	Sigma-Aldrich	C1506
GTP	Sigma-Aldrich	G8877

Supplemental Table 4.1 Key reagents used in this study.(Continued)

Kits	Source	Cat no.
HiScribe T7 High Yield RNA Kit	New England Biolabs	E2040
Qubit ssDNA Assay Kit	Invitrogen (Thermo Fisher Scientific)	Q10212
Quant-iT dsDNA Assay Kit (HS)	Invitrogen (Thermo Fisher Scientific)	Q33120
Quant-iT RNA Assay Kit	Invitrogen (Thermo Fisher Scientific)	Q33140
Columns	Source	Cat no.
HiLoad 16/600 Superdex 200 pg (1x120 ml)	Cytiva	28989335
HisTrap HP	Cytiva	17524802
HiTrap Heparin HP	Cytiva	17040703
Other chemicals and reagents	Source	Cat no.
10 mg/ml Dynabeads MyOne Streptavidin C1 b	Invitrogen (Thermo Fisher Scientific)	65001
Agencourt AMPure XP beads	Beckman Coulter	A63881
Benzoyl Cyanide (BzCN) 98%	Sigma-Aldrich	115959
Chloroform	Fisher	C298-500
Dimethyl Sulfate (DMS) 99+%	Sigma-Aldrich	D186309
Dimethyl Sulfoxide (DMSO), Anhydrous	Invitrogen (Thermo Fisher Scientific)	D12345
Ethanol, 200 proof	Fisher	BP2818-500
Formamide	Fisher	BP227-500
GlycoBlue Coprecipitant	Invitrogen (Thermo Fisher Scientific)	AM9515
Manganese (II) Chloride Tetrahydrate	Fisher	AA3652636
o-Nitrophenyl-B-D-Galactopyranoside (ONPG)	Thermo Fisher Scientific	34055
Phenol (Buffer Saturated Phenol)	Invitrogen (Thermo Fisher Scientific)	15513-039
Protease Inhibitor Cocktail	Sigma-Aldrich	P2714-1BTL
Rifampicin	Fisher	BP2679250
Trizol Reagent	Invitrogen (Thermo Fisher Scientific)	15596026
SpeedBead Magnetic Carboxylate Beads	Cytiva	65152105050250'
SPRI beads	Prepared as in ref 50	N/A
Nuclease Free Water	Invitrogen (Thermo Fisher Scientific)	10977023
Streptavidin Mag Sepharose	Cytiva	28985799

Supplemental Table 4.2 Oligonucleotides used in this study.

The modification codes correspond to those used for IDT ordering. /5bioSG/: “standard” 5’ biotin. /5Phos/: 5’ phosphate. /3AmMO/: 3’ amino modifier.

Stock #	Name	Sequence (5' to 3')	Purification	Source
#0005	M13/pUCfor	CCCAGTCACGACGTTGTAAAAC	Desalt	IDT
#0026	M13/pUCrev	AGCGGATAACAATTTACACACAGG	Desalt	IDT
#0110	PRE T7 wt template for	GGATCCGATGACACTAATAC	Desalt	Eton
#0139	9N_VRA3_Oligo	/5Phos/rNrNrNrNrNNNGATCGTCGGACTGTAGAACTCTGAA C/3AmMO/	HPLC	IDT
#0155	PRA1_SC1_F_1	acctctggcggTgataatggttgcATGGCCTTCGGGCCAATGCAAAGGGG	Desalt	Eton
#0156	PRA1_SC1_F_2	ttatcaaaaagagtattgacTctttacctctggcggTgataatg	Desalt	Eton
#0157	PRA1_NoMod.F	TTATCAAAAAGAGTATTGACT CTTTACCTCTGGCGGTGATA ATGGTTGCAT	HPLC	IDT
#0159	dRP1_5Bio.R	/5Biosg/AATGATACGGCGACCACCGAGATCTACACGTTCCAGA GTTCTACAGTCCGACGATC	HPLC	IDT
#0160	SC1_Brdg_PLUS	CCTTGGCACCCGAGAATTCCARRRYATGGCCTTCGGGCCA A	HPLC	IDT
#0161	SC1_Brdg_MINUS	CCTTGGCACCCGAGAATTCCAYYYRRATGGCCTTCGGGCCA A	HPLC	IDT
#0162	dRP1_NoMod.R	AATGATACGGCGACCACCGAGATCTACACGTTCCAGAGTTCTA CAGTCCGACGATC	HPLC	IDT
#0163-0178; #0235-0248	RPI (Index 1-30 sequences derived from Illumina Truseq Small RNA Kit)	CAAGCAGAAGACGGCATACGAGAT[INDEX]GTGACTGGAGTTCC CTTGGCACCCGAGAATTCCA	HPLC	IDT
#0187	Alx' extension	GTGCCCCACGACGCCCTGCAACAACAGGTCGATAGCCAGCAT	Desalt	Eton
#0188	alx_extension_rev	GTGCCCCACGACG	Desalt	Eton
#0189	0189_alx_extension_rev_5bi o	/5Biosg/GTGCCCCACGACGCC	Desalt	IDT
#0202	0202_DAN1	AGC GGT ACT GCA GTA ATA CGA CTC ACT ATA G	Desalt	Eton
#0203	0203_DAN2	TAG AAG TTC CTT ACA CAT AAA AAA ACG GCC AAC G	Desalt	Eton
#0204	0204_DAN3	GTA AGT GAG ACC TTG CCG GAA GGC GAG	Desalt	Eton
#0229	0229_PCR1_for	GAGCTCGAATTAATCCCTGCA	Desalt	Eton
#0232	0232_PCR2_rev	CACGACGTTGTAAAACGACGG	Desalt	Eton
#0257	mntP_txn_template_F_clonin g	GAGAACGGGATCCGATCGTTCCGAATATCGACGATCTGG	Desalt	IDT
#0258	mntP_txn_template_R_cloni ng	GCAAGATATTGAATCCATCGACATACCAAACGCAAG	Desalt	IDT
#0259	mntP_PRA1_SC1_F1	acctctggcggTgataatggttgcATGGCCTTCGGGCCAATCATTGGGGAC	Desalt	IDT
#0262	mntP_txn_template_f	CGTTCGAAATATCGACGATCTGG	Desalt	IDT

Supplemental Table 4.2 Oligonucleotides used in this study.(Continued)

#0263	mntP_txn_template_r	CATCGACATACCAAACGCAAG	Desalt	IDT
#0264	PRA1_SC1_F_2_mntP_clonir	GGAACGGGATCCGATttatcaaaaagagtattgacTCttttacctctggcggTgata	Desalt	IDT
#0265	cotrans_mntP_cloning_R_2	GCAAGATATTGAATTCCATCGACATACC	Desalt	IDT
#0267	mntP_cotrans_template_R_5 bio	/5Biosg/GCAAGATATTGAATTCCATCGACATACCAAACGCAAGAA G	Desalt	IDT
#0273	0273_alx_aptamer_txn_rev	CCGGCAAGGTCTCACTTA	Desalt	Eton
#0274	0274_mntP_aptamer_txn_re v	CATATGGTCTCGCCTGACTG	Desalt	Eton
#0277	0277_RT_primer_SHAPEcas sette	gaaccggaccgaagcccg	Desalt	Eton
#0278	0278_Step1_rev_SHAPEcas sette	CCCTACACGACGCTCTTCCGATCTNNNNNGaaccggaccgaagcccg	Desalt	IDT
#0279	0279_Step1_fwd_SHAPEcas sette	CCTTGGCACCCGAGAATTCANNNNNGgccttcgggcca	Desalt	IDT
#0287	mntP_elemental_mapping_1 _RNA	rArCrGrUrArCrGrG	Desalt	IDT
#0288	mntP_elemental_mapping_1 _tDNA	CGATTGATGTGCATATGGTCTCGCCTGACTGAAAGTATTCAGTC CGTACGTAGGACGC	Desalt	IDT
#0289	mntP_elemental_mapping_1 _ntDNA	GCGTCCTACGTACGGACTGAATACTTTCAGTCAGGCGAGACCAT ATGCACATCAATCG	Desalt	IDT
#0290	mntP_elemental_mapping_2 _RNA	rCrCrGrUrCrArGrGrCrG	Desalt	IDT
#0291	mntP_elemental_mapping_2 _tDNA	GCAGGCATAGCGATTGATGTGCATATGGTCTCGCCTGACTGAAA GTATTCAAGGACGC	Desalt	IDT
#0292	mntP_elemental_mapping_2 _ntDNA	GCGTCCTTGAATACTTTCAGTCAGGCGAGACCATATGCACATCA ATCGCTATGCTGCG	Desalt	IDT
#0297	mntP_FL_SHAPE_txn_templ ate_F	GGTACCTCGCGAATGCATCTAG	Desalt	IDT
#0298	mntP_FL_SHAPE_txn_templ ate_R	GAACCGGACCGAAGCCC	Desalt	IDT
#0305	mntP_mapping_scaffold_3_n tDNA	CAGGCGAGACCATATGCACATCAATCGCTATGCCTGCATGAAC GCGAAAACCATCACGG	Desalt	IDT
#0306	mntp_mapping_scaffold_3_t DNA	CCGTGATGGTTTTTCGCGTTCATGCAGGCATAGCGATTGATGTGC ATATGGTCTCGCCTG	Desalt	IDT
#0307	mntP_mapping_scaffold_3_ RNA1	rCrCrArUrArUrGrCrArC	Desalt	IDT
#0308	mntP_mapping_scaffold_3_ RNA2	rCrUrArUrGrCrCrUrGrC	Desalt	IDT
#0452	0452_PRE_SL_RBS_termbl o	/5Biosg/GCAACGGCGTGCCGAC	None	IDT
#0455	0455_mntP_lacZ_fwd	CAT CAA TCG CTA TGC CTG CAT GAA CG	Desalt	Eton
#0458	0458_PRE_A37U_F	CGGTCGATCGTCTTTACGATGTGTG	Desalt	Eton
#0459	0459_PRE_A37U_R	CACACATCGTAAAGACGATCGACCG	Desalt	Eton
RAV23	RAV23	CATGGCATGTGGATCCGCCCCACGACGCCCCCTGCAACAACA	Desalt	Eton
RAV126	RAV126	CTCGTTCCGTGCGACTCCATCGACATACCAAAC	Desalt	Eton

Supplemental Table 4.3 DNA templates used in this study.

The table below summarizes the linear double-stranded DNA templates prepared for this study, including the primers used for PCR amplification, the template for PCR, modifications (if applicable), the purification method for the DNA template, and the figure(s) in which each DNA template was used.

ID	Description	Fwd primer	Rev primer	Template	Modifications	Clean-up	Used in Fig(s)
1	Template contains PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>alx</i> riboswitch (nt 17-223), and nt 1-100 of the <i>alx</i> gene.	#0157	#0188	pMIS313	N/A	Phenol extraction/Ethanol precipitation	N/A
2	Template contains PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>alx</i> riboswitch (nt 17-223), and nt 1-100 of the <i>alx</i> gene.	#0157	#0189	Template 1	Terminal biotin	Phenol extraction/Ethanol precipitation	N/A
3	Template contains PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>alx</i> riboswitch (nt 17-223), and nt 1-100 of the <i>alx</i> gene.	#0157	#0189	Template 2	Internal biotin-11 dNMPs, Terminal biotin	Phenol extraction/ Ethanol precipitation	Figs. 4, 5, 7, and 8; SI Figs. 7,8,9,10,11,15,16,17, 19
4	Template contains wild-type <i>mntP</i> promoter, wild-type <i>mntP</i> riboswitch (nt 1-225), and nt 1-45 of the <i>mntP</i> gene.	#0262	#0263	pMIS309	N/A	Monarch PCR & DNA Cleanup Kit (New England Biosciences)	SI Fig. 4
5	Template contains PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>mntP</i> riboswitch (nt 17-241), and nt 1-45 of the <i>mntP</i> gene.	#0157	#0267	pMIS315	Terminal biotin	Phenol extraction/ Ethanol precipitation	N/A
6	Template contains PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>mntP</i> riboswitch (nt 17-241), and nt 1-45 of the <i>mntP</i> gene.	#0157	#0267	Template 5	Internal biotin-11 dNMPs, Terminal biotin	Phenol extraction/ Ethanol precipitation	Figs. 3,5, and 6; SI Figs. 3,6,9,10,11,12,13,18
7	Template contains PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>alx</i> riboswitch (nt 17-223), and nt 1-25 of the <i>alx</i> gene	#0157	#0452	Template 1	Terminal biotin	Monarch PCR & DNA Cleanup Kit (New England Biosciences) or Phenol extraction/ Ethanol precipitation	SI Fig. 19B
8	Template contains T7 promoter, 5' SC1 hairpin (nt 1-14), wild-type <i>alx</i> riboswitch (nt 15-238), and a 3' linker and RT primer anneal site (nt 239-281)	#0026	#0277	pMIS511	N/A	DpnI digest	SI Fig. 20
9	Template contains T7 promoter, 5' SC1 hairpin (nt 1-14), wild-type <i>alx</i> riboswitch aptamer (nt 15-144), and a 3' linker and RT primer anneal site (nt 145-187)	#0026	#0277	pMIS513	N/A	DpnI digest	Fig. 2 and SI Fig. 1
10	Template contains T7 promoter, 5' SC1 hairpin (nt 1-14), wild-type <i>mntP</i> riboswitch (nt 15-239), and a 3' linker for RT primer annealing (nt 240-264)	#0297	#0298	pMIS314	N/A	DpnI digest	SI Fig. 14

Supplemental Table 4.4 *E. coli* K-12 strains used in this study.

All strains employed in this study are *E. coli* K12 derivatives. The $\Delta alx::Kan$ deletion mutation was sourced from the library of Keio collection (SI ref 8).

Strain	Genotype
MC4100	$\Delta(argF-lac)U169 rpsL150$ $relA1spoT1araD139 flbB5301 deoC1$ $ptsF25$
RAS31	MC4100 $\Delta alx::Kan$

Supplemental Table 4.5 Plasmids used in this study.

Plasmid	Source	Description
pMIS301	This work (purchased from Bio Basic)	pUC57-Kan plasmid with synthetic construct encoding wild-type <i>alx</i> riboswitch (nt 1-207) and the first 59 nt of <i>alx</i> , transcribed from the <i>alx</i> promoter
pMIS306	Landick Lab, University of Wisconsin Madison	Overexpression plasmid for wild-type <i>E. coli</i> σ^{70} with an N-terminal His ₆ tag
pMIS309	This work	pUC19-Amp plasmid with insert encoding wild-type <i>mntP</i> riboswitch (nt 1-225) and the first 45 nt of <i>mntP</i> , transcribed from the <i>mntP</i> promoter
pMIS313	This work	pUC19-Amp plasmid with insert encoding PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>alx</i> riboswitch (nt 17-223) and the first 100 nt of <i>alx</i>

Supplemental Table 4.5 Plasmids used in this study.(Continued)

pMIS314	This work (purchased from Bio Basic)	pUC57-Kan plasmid with insert encoding T7 promoter, 5' SC1 hairpin (nt 1-14), wild-type <i>mntP</i> riboswitch (nt 15-239), and a 3' linker for RT primer annealing (nt 240-264)
pMIS315	This work	pUC19-Amp plasmid with insert encoding PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>mntP</i> riboswitch (nt 17-241) and the first 45 nt of <i>mntP</i> .
pMIS349	This work	pMU2386 containing translational reporter of <i>alx</i> clone in the PstI and BamHI sites. Contains A37T mutation in the <i>alx</i> riboswitch sequence
pMIS511	This work	pUC57-Kan plasmid with insert encoding T7 promoter, 5' SC1 hairpin (nt 1-14), wild-type <i>alx</i> riboswitch (nt 15-238), and a 3' linker and RT primer anneal site (nt 239-281)

Supplemental Table 4.5 Plasmids used in this study.(Continued)

pMIS513	This work	pUC57-Kan plasmid with insert encoding T7 promoter, 5' SC1 hairpin (nt 1-14), wild-type alx riboswitch aptamer (nt 15-144), and a 3' linker and RT primer anneal site (nt 145-187)
pMIS447	This work (purchased from BioBasic)	pUC57-Kan plasmid with insert encoding a 50 nt spacer, -35 element, -10 element, 5' SC1 hairpin (nt 1-14), wild-type alx riboswitch (nt 15-298), 3' linker and RT primer anneal site (nt 299-341), and rrnB T1 and T2 terminators
pMIS518	This work	pUC19-Amp plasmid with insert encoding a 50 nt spacer, -35 element, -10 element, 5' SC1 hairpin (nt 1-14), wild-type alx riboswitch (nt 15-298), 3' linker and RT primer anneal site (nt 299-341), and rrnB T1 and T2 terminators
pMIS501	This work	pMU2386 containing translational reporter of T7 promoter- <i>alx</i> in <i>Pst</i> I and <i>Bam</i> HI sites

Supplemental Table 4.5 Plasmids used in this study.(Continued)

pMIS517	This work	pMU2386 containing translational reporter of T7 promoter- <i>mntP</i> in <i>PstI</i> and <i>Sall</i> sites
pRA46	This work	Overexpression plasmid for wild-type <i>E. coli</i> NusA with an N-terminal His ₆ tag
pRA53	Sharma et. al., Journal of Bacteriology, 2024 (ref 19)	promoter-less pMU2386 (empty vector control for lacZ reporter fusion assays)
pRA54	Sharma et. al., Journal of Bacteriology, 2024 (ref 19)	pMU2386 containing translational reporter of <i>alx</i> clone in the <i>PstI</i> and <i>BamHI</i> sites
pRA57	Sharma et. al., Journal of Bacteriology, 2024 (ref 19)	pMU2386 containing translational reporter of <i>mntP</i> clone in the <i>PstI</i> and <i>Sall</i> sites
pRM756	Landick Lab, University of Wisconsin Madison	Wild-type <i>E. coli</i> RNAP ($\alpha 2\beta\beta'\omega$) overexpression plasmid. His ₁₀ -ppx tag at <i>rpoC</i> -C-terminus

Supplemental Table 4.6 Sequences of DNA templates used in this study.

ID	Template description	Sequence
#1-3	Template contains PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>alk</i> riboswitch (nt 17-223), and nt 1-100 of the <i>alk</i> gene.	ttatcaaaaagagtattgacTctttacctctggcgTgataatggttgcATGGCCTTCGGGCCAATGCAAGGGGAGTAA CTTCATTGCCGGTCGATCGTCATTACGATGTGTGAAAAACACATCCGGTCACCGGGCAACCCGAAAGG AATACGCAGACGTATTCCTTTTTTGTGTAAGTGAGACCTTGCCGGAAGGCAGGCTCTATGCATAAAAGCA GCGGCTGACGTCTTCGACGTTGGCCGTTTTTTATGTGAAGGAACCTCTATGAATACTGTCGGCAGGCC GTTGCTATGGGGCGGATTGCTGTTGTTGTCGCCATTATGCTGGCTATCGACCTGTTGTTGACAGGGGCGT CGTGGGGCAC
#4	Template contains wild-type <i>mntP</i> promoter, wild-type <i>mntP</i> riboswitch (nt 1-225), and nt 1-45 of the <i>mntP</i> gene.	CGTTCGAAATATCGACGATCTGGAATAATTATAAGCTTCTCGTTCAACTCAATAAGTTATGAATTAGCCAA AGCTATGTTTAGTGATTTTTTAATAATCAGACATAGCTTAGGCTATATTACCTCTCCCTTATTGTTATTTTAA ACGTTTCATTGATATATAAATCCAAATGAAATCGTTATCAATAAAGCAATGAAATAATATATTCCAACAGTTGTTT TATATTCTCAAATATGTTAAGGTTGCGCCCTCATTGGGGAGTAGCCGATTCCAGATTCCGGAATGTAC GTGTCAACATACTGTTGCAAAACGTGGCAGCTACGGACTGAATCTTTCAGTCAGGCGAGACCATATGC ACATCAATCGCTATGCCCTGCATGAACGCGAAAACCATCACGGCTTGTGTTTCATGTATGCATGTTTCTGCTGG GGCGCATGATGTGTTTATGGATACCCCGGTCAGGACATTGTCatGAATACACTGCTACTGTTCTTCTTGC GTTGGTATGTCGATG
#5-6	Template contains PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>mntP</i> riboswitch (nt 17-241), and nt 1-45 of the <i>mntP</i> gene.	ttatcaaaaagagtattgacTctttacctctggcgTgataatggttgcATGGCCTTCGGGCCAATCATTGGGGAGTAG CCGATTCCAGATTCCGGAATGTACGTGTCAACATACTCGTTGCAAAACGTGGCAGCTACGGACTGAATA CTTCAGTCAGGCGAGACCATATGCACATCAATCGCTATGCCCTGCATGAACGCGAAAACCATCACGGCTT GTTGTTTCATGTATGCATGTTTGTGTTGGGGCGATGATGTGTTTATGGATACCCCGGTCAGGACATTGTCatg AATACACTGCTACTGTTCTTCTTGGCTTGGTATGTCGATG
#7	Template contains PRA1 promoter, 5' SC1 hairpin (nt 1-16), wild-type <i>alk</i> riboswitch (nt 17-223), and nt 1-25 of the <i>alk</i> gene	ttatcaaaaagagtattgacTctttacctctggcgTgataatggttgcATGGCCTTCGGGCCAATGCAAGGGGAGTAA CTTCATTGCCGGTCGATCGTCATTACGATGTGTGAAAAACACATCCGGTCACCGGGCAACCCGAAAGG AATACGCAGACGTATTCCTTTTTTGTGTAAGTGAGACCTTGCCGGAAGGCAGGCTCTATGCATAAAAGCA GCGGCTGACGTCTTCGACGTTGGCCGTTTTTTATGTGAAGGAACCTCTATGAATACTGTCGGCAGGCC GTTGC
#8	Template contains T7 promoter, 5' SC1 hairpin (nt 1-14), wild-type <i>alk</i> riboswitch (nt 15-238), and a 3' linker and RT primer anneal site (nt 239-281)	TAATACGACTCACTATAGGGggccttcgggcaaTGCAAGGGGAGTAACCTCATTGCCGGTCGATCGTCATTA CGATGTGTGAAAAACACATCCGGTCACCGGGCAACCCGAAAGGAATACGCAGACGTATTCCTTTTTTGT GTAAGTGAGACCTTGCCGGAAGGCGAGGTCTATGCATAAAAGCAGCGGCTGACGTCTTCGACGTTG GCCGTTTTTATGTGAAGGAACCTCTATGAATACTGTCGATCCTCGATCCGGTTCGCCGATCCAAATcgg gcttcggtccggttc
#9	Template contains T7 promoter, 5' SC1 hairpin (nt 1-14), wild-type <i>alk</i> riboswitch aptamer (nt 15-144), and a 3' linker and RT primer anneal site (nt 145-187)	TAATACGACTCACTATAGGGggccttcgggcaaTGCAAGGGGAGTAACCTCATTGCCGGTCGATCGTCATTA CGATGTGTGAAAAACACATCCGGTCACCGGGCAACCCGAAAGGAATACGCAGACGTATTCCTTTTTTGT GTAAGTGAGACCTTGCCGGTTCGATCCGGTTCGCCGGATCCAAATcgggcttcggtccggttc
#10	Template contains T7 promoter, 5' SC1 hairpin (nt 1-14), wild-type <i>mntP</i> riboswitch (nt 15-239), and a 3' linker for RT primer annealing (nt 240-264)	TTAATACGACTCACTATAGGGGGCTTCGGGCCAATCATTGGGGAGTAGCCGATTCCAGATTCCGGAA ATGTACGTGTCAACATACTCGTTGCAAAACGTGGCAGCTACGGACTGAATCTTTCAGTCAGGCGAGACC ATATGCACATCAATCGCTATGCTGCATGAACGCGAAAACCATCACGGCTTGTGTTTCATGTATGCATGTTT GCTGGGGCGATGATGTTTATGGATACCCCGGTCAGGACATTGCTCGATCCGGTTCGCCGGATCC AAATCGGGCTTCGGTCCGGTTC

Supplemental Table 4.7 Biosample information for all co-transcriptional chemical probing experiments performed in this work.

[illegible]

Supplemental References

1. Bachas,S.T. and Ferré-D'Amaré,A.R. (2018) Convergent Use of Heptacoordination for Cation Selectivity by RNA and Protein Metalloregulators. *Cell Chem. Biol.*, **25**, 962-973.e5.
2. Suddala,K.C., Price,I.R., Dandpat,S.S., Janeček,M., Kührová,P., Šponer,J., Banáš,P., Ke,A. and Walter,N.G. (2019) Local-to-global signal transduction at the core of a Mn²⁺ sensing riboswitch. *Nat. Commun.*, **10**, 4304.
3. Price,I.R., Gaballa,A., Ding,F., Helmann,J.D. and Ke,A. (2015) Mn²⁺-Sensing Mechanisms of *yybP-ykoY* Orphan Riboswitches. *Mol. Cell*, **57**, 1110–1123.
4. Kang,J.Y., Mishanina,T.V., Bellecourt,M.J., Mooney,R.A., Darst,S.A. and Landick,R. (2018) RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. *Mol. Cell*, **69**, 802-815.e5.
5. Kang,J.Y., Mishanina,T.V., Landick,R. and Darst,S.A. (2019) Mechanisms of Transcriptional Pausing in Bacteria. *J. Mol. Biol.*, **431**, 4007–4029.
6. Chauvier,A., Dandpat,S.S., Romero,R. and Walter,N.G. (2024) A nascent riboswitch helix orchestrates robust transcriptional regulation through signal integration. *Nat. Commun.*, **15**, 3955.
7. Stephen,C. and Mishanina,T.V. (2022) Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the *Escherichia coli* pH-responsive riboswitch. *J. Biol. Chem.*, **298**, 102302.
8. Baba,T., Ara,T., Hasegawa,M., Takai,Y., Okumura,Y., Baba,M., Datsenko,K.A., Tomita,M., Wanner,B.L. and Mori,H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.0008.
9. Heckman,K.L. and Pease,L.R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.*, **2**, 924–932.
10. Sharma,R. and Mishanina,T.V. (2024) A riboswitch-controlled TerC family transporter Alx tunes intracellular manganese concentration in *Escherichia coli* at alkaline pH. *J. Bacteriol.*, **206**, e00168-24.
11. Strobel,E.J., Watters,K.E., Nedialkov,Y., Artsimovitch,I. and Lucks,J.B. (2017) Distributed biotin–streptavidin transcription roadblocks for mapping cotranscriptional RNA folding. *Nucleic Acids Res.*, **45**, e109–e109.
12. Saba,J., Chua,X.Y., Mishanina,T.V., Nayak,D., Windgassen,T.A., Mooney,R.A. and Landick,R. (2019) The elemental mechanism of transcriptional pausing. *eLife*, **8**, e40981.

7. Acknowledgements

Chapter 4, in full, is a reprint of the material in peer review as it appears in Structurally distinct manganese-sensing riboswitch aptamers regulate diverse expression platform architectures. Stephen, Christine; Palmer, Dana; Bautista, Clarisa; Mishanina, Tatiana, in peer review at Nucleic Acids Research, 2025. The dissertation author was the primary investigator and author of this material.

Chapter 5 Opportunities for riboswitch inhibition by targeting co-transcriptional RNA folding events

Opportunities for Riboswitch Inhibition by Targeting Co-Transcriptional RNA Folding Events

Christine Stephen [†], Dana Palmer [†] and Tatiana V. Mishanina *

Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA 92093, USA; depalmer@ucsd.edu (D.P.)

* Correspondence: tmishanina@ucsd.edu

[†] These authors contributed equally to this work.

Citation: Stephen, C.; Palmer, D.; Mishanina, T.V. Opportunities for Riboswitch Inhibition by Targeting Co-Transcriptional RNA Folding Events. *Int. J. Mol. Sci.* **2024**, *25*, 10495. <https://doi.org/10.3390/ijms251910495>

Academic Editor: Constantinos Stathopoulos

Received: 21 August 2024 Revised: 27 September 2024 Accepted: 27 September 2024

Published: 29 September 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

0. Abstract

Antibiotic resistance is a critical global health concern, causing millions of prolonged bacterial infections every year and straining our healthcare systems. Novel antibiotic strategies are essential to combating this health crisis and bacterial non-coding RNAs are promising targets for new antibiotics. In particular, a class of bacterial non-coding RNAs called riboswitches has attracted significant interest as antibiotic targets. Riboswitches reside in the 5'-untranslated region of an mRNA transcript and tune gene expression levels *in cis* by binding to a small-molecule ligand. Riboswitches often control expression of essential genes for bacterial survival, making riboswitch inhibitors an exciting prospect for new antibacterials. Synthetic ligand mimics have predominated the search for new riboswitch inhibitors, which are designed based on static structures of a riboswitch's ligand-sensing aptamer domain or identified by screening a small-molecule library. However, many small-molecule inhibitors that bind an isolated riboswitch aptamer domain with high affinity *in vitro* lack potency *in vivo*. Importantly, riboswitches fold and respond to the ligand during active transcription *in vivo*. This co-transcriptional folding is often not considered during inhibitor design, and may explain the discrepancy between a low K_D *in vitro* and poor inhibition *in vivo*. In this review, we cover advances in riboswitch co-transcriptional folding and illustrate how intermediate structures can be targeted by antisense oligonucleotides—an exciting new strategy for riboswitch inhibitor design.

Keywords: bacteria; antibiotics; drug target; antisense oligonucleotide; riboswitch; riboswitch inhibitor; transcription; translation; RNA folding

1. Introduction

Antibiotic resistance is a major threat to human health and causes millions of prolonged bacterial infections each year, risking overwhelming our healthcare systems [1]. Due to the

widespread use of antibiotics to control infection, pathogenic bacteria have developed resistance to antibiotics that were previously effective, causing deaths from once-treatable infections [2]. Motivated by the pressing need for innovative antibiotic strategies, the World Health Organization (WHO) creates lists of priority bacterial pathogens to focus on when designing novel antibiotics [3]. One promising avenue for discovering new antibiotics is to target bacterial non-coding RNAs that control the expression of essential genes in pathogenic bacteria, such as riboswitches [4]. Riboswitches are 5'-untranslated regions of mRNA that typically respond to small-molecule ligands to control downstream gene expression via regulating transcription, translation, or RNA decay [5–7]. Importantly, with the exception of the thiamine pyrophosphate (TPP) riboswitch, riboswitches are only found in bacteria and archaea [8], making them promising targets for selective antibiotics with low side effects to the human host. Riboswitches are composed of two domains: an aptamer domain that binds a ligand, typically with high affinity, and an expression platform that undergoes structural changes to affect gene expression (Figure 5.1A,B) [9]. There are at least 55 classes of characterized riboswitches that are grouped by what ligand they bind, typically with a highly conserved aptamer domain [10]. Riboswitch expression platforms exhibit more sequence and structural diversity, and either promote or inhibit gene expression through ligand-dependent structural changes (Figure 5.1A,B) [5]. Expression platforms are designated as “off” or “on” depending on whether ligand binding prompts structural changes that downregulate (“off”) or upregulate (“on”) gene expression levels. “Off” riboswitch expression platforms form either an intrinsic terminator hairpin that prompts transcription termination (Figure 5.1A) or a hairpin stem that sequesters the ribosome binding site, leading to inhibition of translation (Figure 5.1B). Conversely, for “on” riboswitch expression platforms, ligand binding leads to the formation

of an anti-terminator stem or release of the ribosome binding site, promoting downstream gene expression [5,11].

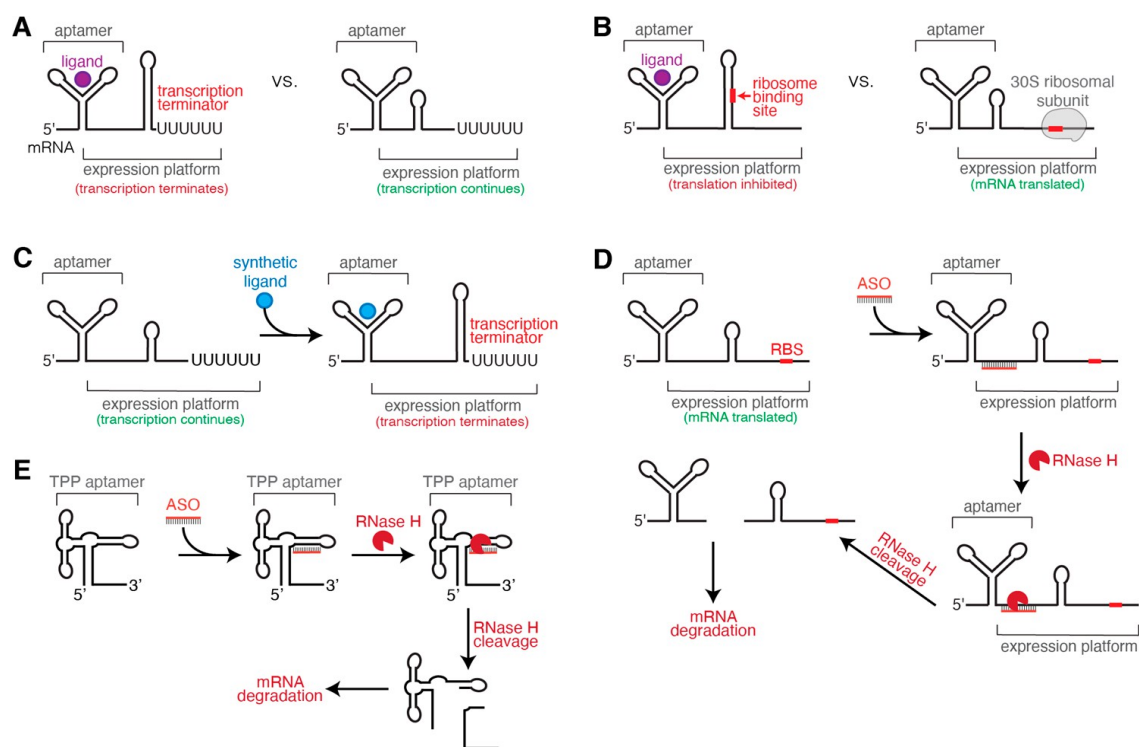


Figure 5.1 Overview of riboswitch-controlled gene expression and riboswitch inhibition strategies.

(A) Riboswitch responding to ligand by terminating transcription, turning gene expression “off”. (B) Riboswitch responding to ligand by inhibiting translation, turning gene expression “off”. (C,D) Mechanism for inhibiting gene expression by targeting a riboswitch with a synthetic ligand analog (C) or a DNA-based antisense oligonucleotide (ASO) that marks mRNA for degradation via RNase H cleavage (D). (E) Example ASO-targeting of the TPP riboswitch aptamer domain, followed by RNase H cleavage and mRNA degradation.

Riboswitches are typically found upstream of essential genes for metabolite biosynthesis and transport, thereby tuning expression of proteins that are essential for survival or virulence in some bacterial pathogens [12,13]. Since riboswitches are not found in human cells and often bind their ligands with high affinity *in vitro* (low nM K_d values) and high selectivity, the design of structural analogs that inhibit bacterial growth has attracted significant research interest [14]. Designing synthetic ligands to bind riboswitches with high affinity is the dominant focus of current research aiming to inhibit gene expression by targeting riboswitches (Figure 5.1C). These specific riboswitch inhibitors have been extensively reviewed in several recent papers [4,14–16].

To be an effective antibiotic that targets a riboswitch, compounds need to have high affinity for the riboswitch RNA, lack off-target effects on other metabolic pathways, have potential to avoid the bacteria developing antibiotic resistance, and pose low toxicity to human cells [15]. Importantly, many compounds that have high affinity for a riboswitch *in vitro* will not end up affecting gene expression *in vivo*, leading some researchers to turn to alternative methods, such as RNA-based gene silencing, to target riboswitches and inhibit bacterial gene expression (Figure 5.1D,E) [17,18]. Riboswitches fold co-transcriptionally and ligands often bind an early RNA folding intermediate, leading to the rearrangement of the expression platform as it is emerging from RNA polymerase (RNAP) [19]. Therefore, an inhibitor compound that has high affinity for the aptamer domain in isolation may lack potency *in vivo* if the inhibitor never reaches thermodynamic equilibrium with the riboswitch target, leading to a discrepancy whereby a low K_d does not indicate an effective inhibitor [17,20]. An exciting new strategy for synthetic RNA silencing of riboswitches, antisense oligonucleotides (ASOs), could potentially bridge this gap [18]. The aim of this review is to illustrate how ASOs can be a new tool for the development of antibiotics that target riboswitches, helping to overcome challenging cases where ligand mimics are unsuccessful. In the following sections, we first review the limitations of riboswitch small-molecule inhibitors, then shift the discussion to RNA-based gene silencing, covering how this occurs in nature and how such discoveries have inspired the design of synthetic oligonucleotides as a therapeutic strategy. From there, we discuss investigations into ASOs currently being researched to inhibit riboswitch-controlled gene expression in pathogenic bacteria. Lastly, we lay the foundation for how advances in riboswitch co-transcriptional folding can be leveraged to expand the available sequence space for ASO inhibitors targeting riboswitches.

2. Limitations in Designing Synthetic Ligands That Target Riboswitches in Pathogenic Bacteria

Despite some success in designing small-molecule riboswitch inhibitors, we introduce here some key riboswitches that are present in pathogenic bacteria for which significant limitations exist in the context of designing small-molecule inhibitors. Further, these riboswitches are excellent candidates for novel inhibition strategies, since they control expression of essential genes in bacteria. In Section 5, we continue the discussion of these riboswitches by highlighting how co-transcriptional folding studies can inspire the design of novel ASO inhibitors. This section is not intended to be a comprehensive review of small-molecule inhibitors targeting riboswitches present in pathogenic bacteria. For recent extensive reviews focused on small-molecule inhibitors, readers are directed to the following references [4,14–16].

2.1. *TPP Riboswitch*

The thiamine pyrophosphate (TPP) riboswitch is one of the most widespread riboswitches, found in a variety of pathogenic bacteria [21]. Thiamine is a critical cofactor for carbohydrate metabolism, and its biosynthetic pathway requires the expression of many unique proteins [22,23]. The TPP riboswitch is the only known riboswitch to occur in eukaryotic cells, with examples in plants and fungi [24]. The TPP aptamer, or THI-box, is highly conserved, and is made up of five stems that compact stably upon TPP binding [25,26]. The TPP riboswitch expression platform differs between bacterial species, and can either promote transcription termination or prevent translation initiation upon TPP binding the aptamer domain [27]. For example, in *E. coli*, there are multiple important TPP riboswitches that control downstream gene expression at both the transcriptional and translational level [28]. To highlight this diversity, the *E. coli thiC* riboswitch regulates the expression of proteins directly involved in TPP biosynthesis,

while the *tpbA* riboswitch regulates the expression of an ABC transporter necessary to import thiamine and TPP into the bacterial cell [22,23].

A synthetic analog of thiamine (pyrithiamine, or PT) binds the thiamine pyrophosphate (TPP) riboswitch with nearly identical affinity to the natural TPP ligand [29]. PT was first characterized as an inhibitor of TPP-dependent enzyme production in *E. coli* in 1967, but mutant *E. coli* strains that resist PT inhibition by derepressing TPP-dependent enzyme synthesis by an unknown mechanism were isolated by 1969 [30,31]. In 2005, it was demonstrated that bacteria develop antibiotic resistance to PT through mutations to the TPP riboswitch, with mutations occurring in both the aptamer domain and expression platform [29]. Another study characterized binding by thiamine, methylene diphosphoric acid (MDP, a pyrophosphate analog), and a covalently linked version of the two molecules (TPPc) to the TPP riboswitch [32]. The individual thiamine and MDP bound with low affinity, but the conjugate molecule exhibited 6-fold higher affinity than TPP itself. The thiamine moiety of TPPc binds the riboswitch in a different conformation than the thiamine moiety of TPP, making it an interesting candidate to test *in vivo* [32]. Other studies have identified limited candidate inhibitors that bind the TPP riboswitch with sufficient affinity *in vitro*, but those have not been studied extensively *in vivo* [15].

2.2. FMN Riboswitch

The flavin mononucleotide (FMN) riboswitch class is widely distributed across bacteria and is found upstream of a variety of genes involved in flavin synthesis and transport [27]. Flavin-based cofactors are essential for proteins involved in the electron transport processes [33]. Since FMN is an essential cofactor for bacterial survival, inhibiting the flavin synthesis or transport pathways has attracted interest for antibiotic development. The FMN riboswitch aptamer domain has a six-stem junction that folds into a butterfly-like scaffold by identically folding the peripheral

domains [34]. Additionally, the FMN riboswitch aptamer is largely folded in the absence of a ligand, and FMN binding only leads to local conformational changes [35]. These smaller-scale conformational changes ultimately alter the fold of the expression platform, leading to transcription termination or inhibition of translation initiation. For example, in *B. subtilis*, the *ribD* riboswitch promotes transcription termination upon FMN binding, leading to a downregulation of several FMN biosynthetic genes located in a single operon controlled by the riboswitch [36].

The FMN riboswitch can successfully be targeted by roseoflavin, which is phosphorylated by cellular proteins, but this requirement limits its efficiency [37]. Additionally, roseoflavin binds many off-target flavoproteins that are also present in humans, which gives it limited potential as an antibiotic [38,39]. Another ligand-mimic, ribocil, targets the FMN riboswitch aptamer with high affinity [40]. However, ribocil-resistant *E. coli* mutants have been identified, and importantly, all mutations conferring resistance mapped to the riboswitch aptamer domain or expression platform [40,41].

2.3. *glmS* Ribozyme

The *glmS* ribozyme is a self-cleaving ribozyme that responds to glucosamine-6-phosphate (GlcN6P) and is located upstream of GlcN6P synthetase [42]. GlcN6P synthetase is responsible for producing GlcN6P [42], which is the essential first step in the pathway to building a bacterial cell wall, so inhibiting this production prevents bacterial growth [43]. GlcN6P binding initiates the rapid self-cleavage of the *glmS* ribozyme, causing the release of a 25-nucleotide (nt) leader mRNA sequence that promotes subsequent degradation of the mRNA by RNase J1 [43]. The *glmS* ribozyme has a compact structure containing multiple pseudoknots that is largely unchanged by ligand binding [44]. Instead, GlcN6P participates directly in catalyzing the self-cleavage event [43,44].

Thousands of ligand analogs for the *glmS* ribozyme have been identified and screened, but very few were found to be active in promoting ribozyme self-cleavage. Of these, none activated the ribozyme as strongly as the natural ligand, making them less promising antibiotics [45–47].

2.4. Purine Riboswitch

Purine riboswitches respond to purine molecules and include guanine and adenine riboswitches [48]. These riboswitches occur upstream of genes involved in essential purine metabolism pathways, purine salvage pathways, and purine transport [48,49]. Thus, designing inhibitors for these essential pathways would lead to a loss of bacterial virulence [50,51]. Purine riboswitches have conserved aptamer domains consisting of three helices that nearly fully surround the three-way junction that coordinates the ligand [52]. In fact, the only difference between the adenine and guanine riboswitch aptamer binding pockets is one nucleobase that directly base-pairs with the ligand [53].

Purine riboswitches can be successfully targeted by rationally designed small molecules [54]. However, these compounds often show limited affinity for the riboswitch aptamer [49]. The compounds that are identified to have high affinity often show little to no effect on gene expression [17,20]. One example study generated 16 rationally designed small molecules and identified 6-*N*-hydroxylaminopurine (6-*N*-HAP) as a promising guanine analog that binds the guanine riboswitch aptamer with high affinity *in vitro* [49]. Unfortunately, when added to media, 6-*N*-HAP demonstrated low potency, since a high concentration was required to inhibit bacterial growth. Additionally, other designed compounds in this study inhibited growth, but exhibited off-target effects [49], and 6-*N*-HAP has even been demonstrated to be a general and potent mutagen in bacterial and mammalian cells [55,56]. These limitations to using small-molecule

ligand mimics to inhibit riboswitch-controlled gene expression have led some researchers to explore RNA-based silencing methods, which will be reviewed in the next sections.

3. RNA-based Silencing in Natural Contexts

RNA-based gene silencing (RNA interference, or RNAi) was described in eukaryotes in 1998, when double-stranded RNAs were identified as post-transcriptional gene silencing elements in *C. elegans* [57,58]. RNAi is a conserved eukaryotic mechanism that regulates gene expression at the transcriptional or post-transcriptional level [59]. RNA silencing relies on double-stranded RNA (dsRNA) cleavage that results in small RNAs [60,61]. These small RNAs are then recruited to guide protein complexes to silence the target genes [62]. This silencing process is a key mechanism by which eukaryotes defend against viruses and transposable elements [63], but is also used to regulate endogenous genes within eukaryotic cells [64].

In prokaryotes, small non-coding RNAs (sRNAs) are responsible for diverse gene silencing events [65]. The expression of sRNAs allows for the fine-tuning of gene expression in response to various environmental signals [66]. These sRNAs can prevent translation of genes by binding and occluding ribosome binding sites, or by affecting RNA stability and promoting degradation by RNase enzymes [67,68]. Some sRNAs associate with chaperone proteins, such as Hfq or ProQ, to achieve this translation inhibition and RNase-dependent degradation [69,70]. Other sRNAs sequester CsrA-like RNA-binding proteins that affect mRNA stability and translation [71]. Additionally, sRNAs associate with Cas proteins to promote the degradation of DNA or RNA targets as part of the CRISPR/Cas defense system against viral infection [65,72,73].

Diverse RNA-based gene silencing events are thus widespread among all kingdoms of life and allow for versatile and specific changes in gene expression in response to environmental

signals. This versatility makes sRNAs a promising avenue for designing drugs to target and silence specific genes, especially in bacteria [57].

4. Synthetic RNA-Based Silencing

Drawing inspiration from natural RNA-based gene silencing, synthetic antisense oligonucleotides (ASOs) can be designed to achieve silencing of target genes. ASOs are single-stranded synthetic oligonucleotides that are complementary to a target gene [74]. ASOs can be used to silence specific genes by Watson–Crick base-pairing with target RNA transcripts and inhibiting gene expression as a result [18]. For example, an ASO may inhibit translation by targeting the RNA transcript for cleavage (Figure 5.2A) or by creating a double-stranded region that serves as a steric block for translating ribosomes or for proper RNA folding (Figure 5.2B) [75]. The first ASO shown to inhibit gene expression *in vivo* was discovered in 1978 and targeted viral RNA translation [76]. Since then, the development of ASOs as potential drugs has been widely pursued, and over a dozen ASOs have been approved by the FDA and others are in clinical trial [77,78].

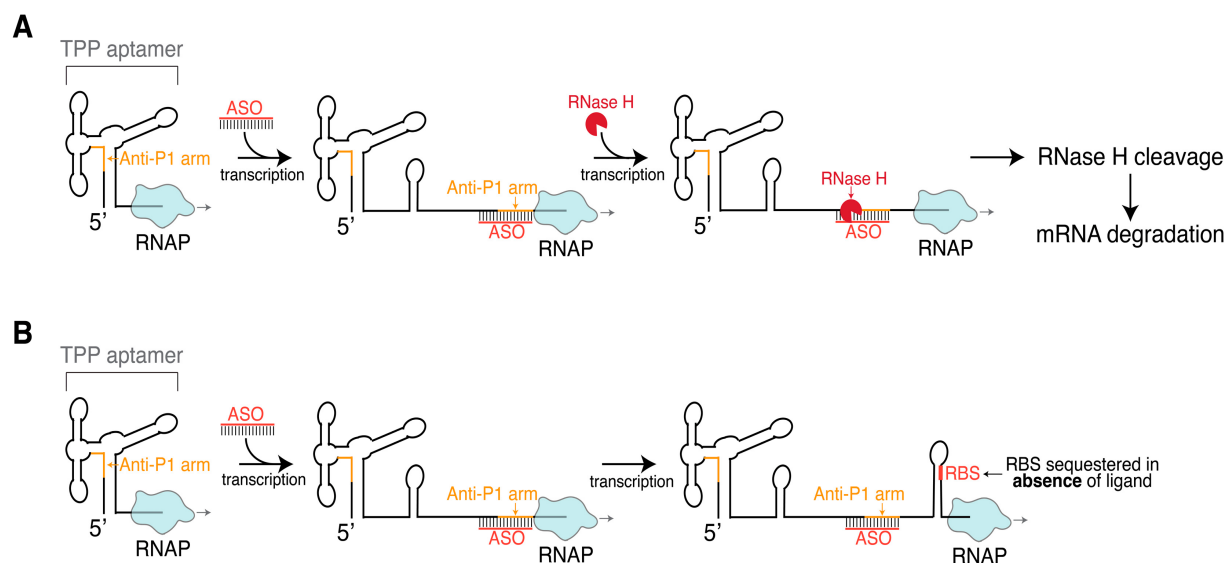


Figure 5.2 Overview of ASO-directed co-transcriptional silencing of target riboswitch-controlled genes, using the TPP riboswitch as an example.

(A) ASO binding recruits RNase H, which cleaves the DNA-RNA hybrid, prompting mRNA degradation. (B) ASO binding creates a steric block that prevents a structural rearrangement and favors the formation of the RBS sequestering stem.

Studies in the Penchovsky lab have shown that ASOs can be designed to target specific riboswitches (Figure 5.1D), including SAM-I, *glmS*, FMN, and TPP riboswitches (Figure 5.1E) from pathogenic bacteria [77]. These riboswitch-targeting ASOs have been previously reviewed [27], and the main findings are summarized briefly below.

All four riboswitch-targeting synthetic DNA ASOs were designed to target a portion of the conserved riboswitch aptamer domains [27]. These ASOs enabled cleavage of the riboswitch RNA by RNase H to degrade the RNA and prevent downstream gene expression (Figure 5.1D). For each riboswitch, the ASO was designed to be fully complementary to a 13–16 nt single-stranded sequence of riboswitch RNA that does not self-hybridize. For all four ASOs, the average minimum concentration of ASO inhibitor needed to inhibit 80% of bacterial growth (MIC_{80}) was around 700 nM [27]. To avoid toxicity of the ASO, the sequences were chosen to be dissimilar to sequences occurring in human cells, and the ASOs were shown to be non-toxic at their MIC_{80}

to a lung cancer human cell line (A549) [27]. A few different methods were used to demonstrate that the ASO was specific for a particular riboswitch. For the TPP and SAM-I riboswitches, the ASO was shown to be ineffective in targeting *E. coli*, which was expected as *E. coli* lacks these target riboswitches [21,79]. For the FMN riboswitch, a second ASO was designed with eight mismatches with the target RNA to be used as a negative control, and this mismatched ASO was ineffective at binding the FMN riboswitch [80]. Finally, the *glmS* ribozyme ASO was designed to be fully complementary to the aptamer domain of a target *glmS* ribozyme in *S. aureus*. This ASO did not inhibit the growth of *E. coli* or *B. subtilis*, demonstrating the desired specificity for the *glmS* ribozyme in *S. aureus* [81].

The advances in ASO-directed degradation of riboswitch RNAs show that riboswitches are a promising target for new antibacterials [27]. However, the requirements that the target RNA be unable to self-hybridize, single-stranded, and absent from human cells limit the options for ASO sequences that can be designed and tested for efficacy [27]. An exciting potential direction for antibacterial ASO design is to target single-stranded regions of RNA co-transcriptional folding intermediates to inhibit riboswitch activity and degrade the mRNA [82].

5. Advances in Riboswitch Co-transcriptional Folding for Riboswitch Families Found in Pathogenic Bacteria

RNA folding is by nature a co-transcriptional process, since RNA secondary structural elements can form as fast as $\sim 10^{-6}$ s [83]—multiple orders of magnitude faster than the transcription elongation rate by RNAPs, e.g., $\sim 10^{-3}$ s/nt for *E. coli* RNAP [84]. Experimental evidence pointing towards the *in vivo* relevance of co-transcriptional RNA folding emerged from early studies on the *Tetrahymena* group I intron, where it was observed that the autocatalyzed splicing of this intron occurs in a few seconds [85] and thus must take place within the timeframe of transcription. In stark contrast, when the same *Tetrahymena* group I intron RNA is synthesized *in vitro*, denatured and refolded, the majority of the RNA becomes “trapped” in intermediate folds

that rearrange into the native state on a slow timescale (~1–100 min) [86]. Therefore, developing a model of how an RNA acquires its functionally relevant folds requires incorporating the transcription machinery into experimental methods.

Many structural studies of riboswitch aptamer domains follow classic *in vitro* protocols: the RNA is transcribed with T7 RNAP, followed by denaturation and refolding for structural characterization [87,88]. While such work has provided seminal insights into the architecture of riboswitch aptamer domains [89], this does not recapitulate how RNA folds *in vivo*. A key example stems from initial efforts to crystallize the *B. subtilis* *yvrC* cobalamin riboswitch, where the traditional method of denaturing gel purification, refolding, and incubation with a cobalamin ligand did not yield the ligand-bound aptamer for X-ray crystallography [90]. These traditional *in vitro* methods used to study folded RNA differ from how an RNA folds in the cellular environment in a few key ways. First, RNA is transcribed 5′-3′, and local secondary structures can begin forming before the full-length transcript is synthesized. Rather than adhering to a single free-energy landscape wherein the full-length transcript folds all at once, for RNA, the folding free energy landscape changes as the transcript is lengthened by RNAP [91]. Second, the rate of RNAP transcription alters RNA's folded structure by dictating the time window for forming structural motifs. In particular, transcriptional pausing—a temporary halt in nucleotide addition during transcript elongation [92]—creates a temporal window for key folding transitions and subsequent biological activity [93]. Lastly, the experimental conditions for studying RNA folding in a test tube do not recreate the local environment experienced by an RNA in the cell, which is a topic that has been extensively reviewed elsewhere [94].

To better recapitulate how riboswitches fold and respond to a ligand in a native context, more laboratories are incorporating the transcriptional machinery into their experimental methods. Structural approaches such as Selective 2′-Hydroxyl Acylation analyzed by Primer

Extension and Mutational Profiling, or SHAPE-MaP [95], have been updated to incorporate RNAP, permitting the interrogation of riboswitch RNA structures in the context of transcription complexes [96]. Single-molecule approaches such as optical tweezers [97] and single-molecule FRET (smFRET) [98] have been used to study the folding events and dynamics of riboswitch constructs in the context of transcription complexes. Such studies have provided invaluable insights into structural intermediates traversed by riboswitch RNAs during active transcription that illustrate how a riboswitch aptamer domain “talks” to the expression platform [19,99]. In light of these advances, we posit that further consideration of riboswitch folding events during active transcription will present opportunities for designing novel inhibitors. We first review methods that hold significant potential to aid rational ASO design for targeting riboswitches co-transcriptionally: Co-transcriptional SHAPE-seq [100] and, more recently, variable length Transcription Elongation Complex RNA structure probing (TECprobe-VL) [19]. Further, we present specific examples of riboswitch co-transcriptional folding studies with the aim of highlighting the diversity of techniques that have uncovered novel insights into kinetic control of riboswitch folding. Table 5.1 presents all riboswitch classes found in pathogenic bacteria where specific family members have had their co-transcriptional folding characterized, with references listed for further reading. Additionally, the riboswitches we focus on in this section are poised as excellent targets for novel inhibition strategies, since they are all present in pathogenic bacteria and control the expression of essential genes for bacterial survival. Thus, upon repression of gene expression by a riboswitch inhibitor, the bacterial cell will be starved of an essential metabolite and die [12].

Table 5.1 Riboswitch classes in pathogenic bacteria whose co-transcriptional folding has been experimentally characterized.

Riboswitch Class	Occurrence in Human Pathogenic Bacteria	Co-transcriptional Folding Studies	Experimentally Validated Secondary/Tertiary Structures for Folding Intermediates
TPP	Present in 48 species of human pathogenic bacteria [24]	<i>E. coli thiC</i> riboswitch: [101,102], <i>E. coli tbpA</i> (a.k.a. <i>thiB</i>) riboswitch: [103,104]	<i>E. coli thiC</i> riboswitch: [101,102], <i>E. coli tbpA</i> (a.k.a. <i>thiB</i>) riboswitch: [103]
FMN	Present in 41 species of human pathogenic bacteria [24]	<i>B. subtilus ribD</i> riboswitch: [105]	N/A
<i>glmS</i> ribozyme	Present in 26 species of human pathogenic bacteria [24]	<i>B. subtilus glmS</i> ribozyme [106,107]	N/A
Purine	Present in 17 species of human pathogenic bacteria [24]	<i>B. subtilus yxjA</i> and <i>xpt-pbuX</i> riboswitches: [108,109] <i>B. subtilus pbuE</i> and <i>V. vulnificus add</i> riboswitches: [97,110,111]	Guanine: <i>B. subtilus yxjA</i> and <i>xpt-pbuX</i> riboswitches: [108,109] Adenine: N/A
Cobalamin	Present in 36 species of human pathogenic bacteria [24]	<i>E. coli btuB</i> riboswitch: [112]	N/A
c-di-GMP	Present in 7 species of human pathogenic bacteria [113]	<i>C. difficile</i> c-di-GMP riboswitch: [114]	<i>C. difficile</i> c-di-GMP riboswitch: [114]
preQ₁	Present in 15 species of human pathogenic bacteria [113]	<i>B. subtilus preQ₁</i> riboswitch: [115,116]	<i>B. subtilus preQ₁</i> riboswitch: [116]

5.1. Co-transcriptional SHAPE-seq and TECprobe-VL

The Lucks and Strobel laboratories have been pioneers in their development of strategies to model RNA structure in the context of transcription elongation complexes. Initially, Strobel and

colleagues reported on the co-transcriptional SHAPE-seq method [100], which involves the following steps: (1) Preparation of a DNA template encoding the RNA sequence of interest with biotinylated dNMPs randomly distributed throughout the template. Streptavidin binds at the randomly distributed biotinylation sites, creating roadblocks to RNAP movement across every position on the DNA template. (2) Reconstitution of transcription *in vitro*, resulting in the formation of stalled transcription elongation complexes that are exposing the nascent, structured RNA transcript. (3) Chemical probing of the displayed RNA with either a SHAPE reagent (e.g., benzoyl cyanide) or a base-specific reagent (e.g., dimethyl sulfate) followed by RNA processing steps to read out where the modification sites are and inform structure models [100]. More recently, SHAPE-seq has been superseded by the TECprobe-VL method, which improved upon the co-transcriptional SHAPE-seq by simplifying the library preparation steps to make the method more broadly accessible [19]. The introduction of TECprobe-VL by Szyjka and Strobel studied the intermediate folds traversed early on in transcription of the *C. bacterium* ppGpp riboswitch, demonstrating regions of the RNA that are initially single-stranded and ultimately rearrange to be base-paired in the full-length folded riboswitch [19]. While the ppGpp riboswitch itself is not an attractive antibiotic candidate since it does not regulate the expression of an essential gene for bacterial survival [117], the accessibility of TECprobe-VL permits its application to any riboswitch sequence of interest.

5.2. TPP Riboswitches

5.2.1. *E. coli thiC* Riboswitch

The Lafontaine lab applied a suite of complementary biochemical approaches to define how ligand binding to the TPP-sensing *E. coli thiC* riboswitch induces riboswitch structural rearrangements during active transcription [101]. Oligonucleotide hybridization/RNase H

digestion assays were performed to compare ligand-dependent structural changes when TPP is added co-transcriptionally vs. post-transcriptionally. Measurements of ligand binding kinetics from these data demonstrated that co-transcriptional ligand binding is approximately four-fold more efficient compared to post-transcriptional ligand binding [101]. Further, *in vitro* transcription–translation assays elucidated that co-transcriptional TPP binding is required for the riboswitch to fold into a translationally inactive conformation [101]. These observations strongly suggest that ligand sensing and subsequent structural rearrangements are coupled to transcription [101].

This study identified three transcriptional pauses encountered by RNAP during synthesis of the riboswitch expression platform [101]. Motivated by the known importance of transcriptional pausing for RNA folding and gene regulation [92], the researchers explored the structure of nascent RNA transcripts at these pause sites using two complementary *in vitro* techniques that interrogate RNA structure—oligonucleotide hybridization/RNase H digestion assays and SHAPE [118]. These data showed that while the nascent RNA adjacent to the first two pauses was competent to sense TPP (Figure 5.3A), the riboswitch structure at the third pause was no longer able to sense TPP; this loss of sensitivity is due to the nucleation of a hairpin stem that disrupts the folded aptamer domain in favor of an alternative fold (Figure 5.3B) [101]. Taken together, this study defined a key time window for ligand-sensing and subsequent structural rearrangements that are “bookmarked” by transcriptional pauses; therefore, a rationally designed ASO that prevents the disruption of the folded aptamer domain within this window could bias the riboswitch towards the “off” state regardless of the intracellular TPP concentration (Figure 5.3C). Alternatively, an ASO containing DNA bases can anneal to a riboswitch folding intermediate and prompt RNase H degradation (Figure 5.3C).

5.2.2. *E. coli* *tbpA* Riboswitch

In the case of the TPP-sensing *E. coli* *tbpA* riboswitch (a.k.a. *thiB* riboswitch), smFRET was used by the Lafontaine lab to monitor the dynamics of nascent riboswitch transcripts in the context of transcription complexes [104]. The *tbpA* riboswitch has a critical feature in common with the *thiC* riboswitch above, in that there are three transcriptional pauses during expression platform folding. Since prior work had found pausing to set a temporal window for ligand-sensing for the *thiC* riboswitch [101], this smFRET study investigated RNA dynamics at the *tbpA* pauses using roadblocked transcription complexes. Importantly, this study found that while nascent riboswitch transcripts were competent to bind the ligand at the first two pauses, ligand-induced structural changes were no longer observed at the third pause [104]. The analysis of riboswitch structural dynamics in this smFRET setup was expanded to monitor the folding of nascent riboswitch RNAs in real time, providing insight into TPP-induced folding transitions during active transcription [104]. These data reveal that TPP sensing occurs within a precise transcriptional window bound by the nucleation of a hairpin stem that competes with a paired region in the aptamer domain [104]. This insight mirrors the findings highlighted above for the *thiC* riboswitch, demonstrating the generality of an ASO-targeting approach that focuses on nucleating RNA structures in the vicinity of transcriptional pause sites (Figure 5.3C).

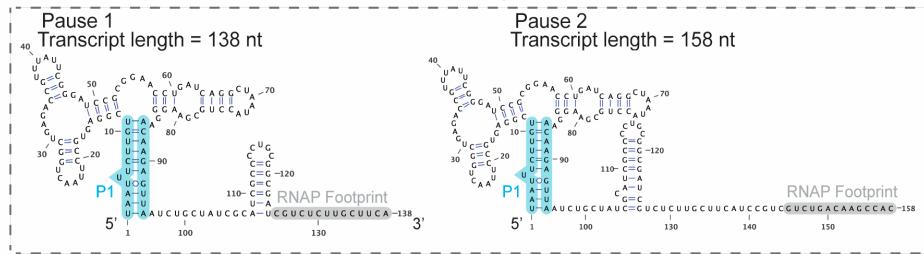
Figure 5.3 Folding pathway for *E. coli thiC* riboswitch in the presence and absence of TPP ligand and its proposed targeting with ASO.

(A) The riboswitch structural intermediates in the vicinity of RNAP pause sites are modeled based on published structure-probing experiments. (B) In the absence of TPP ligand, the Anti-P1 arm invades the already-folded TPP-sensing aptamer domain, promoting a translationally active conformation. In the presence of TPP ligand, TPP stabilizes P1 in the aptamer domain, blocking strand invasion by the Anti-P1 arm and promoting a translationally inactive conformation. (C) Opportunities for ASO inhibition by sequestering the Anti-P1 arm and preventing its invasion of the folded aptamer domain. One strategy is an RNase H-competent ASO, which contains a complementary DNA strand flanked by 2'-OMe nucleotides at its 5' and 3' ends to enhance stability. Alternatively, an ASO can act as a steric block that remains stably base-paired to the riboswitch RNA. Since the ASO remains base-paired, strand invasion by the Anti-P1 arm is blocked, favoring the folding of a hairpin stem that sequesters the RBS. A phosphorodiamidate morpholino oligonucleotide (PMO) is a charge-neutral nucleic acid that is commonly used for steric block ASOs and is described further in Section 7.

A.

Early riboswitch folding intermediates

TPP-sensing window

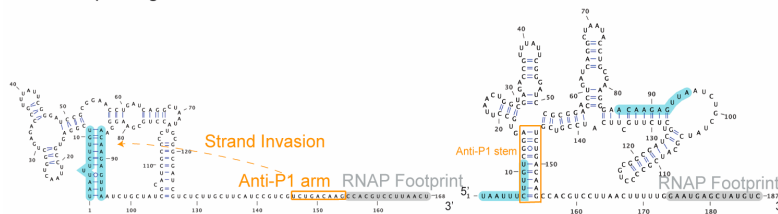


B.

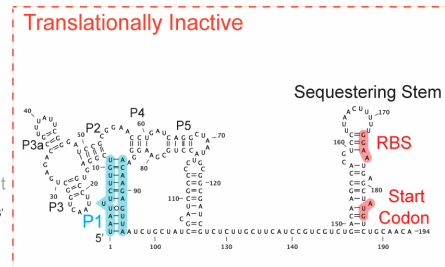
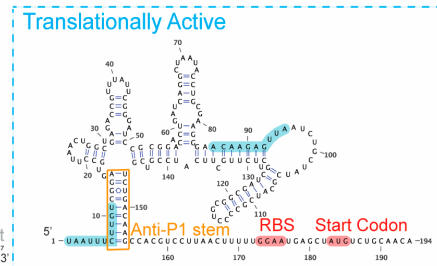
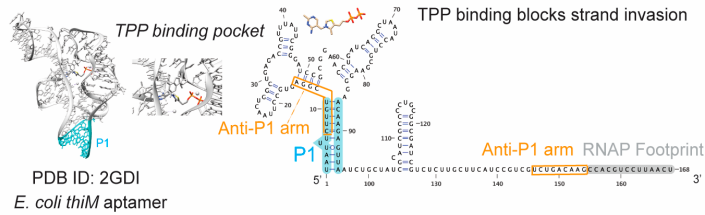
Aptamer disruption (-TPP)

Complementary Anti-P1 arm exits RNAP
Transcript length = 168 nt

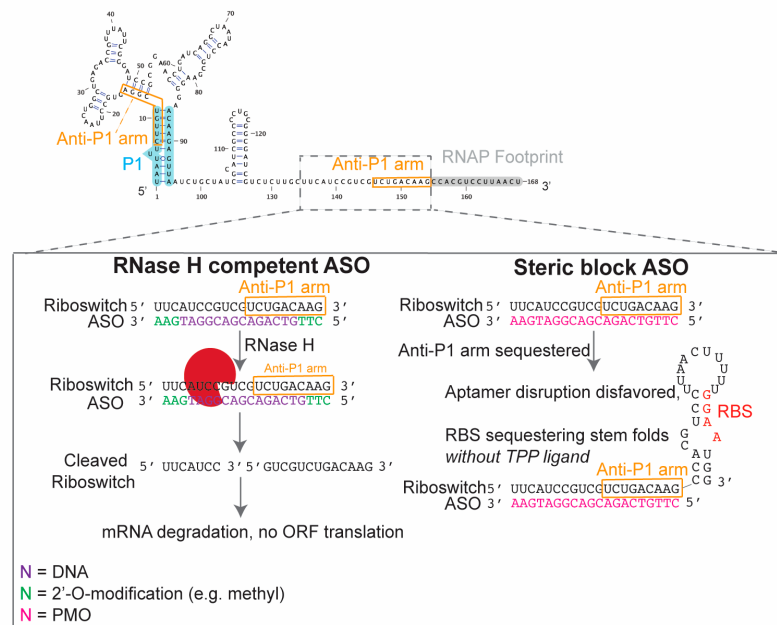
Pause 3
Transcript length = 187 nt



Aptamer stabilized (+TPP)



C.



The Lucks lab reported new insights into intermediate folds traversed by the *E. coli* *tbpA* riboswitch. The co-transcriptional SHAPE-seq approach [100] was applied to nascent *tbpA* transcripts to define key positions during riboswitch folding that commit the RNA structure to either a translationally active or inactive conformation. These data reveal that the linker region between the aptamer domain and expression platform is involved in nucleating a hairpin stem. The nucleation of this hairpin stem requires breaking a key pairing interaction in the aptamer domain, and the completion of the hairpin folding results in RBS exposure [103]. Unlike the smFRET study highlighted above, the co-transcriptional SHAPE-seq experiments performed in this work map intermediate RNA folds to specific riboswitch transcript lengths. Therefore, analysis of the transcript lengths where competing RNA structures begin to fold will define specific sequence targets that, upon annealing of an ASO, could bias RNA folding towards the translationally inactive conformation, starving the bacterial cell of an essential cofactor. We present secondary structure models of these experimentally validated *tbpA* riboswitch folding intermediates in Figure 5.4 and highlight how these intermediates could be targeted by an ASO.

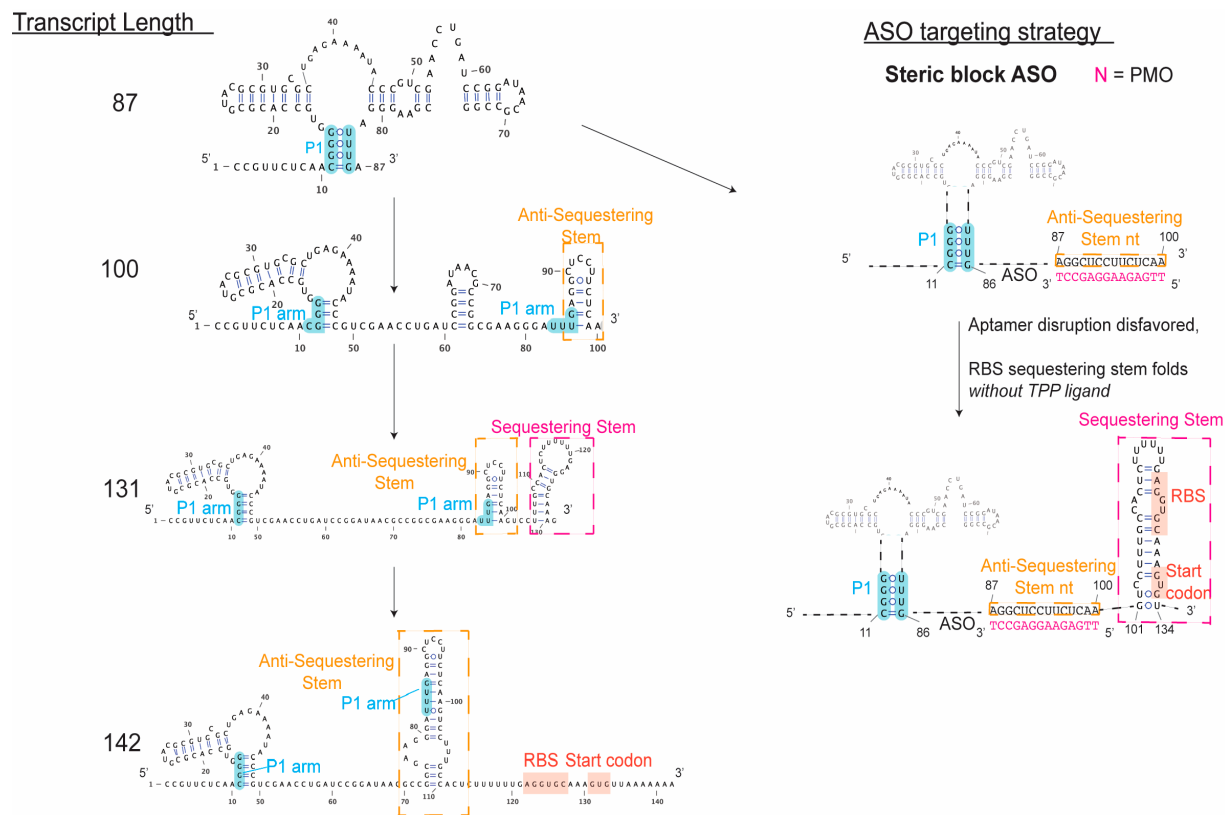


Figure 5.4 *E. coli* *tbpA* riboswitch folding intermediates experimentally validated by co-transcriptional SHAPE-seq.

A steric block ASO (e.g., PMO) can prevent formation of the anti-sequestering stem and favor instead folding of the sequestering stem, leaving RBS occluded in the absence of TPP ligand.

5.3. FMN Riboswitches

B. subtilis *ribD* Riboswitch

The Breaker laboratory clearly demonstrated that FMN-sensing by the *ribD* riboswitch in *B. subtilis* is kinetically driven, wherein ligand binding must occur within a precise time window to prompt transcription termination [105]. Specifically, transcriptional pauses were found to facilitate coordination between ligand binding and expression platform folding [105], similar to the TPP riboswitch examples highlighted above. Further, this study ascertained that FMN never reaches thermodynamic equilibrium with the riboswitch aptamer due to the kinetic constraints on

the folding regime [105], highlighting the importance of transcription kinetics for effective riboswitching and downstream gene regulation.

Notably, further efforts to identify FMN riboswitch structural intermediates have not yet been reported. Approaches described earlier in this section such as co-transcriptional SHAPE-seq [100], and more recently TECprobe-VL [19], will be useful strategies to precisely characterize intermediate RNA structures in the vicinity of transcriptional pause sites and provide new targets for ASO inhibitor design.

5.4. *glmS* Ribozyme

B. subtilis glmS ribozyme

Recently, the Woodson lab explored whether the *glmS* ribozyme executes its self-cleavage co-transcriptionally [107]. Prior to the publication of this study, few key datapoints suggested that co-transcriptional self-cleavage was possible: (1) the ribozyme self-cleavage reaction is fast (80 min^{-1}) [119], so it is a reasonable assumption that cleavage occurs before the 1803-nt *glmS* open reading frame (ORF) is transcribed; (2) co-transcriptional cleavage permits RNase J1 to promptly begin mRNA degradation and prevent ribosomes from initiating translation of the *glmS* mRNA. If cleavage were to occur after ORF transcription, initiating ribosomes may impede RNase J1 degradation [120–122].

Motivated by the above observations, the Woodson lab employed a combination of single-molecule fluorescence microscopy and biochemical methods to analyze co-transcriptional ribozyme cleavage [107]. First, *in vitro* transcription reactions with radiolabeled RNAs showed that the ribozyme was able to self-cleave during transcription [107]. However, this bulk ensemble approach could not define the sequence window where co-transcriptional cleavage began. To address this, a single-molecule TIRF (smTIRF) microscopy assay was developed to monitor

transcription and ribozyme activity in real time. These data were fit to a quantitative model, which ultimately demonstrated that the nascent ribozyme can initiate cleavage before completion of its synthesis [107]. Armed with the knowledge that the *glmS* ribozyme induces self-cleavage co-transcriptionally, the modeling of structural intermediates traversed by this ribozyme could identify sequence regions amenable to ASO targeting.

5.5. Purine Riboswitches

5.5.1. Adenine Riboswitches

The Block lab employed an optical-trapping assay to monitor co-transcriptional folding events for the *B. subtilis pbuE* adenine-sensing riboswitch [97]. These assays clearly demonstrated that the adenine-induced formation of an anti-terminator stem occurs co-transcriptionally [97], highlighting a recurring theme for riboswitches that control transcription termination/anti-termination: the RNA structural switch must occur within the time-frame of transcription [123]. Once an anti-terminator stem is formed, RNAP continues transcribing the mRNA, and any subsequent RNA structural rearrangements in the 5'-UTR can no longer exert control over transcription.

More recent work from the Batey laboratory employed a structure-guided mutagenesis approach to characterize specific regions of the *pbuE* riboswitch that confer ligand-dependent regulatory activity [110]. First, it was observed that a hairpin stem in the aptamer domain facilitates strand invasion by the expression platform to promote the formation of an intrinsic terminator hairpin [110]. Second, the introduction of a strong transcriptional pause towards the end of aptamer domain synthesis produced the greatest induction of reporter gene expression of all the riboswitch variants tested in this study [110]. This finding underscores the importance of pause sites for effective riboswitching, and provides new clues for the rational design of ASO

inhibitors. For example, since pauses provide a temporal window for RNA structural elements to form [93], transient single-stranded regions of RNA in the vicinity of a pause site could be targeted by an ASO inhibitor, as illustrated for the TPP-sensing riboswitches described in Section 5.2 (Figures 5.3 and 5.4).

5.5.2 Guanine Riboswitches

The Lucks lab investigated the co-transcriptional folding mechanism of the *B. subtilis* *yxjA* guanine-sensing riboswitch, which is an example of a riboswitch where there are no overlapping sequences between the aptamer domain and expression platform [108], unlike the above *pbuE* adenine-sensing riboswitch example. Interestingly, the riboswitch relies on the formation of a central helix that dictates the switching mechanism depending on whether guanine is bound to the aptamer domain (Figure 5.5A) [108]. Binding of guanine effectively prevents this central helix from folding, biasing the downstream riboswitch folding toward the “off” state [108]. The use of co-transcriptional SHAPE-seq defined the precise transcriptional window where this key central helix forms (Figure 5.5B) and led the authors to the hypothesis that this central helix forms through a strand-invasion process that is disfavored in the presence of ligand [108]. This hypothesis was further tested using gene reporter fusion assays and MD simulations, supporting a model wherein the formation of this central helix competes with a stable pairing interaction in the aptamer domain [108]. These findings underscore the importance of considering a riboswitch’s intermediate folds in efforts to bias its structural ensemble to the “off” state using novel inhibitors. Since this work on the *yxjA* riboswitch has experimentally validated folding intermediates, we present some of the secondary structure models in Figure 5.5B,C to highlight how these intermediates could be targeted by an ASO. Specifically, we propose that a steric block ASO designed to mimic the central helix could promote the folding of the downstream terminator hairpin in the absence of guanine ligand.

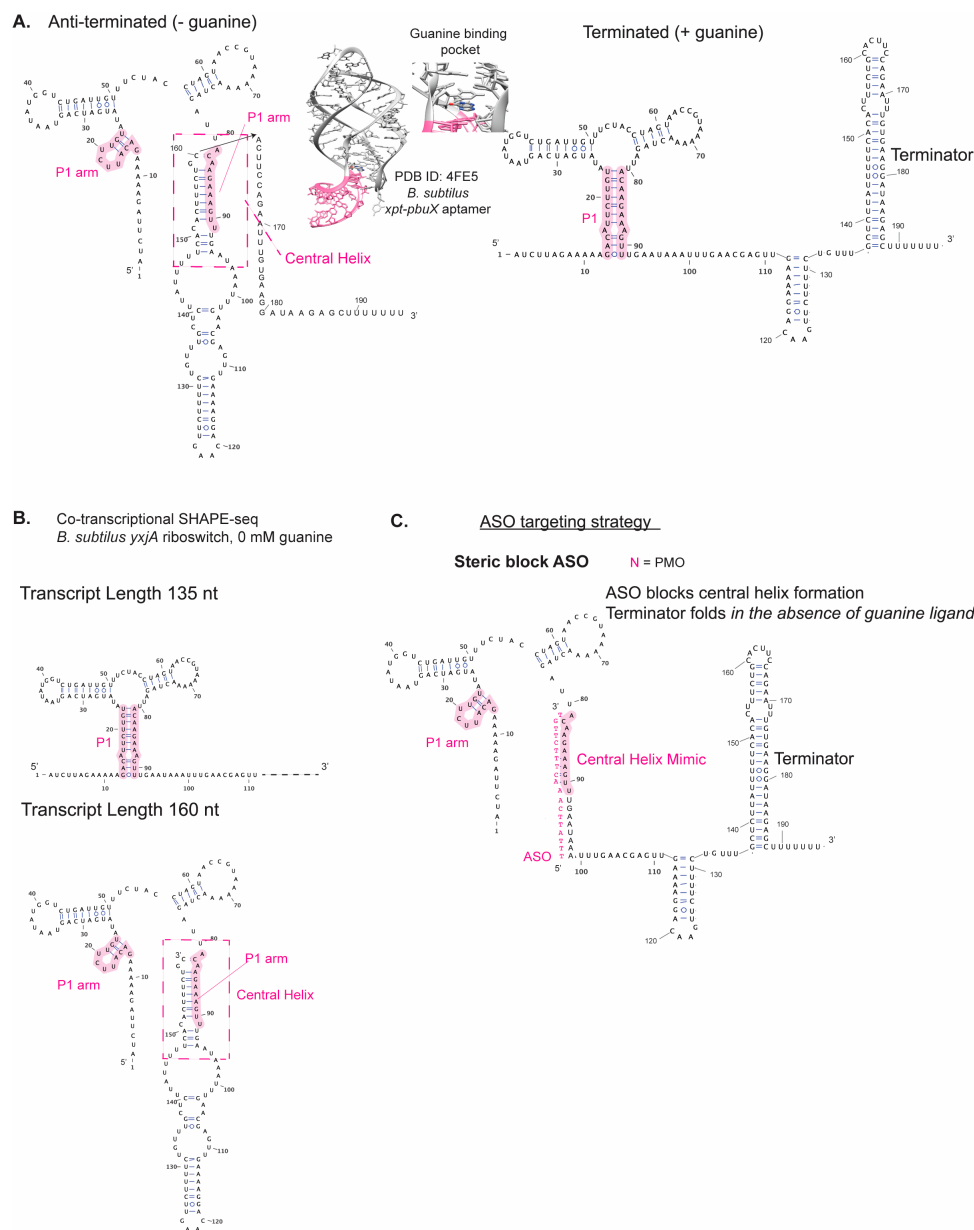


Figure 5.5 Folding outcomes for *B. subtilis* yxjA riboswitch in the presence and absence of guanine ligand and its proposed targeting with ASO.

(A) Alternative structures of the *B. subtilis* yxjA riboswitch in the presence and absence of guanine ligand. (B) Riboswitch structural intermediates based on co-transcriptional SHAPE-seq experiments. (C) ASO targeting strategy to mimic the central helix, which can promote the folding of the downstream terminator in the absence of guanine ligand.

5.6. Cobalamin, c-di-GMP, and preQ₁ Riboswitches

While less widely distributed in pathogenic bacteria, the cobalamin, c-di-GMP, and preQ₁ riboswitches present an enticing opportunity for narrow-spectrum inhibitors that target specific bacterial strains. The co-transcriptional folding of the *E. coli btub* cobalamin-sensing riboswitch was first characterized by the Pan lab in a study that used oligonucleotide hybridization/RNase H digestion assays. First off, comparison of cobalamin-dependent structural changes between denatured and refolded and co-transcriptionally folded RNA elucidated that ligand-dependent structural rearrangements occur during active transcription [112]. Further, transcriptional pausing was found to facilitate cobalamin-dependent structural rearrangements in the expression platform [112]. While the oligonucleotide hybridization/RNase H digestion assays performed in this work were not intended to be applied to riboswitch inhibitor design, a revisitation of these biochemical studies could inspire initial efforts towards ASO design. For example, since this study employed 57 different oligo probes, a fresh look at these data could provide useful starting points for ASOs that can anneal to the nascent riboswitch RNA and prompt degradation during active transcription.

The Schwalbe lab has employed NMR to study folding intermediates of the *C. difficile* c-di-GMP riboswitch. Using a method to rapidly produce and purify a library of intermediate RNA transcripts [124], this study measured K_d values for riboswitch folding intermediates with both NMR and ITC [114]. These data defined a 61-nt window for c-di-GMP sensing by the riboswitch, and the K_d values obtained for the riboswitch transcripts within this window spanned two orders of magnitude [114]. The wide range of K_d values was explained by the presence of competing riboswitch folds present at specific transcript lengths, altering the riboswitch structural ensemble and in turn ligand binding [114].

The Walter lab has employed both smFRET and cryo-EM to investigate the folding and dynamics of the *B. subtilis* preQ₁ riboswitch. Investigation of the preQ₁ riboswitch with smFRET uncovered that a transcriptional pause situated between the aptamer domain and expression platform is stabilized by a pseudoknot that directly contacts RNAP [115]. Further, it was observed that binding of the preQ₁ ligand signals RNAP release from the pause, underscoring the coupling between riboswitch folding and transcriptional pausing [115]. Elaborating on this work, cryo-EM structures of paused elongation complexes containing the nascent preQ₁ riboswitch RNA were solved in both the ligand-free and ligand-bound states [116]. This work revealed that in the absence of preQ₁, the local riboswitch RNA structure “retracts” the RNA 3′ end within the RNAP active site, leading to a halt in nucleotide incorporation [116]. However, preQ₁ binding to the riboswitch induces large-scale conformational changes in RNAP that reorient the RNA 3′ end at the active site and release the pause [116].

5.7. Other Advances in Riboswitch Co-Transcriptional Folding

In this section, we summarize additional riboswitch co-transcriptional folding studies that point to recurring themes in riboswitch-mediated gene regulation, which will likely prove useful for the design of ASO inhibitors.

There has been significant research effort directed towards the co-transcriptional folding of the ZTP riboswitch class, which is a signaling molecule produced in response to 10-formyltetrahydrofolate deficiency in the bacterial cell [125]. The Hua lab designed a single-molecule vectorial folding assay using an engineered superhelicase that sequentially releases fluorescently labeled riboswitch RNA at a speed comparable to that of bacterial RNAPs (~60 nt/s) [126]. This technique was applied to the *F. ulcerans* ZTP riboswitch, which elucidated that the riboswitch is kinetically controlled, since the rate of RNA release by the helicase altered

ligand-dependent structural changes [126]. The folding of the *C. beijerinckii pfl* ZTP riboswitch has been extensively studied by the Lucks and Strobel laboratories using co-transcriptional SHAPE-seq, and more recently TECprobe-VL [19,99]. These studies uncovered the specific sequence and structural elements that modulate an internal RNA strand displacement, which is disfavored in the presence of a ligand [19,99].

Building upon this work on the *C. beijerinckii pfl* ZTP riboswitch, the Lucks lab discovered that altering the rate of strand displacement via mutagenesis tunes its dynamic range [127], and that altering the expression platform sequence and structure to slow RNA folding enhances riboswitch sensitivity [128]. Additionally, the Schwalbe lab has investigated the *T. carboxydvorans* ZTP riboswitch co-transcriptional folding with NMR to define the precise window of transcript lengths where the ligand can bind [129].

Another well-studied class is the fluoride-sensing riboswitch, which controls expression of a fluoride exporter to prevent fluoride from accumulating to toxic levels in the cytosol [130]. The co-transcriptional folding of the *B. cereus crcB* fluoride-sensing riboswitch has been characterized by the Walter lab using smFRET. These studies highlighted a dynamic interplay between the transcription factor NusA and the fluoride binding the co-transcriptionally folded RNA [131]. A further smFRET study on this riboswitch by the Walter lab characterized the dynamics of intermediate transcript lengths and established the role of Mg^{2+} in fluoride-mediated conformational changes [132]. Another study on this riboswitch from the Lucks lab employed a combination of *in vitro* transcription and cellular gene expression assays to demonstrate the role of pseudoknot base pairs on co-transcriptional structure switching [133]. There are additional recent works on riboswitch co-transcriptional folding for the manganese sensing riboswitch class [134], 2'-dG sensing riboswitch class [135,136], and the SAM-VI riboswitch class [137], which we direct the reader to for further reading on this topic.

6. Antisense Oligonucleotides That Act on RNA Co-Transcriptionally: A Case Study on a Group I Intron

Building off of prior work from the Turner laboratory [138], the Zhang laboratory reported the use of an ASO to target co-transcriptional self-splicing of the group I intron located within the 26S rRNA of *C. albicans*, a fungus [82]. First, *in vitro* transcription assays with radiolabeled RNA were used to monitor the efficacy of different ASO sequences to inhibit co-transcriptional splicing [82]. It was found that eight different ASOs targeting the P3-P7 core of the intron (Figure 5.6A) achieved half-maximal inhibition in the $\sim 0.2\text{--}0.3\ \mu\text{M}$ concentration range [82]. The addition of a 2'-O-Me moiety to an ASO targeting the P3-P7 core (2'-O-Me 240-261R) enhanced splicing inhibition *in vitro* by an order of magnitude, exhibiting an IC_{50} of $\sim 37\ \text{nM}$ [82]. Prior work established that modifications at the ribose 2' position enhance ASO annealing to its target by promoting a 3'-*endo* pucker conformation of the ribose, thus making it more RNA-like and increasing its propensity to base-pair with the RNA target [139,140]. Since the 2'-O-Me 240-261R ASO delivered promising results *in vitro*, it was used for the subsequent *in vivo* experiments described here. Confocal fluorescence microscopy confirmed the presence of the core-targeted ASO in the nucleus of *C. albicans*, and reverse-transcription PCR (RT-PCR) targeting the pre-rRNA confirmed ASO inhibition *in vivo*, with an IC_{50} of $2.30\ \mu\text{M}$ [82]. The authors proposed that splicing inhibition would be lethal to *C. albicans* because if the intron-containing RNA is assembled into the ribosome, association with the small ribosomal subunit is blocked (Figure 5.6B) [141]. The authors demonstrated that the core-targeted ASO specifically inhibits the growth of an intron-containing *C. albicans* and was unable to inhibit growth for a *C. albicans* strain lacking the intron [82]. Further, the ASO was found to inhibit *C. albicans* growth in a dose-dependent manner, with an IC_{50} of $4.7\ \mu\text{M}$ [82].

Taken together, these data illustrate that ASOs can be rationally designed to target RNA co-transcriptional folding. For the group I intron example highlighted above, the best-performing

ASO for splicing inhibition targets a region of the RNA that is base-paired in its final folded structure, and therefore would not have been an obvious target resulting from the inspection of the fully folded intron (Figure 5.6A). Thus, assessing the final-folded RNA structure alone limits opportunities for the design of efficacious ASO inhibitors. In the remaining sections of this review, we cover advances made in the delivery of oligonucleotide therapeutics and progress in the use of ASOs in the clinic as antibacterials.

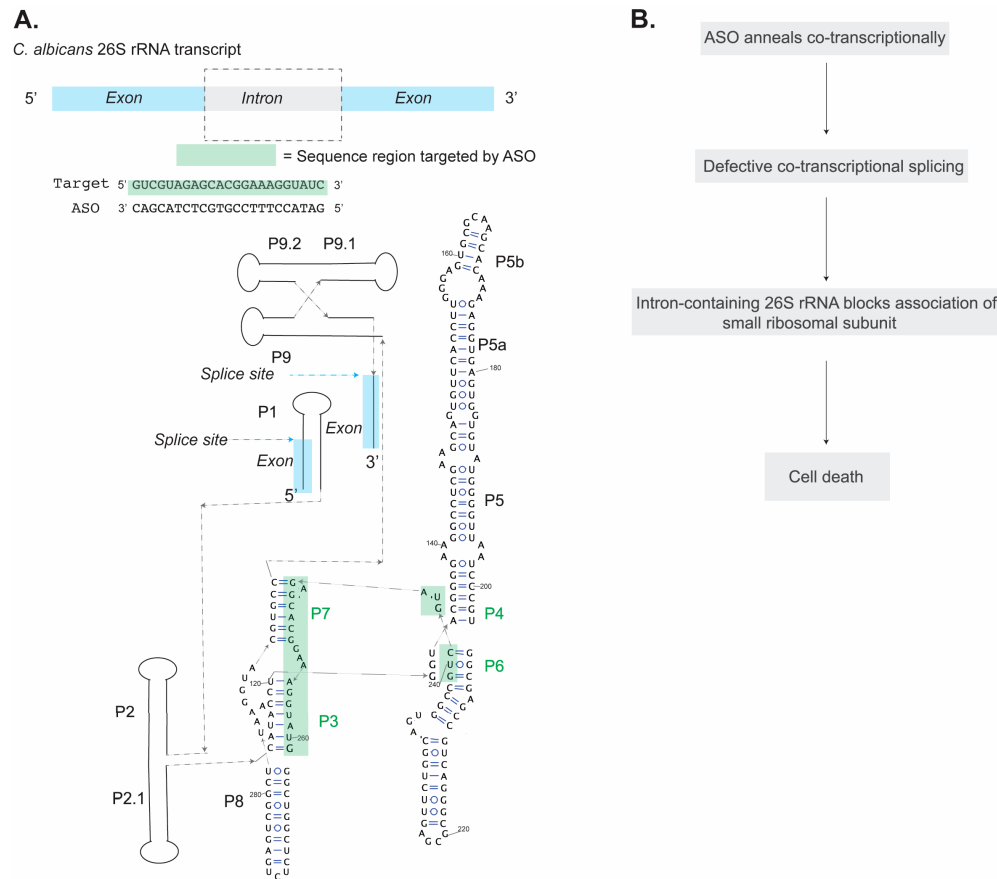


Figure 5.6 ASO targeting of *C. albicans* 26S rRNA group I intron

(A) Secondary structure of *C. albicans* 26S rRNA group I intron, with paired regions of the intron core targeted by the ASO labeled in green. The specific stretch of sequence targeted by the ASO is highlighted in green. Splice sites are denoted with blue arrows. **(B)** Illustration of how the ASO targeting co-transcriptional splicing of the 26S rRNA is lethal to *C. albicans*.

7. Delivery of ASOs

7.1. Chemical Modifications

Chemical modifications have proven to be one of the most effective strategies for enhancing oligonucleotide drug delivery. Modifications of the nucleic acid backbone, ribose sugar and nucleobase have been used extensively in the design of therapeutic ASOs [142]. Such modifications help improve the pharmacokinetics, pharmacodynamics, and biodistribution of ASOs [142]. A commonly employed backbone modification is phosphorothioate linkages (PS), where one of the non-bridging oxygen atoms of the inter-nucleotide phosphate group is replaced with sulfur [142]. PS linkages improve an ASO's therapeutic efficacy by providing nuclease resistance and increasing its association with proteins in plasma and within cells [142]. Another common ASO modification is the addition of a functional group at the 2' position of a ribose sugar, with common examples being 2'-O-methyl, 2'-O-methoxyethyl and 2'-fluoro [142]. These modifications increase resistance to nucleases by substituting the nucleophilic 2'-hydroxyl group, which improves oligonucleotide stability in plasma and increases its half-life in tissues, resulting in extended drug effects [142]. Lastly, rather than using a DNA or RNA-like oligonucleotide, alternative chemistries have been developed, such as phosphorodiamidate morpholino oligonucleotides (PMOs), which are charge-neutral nucleic acids where the five-membered ribose is replaced with a six-membered morpholine ring [142]. The advantage of charge-neutral oligonucleotide chemistries is that they facilitate conjugation to moieties that promote cellular delivery [142].

7.1.2 Bioconjugation Strategies

One strategy to deliver ASOs into bacteria is the use of cell-penetrating peptides (CPPs), which are short (generally <30 amino acids) amphipathic or cationic peptide fragments that originate from protein translocation motifs found in nature, or are based on polymers of basic

amino acids [143]. Peptide–PMO (PPMO) conjugates consisting of a CPP covalently linked to a PMO have been identified as a promising antibacterial delivery strategy [142]. For example, PPMO conjugates that target an essential bacterial gene *acpP* in a murine infection model were demonstrated to be bactericidal [144]. A comprehensive list of PPMO conjugates that have been experimentally validated to cross the cell wall across diverse pathogenic bacteria is presented in a review by Xue *et al.* [145]. Additional strategies that are not based on the covalent modification of the ASO to the delivery system have attracted interest as alternatives, such as encapsulation in lipid nanoparticles or DNA nanostructures, and are also covered in the review by Xue *et al.* [145].

8. Translation of ASO Therapies to the Clinic

There has been significant progress in the translation of oligonucleotide therapies to the clinic, with thirteen ASO-based medicines to treat human diseases approved by the FDA thus far for clinical use [75]. For example, Golodirsen and Eteplirsen are FDA- approved ASOs used to treat Duchenne muscular dystrophy (DMD), which promote the skipping of an exon in the *DMD* gene that would otherwise produce an out-of-frame transcript [142]. Generally speaking, ASOs used to treat human disease can operate through many different mechanisms. For example, some are designed to silence disease-causing or disease-modifying genes by promoting RNase H degradation upon DNA oligo annealing to the target mRNA [142]. Other ASOs are designed to mask a particular stretch of sequence in an mRNA and interfere with downstream RNA–RNA and/or RNA–protein interactions, with a common application being the modulation of alternative splicing, as described for Golodirsen and Eteplirsen [142].

While there are no FDA-approved antisense antibacterials yet, there have been great strides in their development. Multiple preclinical studies delivered promising results on the use of ASOs as potential antibacterial therapeutics [146]. For example, an ASO targeting a gene in

S. aureus that is critical for replication (*ftsZ*) inhibited the growth of methicillin-resistant *S. aureus* in cell culture, and conferred increased survival to mice infected with the Mu50 strain of *S. aureus* [147]. Additionally, in an *S. aureus*-infected cutaneous mouse wound, the use of an ASO targeting the essential *gyrA* mRNA decreased bacterial growth and promoted cutaneous wound healing in the mice [148]. Additional efforts to develop ASOs targeting essential genes in pathogenic bacteria are highlighted in the review by Buthelezi *et al.* [146].

Despite significant advances made in the design of ASO therapeutics, there are many challenges to overcome to achieve their widespread use in the clinic. The effective delivery of ASOs to many tissues is a major hurdle to overcome [142]. For example, injected nucleic acid therapeutics must resist nuclease degradation, bypass renal clearance, and avoid non-productive sequestration by plasma proteins in order to reach the desired target [142]. A combination of chemical modifications and conjugation strategies outlined in Section 7 is typically employed to confer the desired pharmacokinetic and pharmacodynamic properties to an ASO drug [142], and further advances on these fronts will likely expand the number of conditions that are treated with ASO therapeutics.

9. Discussion

Widespread antibiotic resistance in pathogenic bacteria is an urgent problem in modern medicine [1]. To combat this problem, riboswitches have emerged as potential targets for new antibiotics, since they are widespread among pathogenic bacteria but absent in human cells [4]. There have been significant advances in designing small-molecule riboswitch inhibitors, but these strategies still show significant limitations [29,37,49]. Antisense oligonucleotides are beginning to attract attention as a new riboswitch inhibition strategy, and several have been designed to specifically target riboswitches while being non-toxic to human cells [27]. However, there are limited studies leveraging advances in RNA co-transcriptional folding to design ASO

riboswitch inhibitors. The group I intron example (Section 6) demonstrated that regions of RNA that are base-paired in the final folded structure can still be effectively targeted by an ASO. Advances in methodologies to model riboswitch folding intermediates [19,100] provide exciting opportunities to expand the search for efficacious ASO inhibitors.

Experimental work on the riboswitches outlined in Section 5 has built foundational principles for riboswitch co-transcriptional folding that can guide the search for novel ASO inhibitors. For example, co-transcriptional SHAPE-seq studies on the *tbpA* riboswitch illustrated how in the absence of the TPP ligand, a strand invasion event disrupts the riboswitch aptamer domain, leading to the riboswitch “on” state. We illustrate how a competitor ASO could prevent effective strand displacement of the riboswitch, biasing the structural ensemble towards the “off” state even in the absence of TPP ligand (Figures 5.2 and 5.3). Additionally, biochemical study of TPP riboswitches can provide inspiration for novel ASO design, with an example being ASOs that target key RNA structural transitions in the vicinity of transcriptional pause sites during expression platform folding (Figure 5.2).

Additionally, there have been exciting advances in targeting dsRNA–ssRNA junctions using ASOs, which can form a triplex adjacent to a duplex at a dsRNA–ssRNA junction [149]. Recent work illustrated how targeting these junctions with an ASO can stimulate ribosomal frameshifting and inhibit Dicer cleavage activity of a pre-miRNA [149]. This work represents an important advance that could inspire the design of ASOs that target riboswitches at dsRNA–ssRNA junctions, further expanding the available sequence space for inhibitor design.

The studies referenced in Sections 7 and 8 demonstrate how ASOs can be modified and/or conjugated to facilitate efficient delivery into the cell [142,143,145]. In fact, thirteen ASO-based drugs are already FDA-approved for clinical use, with more in various stages of clinical

trial [75]. Using ASOs to treat infectious diseases, including bacterial infections, is an active area of research with several promising candidates already developed [146]. Taken together, we propose that the methods used to study RNA co-transcriptional folding could be key for designing novel ASOs to target riboswitch RNAs in pathogenic bacteria.

Author Contributions: Conceptualization, C.S. and D.P.; writing—original draft preparation, C.S. and D.P.; writing—review and editing, C.S., D.P. and T.V.M.; visualization, C.S. and D.P.; supervision, T.V.M.; funding acquisition, T.V.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Institutes of Health (NIGMS ESI grant R35GM142785), Burroughs Wellcome Fund CASI (grant no.: 1016945) to T.V.M., UCSD institutional support, and Yinan Wang Memorial Chancellor’s Endowed Junior Faculty Fellowship to T.V.M. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-2038238 to D.P. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Conflicts of Interest: The authors declare no conflicts of interest.

10. References

1. Prestinaci, F.; Pezzotti, P.; Pantosti, A. Antimicrobial Resistance: A Global Multifaceted Phenomenon. *Pathog. Glob. Health* **2015**, *109*, 309–318. <https://doi.org/10.1179/2047773215Y.00000000030>.
2. Brown, E.D.; Wright, G.D. Antibacterial Drug Discovery in the Resistance Era. *Nature* **2016**, *529*, 336–343. <https://doi.org/10.1038/nature17042>.
3. Tacconelli, E.; Carrara, E.; Savoldi, A.; Harbarth, S.; Mendelson, M.; Monnet, D.L.; Pulcini, C.; Kahlmeter, G.; Kluytmans, J.; Carmeli, Y.; Ouellette, M.; Outtersson, K.; Patel, J.; Cavaleri, M.; Cox, E.M.; Houchens, C.R.; Grayson, M.L.; Hansen, P.; Singh, N.; Theuretzbacher, U.; Magrini, N.; WHO Pathogens Priority List Working Group. Discovery, Research, and Development of New Antibiotics: The WHO Priority List of Antibiotic-Resistant Bacteria and Tuberculosis. *Lancet Infect. Dis.* **2018**, *18*, 318–327. [https://doi.org/10.1016/S1473-3099\(17\)30753-3](https://doi.org/10.1016/S1473-3099(17)30753-3).
4. Ellinger, E.; Chauvier, A.; Romero, R.A.; Liu, Y.; Ray, S.; Walter, N.G. Riboswitches as Therapeutic Targets: Promise of a New Era of Antibiotics. *Expert Opin. Ther. Targets* **2023**, *27*, 433–445. <https://doi.org/10.1080/14728222.2023.2230363>.
5. Mandal, M.; Breaker, R.R. Gene Regulation by Riboswitches. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 451–463. <https://doi.org/10.1038/nrm1403>.
6. Richards, J.; Belasco, J.G. Riboswitch Control of Bacterial RNA Stability. *Mol. Microbiol.* **2021**, *116*, 361–365. <https://doi.org/10.1111/mmi.14723>.
7. Sherwood, A.V.; Henkin, T.M. Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. *Annu. Rev. Microbiol.* **2016**, *70*, 361–374. <https://doi.org/10.1146/annurev-micro-091014-104306>.
8. Pavlova, N.; Kaloudas, D.; Penchovsky, R. Riboswitch Distribution, Structure, and Function in Bacteria. *Gene* **2019**, *708*, 38–48. <https://doi.org/10.1016/j.gene.2019.05.036>.
9. Winkler, W.C.; Breaker, R.R. Regulation of Bacterial Gene Expression by Riboswitches. *Annu. Rev. Microbiol.* **2005**, *59*, 487–517. <https://doi.org/10.1146/annurev.micro.59.030804.121336>.
10. Kavita, K.; Breaker, R.R. Discovering Riboswitches: The Past and the Future. *Trends Biochem. Sci.* **2022**, *48*, 119–141. <https://doi.org/10.1016/j.tibs.2022.08.009>.
11. Tucker, B.J.; Breaker, R.R. Riboswitches as Versatile Gene Control Elements. *Curr. Opin. Struct. Biol.* **2005**, *15*, 342–348. <https://doi.org/10.1016/j.sbi.2005.05.003>.
12. Blount, K.F.; Breaker, R.R. Riboswitches as Antibacterial Drug Targets. *Nat. Biotechnol.* **2006**, *24*, 1558–1564. <https://doi.org/10.1038/nbt1268>.
13. Zhang, R.; Ou, H.-Y.; Zhang, C.-T. DEG: A Database of Essential Genes. *Nucleic Acids Res.* **2004**, *32*, D271–D272. <https://doi.org/10.1093/nar/gkh024>.

14. Panchal, V.; Brenk, R. Riboswitches as Drug Targets for Antibiotics. *Antibiotics* **2021**, *10*, 45. <https://doi.org/10.3390/antibiotics10010045>.
15. Giarimoglou, N.; Kouvela, A.; Maniatis, A.; Papakyriakou, A.; Zhang, J.; Stamatopoulou, V.; Stathopoulos, C. A Riboswitch-Driven Era of New Antibacterials. *Antibiotics* **2022**, *11*, 1243. <https://doi.org/10.3390/antibiotics11091243>.
16. Childs-Disney, J.L.; Yang, X.; Gibaut, Q.M.R.; Tong, Y.; Batey, R.T.; Disney, M.D. Targeting RNA Structures with Small Molecules. *Nat. Rev. Drug Discov.* **2022**, *21*, 736–762. <https://doi.org/10.1038/s41573-022-00521-4>.
17. Trausch, J.J.; Batey, R.T. A Disconnect between High-Affinity Binding and Efficient Regulation by Antifolates and Purines in the Tetrahydrofolate Riboswitch. *Chem. Biol.* **2014**, *21*, 205–216. <https://doi.org/10.1016/j.chembiol.2013.11.012>.
18. Kilanowska, A.; Studzińska, S. *In Vivo* and *in Vitro* Studies of Antisense Oligonucleotides—A Review. *RSC Adv.* **2020**, *10*, 34501–34516. <https://doi.org/10.1039/D0RA04978F>.
19. Szyjka, C.E.; Strobel, E.J. Observation of Coordinated RNA Folding Events by Systematic Cotranscriptional RNA Structure Probing. *Nat. Commun.* **2023**, *14*, 7839. <https://doi.org/10.1038/s41467-023-43395-9>.
20. Yan, L.-H.; Le Roux, A.; Boyapelly, K.; Lamontagne, A.-M.; Archambault, M.-A.; Picard-Jean, F.; Lalonde-Seguin, D.; St-Pierre, E.; Najmanovich, R.J.; Fortier, L.-C.; Lafontaine, D.; Marsault, E. Purine Analogs Targeting the Guanine Riboswitch as Potential Antibiotics against *Clostridioides Difficile*. *Eur. J. Med. Chem.* **2018**, *143*, 755–768. <https://doi.org/10.1016/j.ejmech.2017.11.079>.
21. Traykovska, M.; Otcheva, L.A.; Penchovsky, R. Targeting TPP Riboswitches Using Chimeric Antisense Oligonucleotide Technology for Antibacterial Drug Development. *ACS Appl. Bio Mater.* **2022**, *5*, 4896–4902. <https://doi.org/10.1021/acsabm.2c00628>.
22. Begley, T.P.; Downs, D.M.; Ealick, S.E.; McLafferty, F.W.; Van Loon, A.P.; Taylor, S.; Campobasso, N.; Chiu, H.-J.; Kinsland, C.; Reddick, J.J.; Xi, J. Thiamin Biosynthesis in Prokaryotes. *Arch. Microbiol.* **1999**, *171*, 293–300. <https://doi.org/10.1007/s002030050713>.
23. Rodionov, D.A.; Vitreschak, A.G.; Mironov, A.A.; Gelfand, M.S. Comparative Genomics of Thiamin Biosynthesis in Prokaryotes: New Genes and Regulatory Mechanisms. *J. Biol. Chem.* **2002**, *277*, 48949–48959. <https://doi.org/10.1074/jbc.M208965200>.
24. Pavlova, N.; Penchovsky, R. Genome-Wide Bioinformatics Analysis of FMN, SAM-I, *glmS*, TPP, Lysine, Purine, Cobalamin, and SAH Riboswitches for Their Applications as Allosteric Antibacterial Drug Targets in Human Pathogenic Bacteria. *Expert Opin. Ther. Targets* **2019**, *23*, 631–643. <https://doi.org/10.1080/14728222.2019.1618274>.
25. Miranda-Ríos, J.; Navarro, M.; Soberón, M. A Conserved RNA Structure (Thi Box) Is Involved in Regulation of Thiamin Biosynthetic Gene Expression in Bacteria. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9736–9741. <https://doi.org/10.1073/pnas.161168098>.

26. 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* **2013**, *110*, 4188–4193. <https://doi.org/10.1073/pnas.1218062110>.
27. Pavlova, N.; Traykovska, M.; Penchovsky, R. Targeting FMN, TPP, SAM-I, and *glmS* Riboswitches with Chimeric Antisense Oligonucleotides for Completely Rational Antibacterial Drug Development. *Antibiotics* **2023**, *12*, 1607. <https://doi.org/10.3390/antibiotics12111607>.
28. Winkler, W.; Nahvi, A.; Breaker, R.R. Thiamine Derivatives Bind Messenger RNAs Directly to Regulate Bacterial Gene Expression. *Nature* **2002**, *419*, 952–956. <https://doi.org/10.1038/nature01145>.
29. Sudarsan, N.; Cohen-Chalamish, S.; Nakamura, S.; Emilsson, G.M.; Breaker, R.R. Thiamine Pyrophosphate Riboswitches Are Targets for the Antimicrobial Compound Pyrithiamine. *Chem. Biol.* **2005**, *12*, 1325–1335. <https://doi.org/10.1016/j.chembiol.2005.10.007>.
30. Kawasaki, T.; Nose, Y. Thiamine Regulatory Mutants in *Escherichia Coli*. *J. Biochem.* **1969**, *65*, 417–425. <https://doi.org/10.1093/oxfordjournals.jbchem.a129029>.
31. Kawasaki, T.; Iwashima, A.; Nose, Y. Regulation of Thiamine Biosynthesis in *Escherichia Coli*. *J. Biochem.* **1969**, *65*, 407–416. <https://doi.org/10.1093/oxfordjournals.jbchem.a129028>.
32. Zeller, M.J.; Nuthanakanti, A.; Li, K.; Aubé, J.; Serganov, A.; Weeks, K.M. Subsite Ligand Recognition and Cooperativity in the TPP Riboswitch: Implications for Fragment-Linking in RNA Ligand Discovery. *ACS Chem. Biol.* **2022**, *17*, 438–448. <https://doi.org/10.1021/acscchembio.1c00880>.
33. Bereswill, S.; Hinkelmann, S.; Kist, M.; Sander, A. Molecular Analysis of Riboflavin Synthesis Genes in *Bartonella Henselae* and Use of the *ribC* Gene for Differentiation of *Bartonella* Species by PCR. *J. Clin. Microbiol.* **1999**, *37*, 3159–3166.
34. Serganov, A.; Huang, L.; Patel, D.J. Coenzyme Recognition and Gene Regulation by a Flavin Mononucleotide Riboswitch. *Nature* **2009**, *458*, 233–237. <https://doi.org/10.1038/nature07642>.
35. Vicens, Q.; Mondragón, E.; Batey, R.T. Molecular Sensing by the Aptamer Domain of the FMN Riboswitch: A General Model for Ligand Binding by Conformational Selection. *Nucleic Acids Res.* **2011**, *39*, 8586–8598. <https://doi.org/10.1093/nar/gkr565>.
36. Winkler, W.C.; Cohen-Chalamish, S.; Breaker, R.R. An mRNA Structure That Controls Gene Expression by Binding FMN. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15908–15913. <https://doi.org/10.1073/pnas.212628899>.
37. Lee, E.R.; Blount, K.F.; Breaker, R.R. Roseoflavin Is a Natural Antibacterial Compound That Binds to FMN Riboswitches and Regulates Gene Expression. *RNA Biol.* **2009**, *6*, 187–194.

38. Langer, S.; Hashimoto, M.; Hobl, B.; Mathes, T.; Mack, M. Flavoproteins Are Potential Targets for the Antibiotic Roseoflavin in *Escherichia Coli*. *J. Bacteriol.* **2013**, *195*, 4037–4045. <https://doi.org/10.1128/jb.00646-13>.
39. Pedrolli, D.B.; Nakanishi, S.; Barile, M.; Mansurova, M.; Carmona, E.C.; Lux, A.; Gärtner, W.; Mack, M. The Antibiotics Roseoflavin and 8-Demethyl-8-Amino-Riboflavin from *Streptomyces Davawensis* Are Metabolized by Human Flavokinase and Human FAD Synthetase. *Biochem. Pharmacol.* **2011**, *82*, 1853–1859. <https://doi.org/10.1016/j.bcp.2011.08.029>.
40. Howe, J.A.; Wang, H.; Fischmann, T.O.; Balibar, C.J.; Xiao, L.; Galgoci, A.M.; Malinverni, J.C.; Mayhood, T.; Villafania, A.; Nahvi, A.; Murgolo, N.; Barbieri, C.M.; Mann, P.A.; Carr, D.; Xia, E.; Zuck, P.; Riley, D.; Painter, R.E.; Walker, S.S.; Sherborne, B.; de Jesus, R.; Pan, W.; Plotkin, M.A.; Wu, J.; Rindgen, D.; Cummings, J.; Garlisi, C.G.; Zhang, R.; Sheth, P.R.; Gill, C.J.; Tang, H.; Roemer, T. Selective Small-Molecule Inhibition of an RNA Structural Element. *Nature* **2015**, *526*, 672–677. <https://doi.org/10.1038/nature15542>.
41. Howe, J.A.; Xiao, L.; Fischmann, T.O.; Wang, H.; Tang, H.; Villafania, A.; Zhang, R.; Barbieri, C.M.; Roemer, T. Atomic Resolution Mechanistic Studies of Ribocil: A Highly Selective Unnatural Ligand Mimic of the *E. Coli* FMN Riboswitch. *RNA Biol.* **2016**, *13*, 946–954. <https://doi.org/10.1080/15476286.2016.1216304>.
42. Winkler, W.C.; Nahvi, A.; Roth, A.; Collins, J.A.; Breaker, R.R. Control of Gene Expression by a Natural Metabolite-Responsive Ribozyme. *Nature* **2004**, *428*, 281–286. <https://doi.org/10.1038/nature02362>.
43. Ferré-D'Amaré, A.R. The *glmS* Ribozyme: Use of a Small Molecule Coenzyme by a Gene-Regulatory RNA. *Q. Rev. Biophys.* **2010**, *43*, 423–447. <https://doi.org/10.1017/S0033583510000144>.
44. Cochrane, J.C.; Lipchock, S.V.; Strobel, S.A. Structural Investigation of the *GlmS* Ribozyme Bound to Its Catalytic Cofactor. *Chem. Biol.* **2007**, *14*, 97–105. <https://doi.org/10.1016/j.chembiol.2006.12.005>.
45. Popova, K.B.; Penchovsky, R. Current Activators of the *glmS* Riboswitch. *Biomed. J. Sci. Tech. Res.* **2021**, *33*, 25528–25530. <https://doi.org/10.26717/BJSTR.2021.33.005348>.
46. Mayer, G.; Famulok, M. High-Throughput-Compatible Assay for *glmS* Riboswitch Metabolite Dependence. *ChemBioChem* **2006**, *7*, 602–604. <https://doi.org/10.1002/cbic.200500490>.
47. Fei, X.; Holmes, T.; Diddle, J.; Hintz, L.; Delaney, D.; Stock, A.; Renner, D.; McDevitt, M.; Berkowitz, D.B.; Soukup, J.K. Phosphatase-Inert Glucosamine 6-Phosphate Mimics Serve as Actuators of the *glmS* Riboswitch. *ACS Chem. Biol.* **2014**, *9*, 2875–2882. <https://doi.org/10.1021/cb500458f>.
48. Mandal, M.; Boese, B.; Barrick, J.E.; Winkler, W.C.; Breaker, R.R. Riboswitches Control Fundamental Biochemical Pathways in *Bacillus Subtilis* and Other Bacteria. *Cell* **2003**, *113*, 577–586. [https://doi.org/10.1016/S0092-8674\(03\)00391-X](https://doi.org/10.1016/S0092-8674(03)00391-X).

49. Kim, J.N.; Blount, K.F.; Puskarz, I.; Lim, J.; Link, K.H.; Breaker, R.R. Design and Antimicrobial Action of Purine Analogues That Bind Guanine Riboswitches. *ACS Chem. Biol.* **2009**, *4*, 915–927. <https://doi.org/10.1021/cb900146k>.
50. Tran, B.; Pichling, P.; Tenney, L.; Connelly, C.M.; Moon, M.H.; Ferré-D'Amaré, A.R.; Schneekloth, J.S.; Jones, C.P. Parallel Discovery Strategies Provide a Basis for Riboswitch Ligand Design. *Cell Chem. Biol.* **2020**, *27*, 1241–1249.e4. <https://doi.org/10.1016/j.chembiol.2020.07.021>.
51. Samant, S.; Lee, H.; Ghassemi, M.; Chen, J.; Cook, J.L.; Mankin, A.S.; Neyfakh, A.A. Nucleotide Biosynthesis Is Critical for Growth of Bacteria in Human Blood. *PLOS Pathog.* **2008**, *4*, e37. <https://doi.org/10.1371/journal.ppat.0040037>.
52. Batey, R.T.; Gilbert, S.D.; Montange, R.K. Structure of a Natural Guanine-Responsive Riboswitch Complexed with the Metabolite Hypoxanthine. *Nature* **2004**, *432*, 411–415. <https://doi.org/10.1038/nature03037>.
53. Serganov, A.; Yuan, Y.-R.; Pikovskaya, O.; Polonskaia, A.; Malinina, L.; Phan, A.T.; Hobartner, C.; Micura, R.; Breaker, R.R.; Patel, D.J. Structural Basis for Discriminative Regulation of Gene Expression by Adenine- and Guanine-Sensing mRNAs. *Chem. Biol.* **2004**, *11*, 1729–1741. <https://doi.org/10.1016/j.chembiol.2004.11.018>.
54. Mulhbachter, J.; Brouillette, E.; Allard, M.; Fortier, L.-C.; Malouin, F.; Lafontaine, D.A. Novel Riboswitch Ligand Analogs as Selective Inhibitors of Guanine-Related Metabolic Pathways. *PLoS Pathog.* **2010**, *6*, e1000865. <https://doi.org/10.1371/journal.ppat.1000865>.
55. Krajewski, S.S.; Isoz, I.; Johansson, J. Antibacterial and Antivirulence Effect of 6-N-Hydroxylaminopurine in *Listeria Monocytogenes*. *Nucleic Acids Res.* **2017**, *45*, 1914–1924. <https://doi.org/10.1093/nar/gkw1308>.
56. Barrett, J.C. Induction of Gene Mutation in and Cell Transformation of Mammalian Cells by Modified Purines: 2-Aminopurine and 6-N-Hydroxylaminopurine. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 5685–5689.
57. Setten, R.L.; Rossi, J.J.; Han, S. The Current State and Future Directions of RNAi-Based Therapeutics. *Nat. Rev. Drug Discov.* **2019**, *18*, 421–446. <https://doi.org/10.1038/s41573-019-0017-4>.
58. Fire, A.; Xu, S.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E.; Mello, C.C. Potent and Specific Genetic Interference by Double-Stranded RNA in *Caenorhabditis Elegans*. *Nature* **1998**, *391*, 806–811. <https://doi.org/10.1038/35888>.
59. Zhao, J.-H.; Guo, H.-S. RNA Silencing: From Discovery and Elucidation to Application and Perspectives. *J. Integr. Plant Biol.* **2022**, *64*, 476–498. <https://doi.org/10.1111/jipb.13213>.
60. Baulcombe, D. RNA Silencing in Plants. *Nature* **2004**, *431*, 356–363. <https://doi.org/10.1038/nature02874>.

61. Baulcombe, D. RNA Silencing. *Trends Biochem. Sci.* **2005**, *30*, 290–293. <https://doi.org/10.1016/j.tibs.2005.04.012>.
62. Meister, G.; Tuschl, T. Mechanisms of Gene Silencing by Double-Stranded RNA. *Nature* **2004**, *431*, 343–349. <https://doi.org/10.1038/nature02873>.
63. Waterhouse, P.M.; Wang, M.-B.; Lough, T. Gene Silencing as an Adaptive Defence against Viruses. *Nature* **2001**, *411*, 834–842. <https://doi.org/10.1038/35081168>.
64. Voinnet, O. Induction and Suppression of RNA Silencing: Insights from Viral Infections. *Nat. Rev. Genet.* **2005**, *6*, 206–220. <https://doi.org/10.1038/nrg1555>.
65. Oliva, G.; Sahr, T.; Buchrieser, C. Small RNAs, 5' UTR Elements and RNA-Binding Proteins in Intracellular Bacteria: Impact on Metabolism and Virulence. *FEMS Microbiol. Rev.* **2015**, *39*, 331–349. <https://doi.org/10.1093/femsre/fuv022>.
66. Felden, B.; Augagneur, Y. Diversity and Versatility in Small RNA-Mediated Regulation in Bacterial Pathogens. *Front. Microbiol.* **2021**, *12*, 719977. <https://doi.org/10.3389/fmicb.2021.719977>.
67. Storz, G.; Vogel, J.; Wassarman, K.M. Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Mol. Cell* **2011**, *43*, 880–891. <https://doi.org/10.1016/j.molcel.2011.08.022>.
68. Saramago, M.; Bárria, C.; dos Santos, R.F.; Silva, I.J.; Pobre, V.; Domingues, S.; Andrade, J.M.; Viegas, S.C.; Arraiano, C.M. The Role of RNases in the Regulation of Small RNAs. *Curr. Opin. Microbiol.* **2014**, *18*, 105–115. <https://doi.org/10.1016/j.mib.2014.02.009>.
69. Aiba, H. Mechanism of RNA Silencing by Hfq-Binding Small RNAs. *Curr. Opin. Microbiol.* **2007**, *10*, 134–139. <https://doi.org/10.1016/j.mib.2007.03.010>.
70. Jørgensen, M.G.; Pettersen, J.S.; Kallipolitis, B.H. sRNA-Mediated Control in Bacteria: An Increasing Diversity of Regulatory Mechanisms. *Biochim. Biophys. Acta BBA—Gene Regul. Mech.* **2020**, *1863*, 194504. <https://doi.org/10.1016/j.bbagr.2020.194504>.
71. Djapgne, L.; Oglesby, A.G. Impacts of Small RNAs and Their Chaperones on Bacterial Pathogenicity. *Front. Cell. Infect. Microbiol.* **2021**, *11*, 604511. <https://doi.org/10.3389/fcimb.2021.604511>.
72. Brouns, S.J.J.; Jore, M.M.; Lundgren, M.; Westra, E.R.; Slijkhuis, R.J.H.; Snijders, A.P.L.; Dickman, M.J.; Makarova, K.S.; Koonin, E.V.; van der Oost, J. Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science* **2008**, *321*, 960–964. <https://doi.org/10.1126/science.1159689>.
73. Wiedenheft, B.; Sternberg, S.H.; Doudna, J.A. RNA-Guided Genetic Silencing Systems in Bacteria and Archaea. *Nature* **2012**, *482*, 331–338. <https://doi.org/10.1038/nature10886>.
74. Gheibi-Hayat, S.M.; Jamialahmadi, K. Antisense Oligonucleotide (AS-ODN) Technology: Principle, Mechanism and Challenges. *Biotechnol. Appl. Biochem.* **2021**, *68*, 1086–1094. <https://doi.org/10.1002/bab.2028>.

75. Collotta, D.; Bertocchi, I.; Chiapello, E.; Collino, M. Antisense Oligonucleotides: A Novel Frontier in Pharmacological Strategy. *Front. Pharmacol.* **2023**, *14*, 1304342. <https://doi.org/10.3389/fphar.2023.1304342>.
76. Stephenson, M.L.; Zamecnik, P.C. Inhibition of Rous Sarcoma Viral RNA Translation by a Specific Oligodeoxyribonucleotide. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 285–288.
77. Penchovsky, R.; Georgieva, A.V.; Dyakova, V.; Traykovska, M.; Pavlova, N. Antisense and Functional Nucleic Acids in Rational Drug Development. *Antibiotics* **2024**, *13*, 221. <https://doi.org/10.3390/antibiotics13030221>.
78. Quemener, A.M.; Bachelot, L.; Forestier, A.; Donnou-Fournet, E.; Gilot, D.; Galibert, M.-D. The Powerful World of Antisense Oligonucleotides: From Bench to Bedside. *WIREs RNA* **2020**, *11*, e1594. <https://doi.org/10.1002/wrna.1594>.
79. Traykovska, M.; Penchovsky, R. Targeting SAM-I Riboswitch Using Antisense Oligonucleotide Technology for Inhibiting the Growth of *Staphylococcus Aureus* and *Listeria Monocytogenes*. *Antibiotics* **2022**, *11*, 1662. <https://doi.org/10.3390/antibiotics11111662>.
80. Traykovska, M.; Penchovsky, R. Engineering Antisense Oligonucleotides as Antibacterial Agents That Target FMN Riboswitches and Inhibit the Growth of *Staphylococcus Aureus*, *Listeria Monocytogenes*, and *Escherichia Coli*. *ACS Synth. Biol.* **2022**, *11*, 1845–1855. <https://doi.org/10.1021/acssynbio.2c00013>.
81. Traykovska, M.; Popova, K.B.; Penchovsky, R. Targeting *glmS* Ribozyme with Chimeric Antisense Oligonucleotides for Antibacterial Drug Development. *ACS Synth. Biol.* **2021**, *10*, 3167–3176. <https://doi.org/10.1021/acssynbio.1c00443>.
82. Zhang, L.; Leibowitz, M.J.; Zhang, Y. Antisense Oligonucleotides Effectively Inhibit the Co-transcriptional Splicing of a *Candida* Group I Intron *in Vitro* and *in Vivo*: Implications for Antifungal Therapeutics. *FEBS Lett.* **2009**, *583*, 734–738. <https://doi.org/10.1016/j.febslet.2009.01.040>.
83. Al-Hashimi, H.M.; Walter, N.G. RNA Dynamics: It Is about Time. *Curr. Opin. Struct. Biol.* **2008**, *18*, 321–329. <https://doi.org/10.1016/j.sbi.2008.04.004>.
84. Vogel, U.; Jensen, K.F. The RNA Chain Elongation Rate in *Escherichia Coli* Depends on the Growth Rate. *J. Bacteriol.* **1994**, *176*, 2807–2813. <https://doi.org/10.1128/jb.176.10.2807-2813.1994>.
85. Brehm, S.L.; Cech, T.R. The Fate of an Intervening Sequence RNA: Excision and Cyclization of the *Tetrahymena* Ribosomal RNA Intervening Sequence *in Vivo*. *Biochemistry* **1983**, *22*, 2390–2397. <https://doi.org/10.1021/bi00279a014>.
86. Woodson, S.A. Folding Mechanisms of Group I Ribozymes: Role of Stability and Contact Order. *Biochem. Soc. Trans.* **2002**, *30*, 1166–1169. <https://doi.org/10.1042/bst0301166>.

87. Ke, A. Crystallization of RNA and RNA Protein Complexes. *Methods* **2004**, *34*, 408–414. <https://doi.org/10.1016/j.ymeth.2004.03.027>.
88. Edwards, A.L.; Garst, A.D.; Batey, R.T. Determining Structures of RNA Aptamers and Riboswitches by X-Ray Crystallography. In *Nucleic Acid and Peptide Aptamers*; Mayer, G., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, USA, 2009; Volume 535, pp. 135–163; ISBN 978-1-934115-89-3.
89. Garst, A.D.; Edwards, A.L.; Batey, R.T. Riboswitches: Structures and Mechanisms. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a003533. <https://doi.org/10.1101/cshperspect.a003533>.
90. Chan, C.W.; Mondragón, A. Crystal Structure of an Atypical Cobalamin Riboswitch Reveals RNA Structural Adaptability as Basis for Promiscuous Ligand Binding. *Nucleic Acids Res.* **2020**, *48*, 7569–7583. <https://doi.org/10.1093/nar/gkaa507>.
91. Bushhouse, D.Z.; Choi, E.K.; Hertz, L.M.; Lucks, J.B. How Does RNA Fold Dynamically? *J. Mol. Biol.* **2022**, *434*, 167665. <https://doi.org/10.1016/j.jmb.2022.167665>.
92. Kang, J.Y.; Mishanina, T.V.; Landick, R.; Darst, S.A. Mechanisms of Transcriptional Pausing in Bacteria. *J. Mol. Biol.* **2019**, *431*, 4007–4029. <https://doi.org/10.1016/j.jmb.2019.07.017>.
93. Zhang, J.; Landick, R. A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure. *Trends Biochem. Sci.* **2016**, *41*, 293–310. <https://doi.org/10.1016/j.tibs.2015.12.009>.
94. Leamy, K.A.; Assmann, S.M.; Mathews, D.H.; Bevilacqua, P.C. Bridging the Gap between *in Vitro* and *in Vivo* RNA Folding. *Q. Rev. Biophys.* **2016**, *49*, e10. <https://doi.org/10.1017/S003358351600007X>.
95. Siegfried, N.A.; Busan, S.; Rice, G.M.; Nelson, J.A.E.; Weeks, K.M. RNA Motif Discovery by SHAPE and Mutational Profiling (SHAPE-MaP). *Nat. Methods* **2014**, *11*, 959–965. <https://doi.org/10.1038/nmeth.3029>.
96. Szyjka, C.E.; Strobel, E.J. Cotranscriptional RNA Chemical Probing. In *Riboregulator Design and Analysis*; Methods in Molecular Biology; Springer USA: New York, NY, USA, 2022; Volume 2518, pp. 291–330 ISBN 978-1-07-162420-3.
97. Frieda, K.L.; Block, S.M. Direct Observation of Cotranscriptional Folding in an Adenine Riboswitch. *Science* **2012**, *338*, 397–400. <https://doi.org/10.1126/science.1225722>.
98. Ray, S.; Chauvier, A.; Walter, N.G. Kinetics Coming into Focus: Single-Molecule Microscopy of Riboswitch Dynamics. *RNA Biol.* **2019**, *16*, 1077–1085. <https://doi.org/10.1080/15476286.2018.1536594>.
99. Strobel, E.J.; Cheng, L.; Berman, K.E.; Carlson, P.D.; Lucks, J.B. A Ligand-Gated Strand Displacement Mechanism for ZTP Riboswitch Transcription Control. *Nat. Chem. Biol.* **2019**, *15*, 1067–1076. <https://doi.org/10.1038/s41589-019-0382-7>.

100. Strobel, E.J.; Watters, K.E.; Nedialkov, Y.; Artsimovitch, I.; Lucks, J.B. Distributed Biotin–Streptavidin Transcription Roadblocks for Mapping Cotranscriptional RNA Folding. *Nucleic Acids Res.* **2017**, *45*, e109. <https://doi.org/10.1093/nar/gkx233>.
101. Chauvier, A.; Picard-Jean, F.; Berger-Dancause, J.-C.; Bastet, L.; Naghdi, M.R.; Dubé, A.; Turcotte, P.; Perreault, J.; Lafontaine, D.A. Transcriptional Pausing at the Translation Start Site Operates as a Critical Checkpoint for Riboswitch Regulation. *Nat. Commun.* **2017**, *8*, 13892. <https://doi.org/10.1038/ncomms13892>.
102. Hien, E.D.M.; Chauvier, A.; St-Pierre, P.; Lafontaine, D.A. Structural Characterization of the Cotranscriptional Folding of the Thiamin Pyrophosphate Sensing *thiC* Riboswitch in *Escherichia Coli*. *Biochemistry* **2024**, *63*, 1608–1620. <https://doi.org/10.1021/acs.biochem.3c00665>.
103. Berman, K.E.; Steans, R.; Hertz, L.M.; Lucks, J.B. A Transient Intermediate RNA Structure Underlies the Regulatory Function of the *E. Coli thiB* TPP Translational Riboswitch. *RNA* **2023**, *29*, 1658–1672. <https://doi.org/10.1261/rna.079427.122>.
104. Chauvier, A.; St-Pierre, P.; Nadon, J.-F.; Hien, E.D.M.; Pérez-González, C.; Eschbach, S.H.; Lamontagne, A.-M.; Penedo, J.C.; Lafontaine, D.A. Monitoring RNA Dynamics in Native Transcriptional Complexes. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2106564118. <https://doi.org/10.1073/pnas.2106564118>.
105. 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* **2005**, *18*, 49–60. <https://doi.org/10.1016/j.molcel.2005.02.032>.
106. Savinov, A.; Block, S.M. Self-Cleavage of the *glmS* Ribozyme Core Is Controlled by a Fragile Folding Element. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 11976–11981. <https://doi.org/10.1073/pnas.1812122115>.
107. Lou, Y.; Woodson, S.A. Co-Transcriptional Folding of the *glmS* Ribozyme Enables a Rapid Response to Metabolite. *Nucleic Acids Res.* **2024**, *52*, 872–884. <https://doi.org/10.1093/nar/gkad1120>.
108. Cheng, L.; White, E.N.; Brandt, N.L.; Yu, A.M.; Chen, A.A.; Lucks, J.B. Cotranscriptional RNA Strand Exchange Underlies the Gene Regulation Mechanism in a Purine-Sensing Transcriptional Riboswitch. *Nucleic Acids Res.* **2022**, *50*, 12001–12018. <https://doi.org/10.1093/nar/gkac102>.
109. Steinert, H.; Sochor, F.; Wacker, A.; Buck, J.; Helmling, C.; Hiller, F.; Keyhani, S.; Noeske, J.; Grimm, S.; Rudolph, M.M.; Keller, H.; Mooney, R.A.; Landick, R.; Suess, B.; Furtig, B.; Wohnert, J.; Schwalbe, H. Pausing Guides RNA Folding to Populate Transiently Stable RNA Structures for Riboswitch-Based Transcription Regulation. *eLife* **2017**, *6*, e21297. <https://doi.org/10.7554/eLife.21297>.
110. Drogalis, L.K.; Batey, R.T. Requirements for Efficient Ligand-Gated Co-Transcriptional Switching in Designed Variants of the B. Subtilis pbuE Adenine-Responsive Riboswitch in *E. coli*. *PLoS ONE* **2020**, *15*, e0243155. <https://doi.org/10.1371/journal.pone.0243155>.

111. Fürtig, B.; Oberhauser, E.M.; Zetzsche, H.; Klötzner, D.-P.; Heckel, A.; Schwalbe, H. Refolding through a Linear Transition State Enables Fast Temperature Adaptation of a Translational Riboswitch. *Biochemistry* **2020**, *59*, 1081–1086. <https://doi.org/10.1021/acs.biochem.9b01044>.
112. Perdrizet, G.A.; Artsimovitch, I.; Furman, R.; Sosnick, T.R.; Pan, T. Transcriptional Pausing Coordinates Folding of the Aptamer Domain and the Expression Platform of a Riboswitch. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 3323–3328. <https://doi.org/10.1073/pnas.1113086109>.
113. Pavlova, N.; Penchovsky, R. Bioinformatics and Genomic Analyses of the Suitability of Eight Riboswitches for Antibacterial Drug Targets. *Antibiotics* **2022**, *11*, 1177. <https://doi.org/10.3390/antibiotics11091177>.
114. Landgraf, T.; Völklein, A.E.; Fürtig, B.; Schwalbe, H. The Cotranscriptional Folding Landscape for Two Cyclic Di-Nucleotide-Sensing Riboswitches with Highly Homologous Aptamer Domains Acting Either as ON- or OFF-Switches. *Nucleic Acids Res.* **2022**, *50*, 6639–6655. <https://doi.org/10.1093/nar/gkac514>.
115. Widom, J.R.; Nedialkov, Y.A.; Rai, V.; Hayes, R.L.; Brooks, C.L.; Artsimovitch, I.; Walter, N.G. Ligand Modulates Cross-Coupling between Riboswitch Folding and Transcriptional Pausing. *Mol. Cell* **2018**, *72*, 541–552.e6. <https://doi.org/10.1016/j.molcel.2018.08.046>.
116. Chauvier, A.; Porta, J.C.; Deb, I.; Ellinger, E.; Meze, K.; Frank, A.T.; Ohi, M.D.; Walter, N.G. Structural Basis for Control of Bacterial RNA Polymerase Pausing by a Riboswitch and Its Ligand. *Nat. Struct. Mol. Biol.* **2023**, *30*, 902–913. <https://doi.org/10.1038/s41594-023-01002-x>.
117. Sherlock, M.E.; Sudarsan, N.; Breaker, R.R. Riboswitches for the Alarmone ppGpp Expand the Collection of RNA-Based Signaling Systems. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6052–6057. <https://doi.org/10.1073/pnas.1720406115>.
118. Merino, E.J.; Wilkinson, K.A.; Coughlan, J.L.; Weeks, K.M. RNA Structure Analysis at Single Nucleotide Resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE). *J. Am. Chem. Soc.* **2005**, *127*, 4223–4231. <https://doi.org/10.1021/ja043822v>.
119. Bingaman, J.L.; Zhang, S.; Stevens, D.R.; Yennawar, N.H.; Hammes-Schiffer, S.; Bevilacqua, P.C. The GlcN6P Cofactor Plays Multiple Catalytic Roles in the *glmS* Ribozyme. *Nat. Chem. Biol.* **2017**, *13*, 439–445. <https://doi.org/10.1038/nchembio.2300>.
120. Mathy, N.; Bénard, L.; Pellegrini, O.; Daou, R.; Wen, T.; Condon, C. 5'-to-3' Exoribonuclease Activity in Bacteria: Role of RNase J1 in rRNA Maturation and 5' Stability of mRNA. *Cell* **2007**, *129*, 681–692. <https://doi.org/10.1016/j.cell.2007.02.051>.
121. Daou-Chabo, R.; Mathy, N.; Bénard, L.; Condon, C. Ribosomes Initiating Translation of the *Hbs* mRNA Protect It from 5'-to-3' Exoribonucleolytic Degradation by RNase J1. *Mol. Microbiol.* **2009**, *71*, 1538–1550. <https://doi.org/10.1111/j.1365-2958.2009.06620.x>.
122. Braun, F.; Durand, S.; Condon, C. Initiating Ribosomes and a 5'/3'-UTR Interaction Control Ribonuclease Action to Tightly Couple *B. Subtilis* *Hbs* mRNA Stability with Translation. *Nucleic Acids Res.* **2017**, *45*, 11386–11400. <https://doi.org/10.1093/nar/gkx793>.

123. Scull, C.E.; Dandpat, S.S.; Romero, R.A.; Walter, N.G. Transcriptional Riboswitches Integrate Timescales for Bacterial Gene Expression Control. *Front. Mol. Biosci.* **2021**, *7*, 607158. <https://doi.org/10.3389/fmolb.2020.607158>.
124. Helmling, C.; Keyhani, S.; Sochor, F.; Fürtig, B.; Hengesbach, M.; Schwalbe, H. Rapid NMR Screening of RNA Secondary Structure and Binding. *J. Biomol. NMR* **2015**, *63*, 67–76. <https://doi.org/10.1007/s10858-015-9967-y>.
125. Kim, P.B.; Nelson, J.W.; Breaker, R.R. An Ancient Riboswitch Class in Bacteria Regulates Purine Biosynthesis and One-Carbon Metabolism. *Mol. Cell* **2015**, *57*, 317–328. <https://doi.org/10.1016/j.molcel.2015.01.001>.
126. Hua, B.; Jones, C.P.; Mitra, J.; Murray, P.J.; Rosenthal, R.; Ferré-D'Amaré, A.R.; Ha, T. Real-Time Monitoring of Single ZTP Riboswitches Reveals a Complex and Kinetically Controlled Decision Landscape. *Nat. Commun.* **2020**, *11*, 4531. <https://doi.org/10.1038/s41467-020-18283-1>.
127. Bushhouse, D.Z.; Lucks, J.B. Tuning Strand Displacement Kinetics Enables Programmable ZTP Riboswitch Dynamic Range *in Vivo*. *Nucleic Acids Res.* **2023**, *51*, 2891–2903. <https://doi.org/10.1093/nar/gkad110>.
128. Bushhouse, D.Z.; Fu, J.; Lucks, J.B. RNA Folding Kinetics Control Riboswitch Sensitivity *in Vivo*. *bioRxiv* **2024**, *preprint*. <https://doi.org/10.1101/2024.03.29.587317>.
129. Binas, O.; Schamber, T.; Schwalbe, H. The Conformational Landscape of Transcription Intermediates Involved in the Regulation of the ZMP-Sensing Riboswitch *from Thermosinus Carboxydivorans*. *Nucleic Acids Res.* **2020**, *48*, 6970–6979. <https://doi.org/10.1093/nar/gkaa427>.
130. Baker, J.L.; Sudarsan, N.; Weinberg, Z.; Roth, A.; Stockbridge, R.B.; Breaker, R.R. Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. *Science* **2012**, *335*, 233–235. <https://doi.org/10.1126/science.1215063>.
131. Chauvier, A.; Ajmera, P.; Yadav, R.; Walter, N.G. Dynamic Competition between a Ligand and Transcription Factor NusA Governs Riboswitch-Mediated Transcription Regulation. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2109026118. <https://doi.org/10.1073/pnas.2109026118>.
132. Yadav, R.; Widom, J.R.; Chauvier, A.; Walter, N.G. An Anionic Ligand Snap-Locks a Long-Range Interaction in a Magnesium-Folded Riboswitch. *Nat. Commun.* **2022**, *13*, 207. <https://doi.org/10.1038/s41467-021-27827-y>.
133. Hertz, L.M.; White, E.N.; Kuznedelov, K.; Cheng, L.; Yu, A.M.; Kakkaramadam, R.; Severinov, K.; Chen, A.; Lucks, J.B. The Effect of Pseudoknot Base Pairing on Cotranscriptional Structural Switching of the Fluoride Riboswitch. *Nucleic Acids Res.* **2024**, *52*, 4466–4482. <https://doi.org/10.1093/nar/gkae231>.

134. Chauvier, A.; Dandpat, S.S.; Romero, R.; Walter, N.G. A Nascent Riboswitch Helix Orchestrates Robust Transcriptional Regulation through Signal Integration. *Nat. Commun.* **2024**, *15*, 3955. <https://doi.org/10.1038/s41467-024-48409-8>.
135. Helmling, C.; Wacker, A.; Wolfinger, M.T.; Hofacker, I.L.; Hengesbach, M.; Fürtig, B.; Schwalbe, H. NMR Structural Profiling of Transcriptional Intermediates Reveals Riboswitch Regulation by Metastable RNA Conformations. *J. Am. Chem. Soc.* **2017**, *139*, 2647–2656. <https://doi.org/10.1021/jacs.6b10429>.
136. Helmling, C.; Klötzner, D.-P.; Sochor, F.; Mooney, R.A.; Wacker, A.; Landick, R.; Fürtig, B.; Heckel, A.; Schwalbe, H. Life Times of Metastable States Guide Regulatory Signaling in Transcriptional Riboswitches. *Nat. Commun.* **2018**, *9*, 944. <https://doi.org/10.1038/s41467-018-03375-w>.
137. Xue, Y.; Li, J.; Chen, D.; Zhao, X.; Hong, L.; Liu, Y. Observation of Structural Switch in Nascent SAM-VI Riboswitch during Transcription at Single-Nucleotide and Single-Molecule Resolution. *Nat. Commun.* **2023**, *14*, 2320. <https://doi.org/10.1038/s41467-023-38042-2>.
138. Childs, J.L.; Disney, M.D.; Turner, D.H. Oligonucleotide Directed Misfolding of RNA Inhibits *Candida Albicans* Group I Intron Splicing. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11091–11096. <https://doi.org/10.1073/pnas.172391199>.
139. Southwell, A.L.; Skotte, N.H.; Bennett, C.F.; Hayden, M.R. Antisense Oligonucleotide Therapeutics for Inherited Neurodegenerative Diseases. *Trends Mol. Med.* **2012**, *18*, 634–643. <https://doi.org/10.1016/j.molmed.2012.09.001>.
140. Manoharan, M. 2'-Carbohydrate Modifications in Antisense Oligonucleotide Therapy: Importance of Conformation, Configuration and Conjugation. *Biochim. Biophys. Acta* **1999**, *1489*, 117–130.
141. Nikolcheva, T.; Woodson, S.A. Association of a Group I Intron with Its Splice Junction in 50S Ribosomes: Implications for Intron Toxicity. *RNA* **1997**, *3*, 1016–1027.
142. Roberts, T.C.; Langer, R.; Wood, M.J.A. Advances in Oligonucleotide Drug Delivery. *Nat. Rev. Drug Discov.* **2020**, *19*, 673–694. <https://doi.org/10.1038/s41573-020-0075-7>.
143. McClorey, G.; Banerjee, S. Cell-Penetrating Peptides to Enhance Delivery of Oligonucleotide-Based Therapeutics. *Biomedicines* **2018**, *6*, 51. <https://doi.org/10.3390/biomedicines6020051>.
144. Geller, B.L.; Marshall-Batty, K.; Schnell, F.J.; McKnight, M.M.; Iversen, P.L.; Greenberg, D.E. Gene-Silencing Antisense Oligomers Inhibit *Acinetobacter* Growth *In Vitro* and *In Vivo*. *J. Infect. Dis.* **2013**, *208*, 1553–1560. <https://doi.org/10.1093/infdis/jit460>.
145. Xue, X.-Y.; Mao, X.-G.; Zhou, Y.; Chen, Z.; Hu, Y.; Hou, Z.; Li, M.-K.; Meng, J.-R.; Luo, X.-X. Advances in the Delivery of Antisense Oligonucleotides for Combating Bacterial Infectious Diseases. *Nanomedicine Nanotechnol. Biol. Med.* **2018**, *14*, 745–758. <https://doi.org/10.1016/j.nano.2017.12.026>.

146. Buthelezi, L.A.; Pillay, S.; Ntuli, N.N.; Gcanga, L.; Guler, R. Antisense Therapy for Infectious Diseases. *Cells* **2023**, *12*, 2119. <https://doi.org/10.3390/cells12162119>.
147. Meng, J.; Da, F.; Ma, X.; Wang, N.; Wang, Y.; Zhang, H.; Li, M.; Zhou, Y.; Xue, X.; Hou, Z.; Jia, M.; Luo, X. Antisense Growth Inhibition of Methicillin-Resistant *Staphylococcus Aureus* by Locked Nucleic Acid Conjugated with Cell-Penetrating Peptide as a Novel FtsZ Inhibitor. *Antimicrob. Agents Chemother.* **2015**, *59*, 914–922. <https://doi.org/10.1128/AAC.03781-14>.
148. Sawyer, A.J.; Wesolowski, D.; Gandotra, N.; Stojadinovic, A.; Izadjoo, M.; Altman, S.; Kyriakides, T.R. A Peptide-Morpholino Oligomer Conjugate Targeting *Staphylococcus Aureus gyrA* mRNA Improves Healing in an Infected Mouse Cutaneous Wound Model. *Int. J. Pharm.* **2013**, *453*, 651–655. <https://doi.org/10.1016/j.ijpharm.2013.05.041>.
149. Lu, R.; Deng, L.; Lian, Y.; Ke, X.; Yang, L.; Xi, K.; Ong, A.A.L.; Chen, Y.; Zhou, H.; Meng, Z.; Lin, R.; Fan, S.; Liu, Y.; Toh, D.-F. K.; Zhan, X.; Krishna, M.S.; Patil, K.M.; Lu, Y.; Liu, Z. Zhu, L.; Wang, H.; Li, G.; Chen, G. Recognition of RNA Secondary Structures with a Programmable Peptide Nucleic Acid-Based Platform. *Cell Rep. Phys. Sci.* **2024**, *5*, 102150. <https://doi.org/10.1016/j.xcrp.2024.102150>.

11. Acknowledgements

Chapter 5, in full, is a reprint of the material as it appears in Opportunities for riboswitch inhibition by targeting co-transcriptional RNA folding events. Stephen, Christine; Palmer, Danea; Mishanina, Tatiana, International Journal of Molecular Sciences, 2024. The dissertation author was the primary co-author of this paper.