

Movement of the *Escherichia coli* RNA polymerase clamp modulates intrinsic transcription
termination

By

Michael J. Bellecourt

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Biochemistry)

At the

UNIVERSITY OF WISCONSIN–MADISON

2018

Date of final oral examination: 08/21/2018

The dissertation is approved by the following members of the Final Oral Committee:

Robert C. Landick, Professor, Biochemistry and Bacteriology

Aaron A. Hoskins, Associate Professor, Biochemistry

M. Thomas Record, Jr., Professor, Biochemistry and Chemistry

James C. Weisshaar, Professor, Chemistry

Movement of the *Escherichia coli* RNA polymerase clamp modulates intrinsic transcription termination

Michael John Allen Bellecourt

under the supervision of Professor Robert Landick

University of Wisconsin–Madison

Abstract

All cellular organisms rely upon multisubunit RNA polymerases (RNAPs) to accurately turn on expression of a gene through the process of transcription. Accurate and timely RNA synthesis requires careful regulation of all steps of the complex transcription cycle. This regulatory framework relies upon the careful interplay of interactions between RNAP and the sequences and structure of template DNA and the nascent RNA transcript as well as myriad transcription factors that associate with both RNAP itself and the chromosome to reprogram the transcription cycle. To ensure that RNAP can extend an entire RNA chain without premature dissociation from the template, the enzyme has evolved dozens of protein–nucleic acid contacts that confer upon it a remarkable stability and processivity. However, it is essential for transcription to end spontaneously and efficiently at the end of a transcription unit through dissociation of both RNA and DNA from RNAP. Failure to dissociate will cause RNAP to elongate into downstream genes, potentially turning on the expression of genes the cell otherwise wants turned off, or even inducing catastrophic collisions with the replication machinery.

Cells have evolved several termination mechanisms that enables balancing of the transcription elongation complex's (EC's) high stability and processivity throughout

transcription elongation with its need to induce dissociate at precise base pairs along the transcription unit. Bacteria have evolved three such mechanisms, two of which have been programmed into the DNA template itself. During intrinsic termination, RNAP transcribes a GC-dyad immediately adjacent to a downstream, U-rich tract. The U-tract induces a brief transcriptional pause that provides time for the palindromic GC-dyad to fold upon itself, situating an RNA stem loop structure called a terminator hairpin (T_{hp}) in RNA exit channel of RNAP. Complete folding of the T_{hp} , a two-step process, inactivates RNAP for nucleotide addition, committing it to a slow release from the chromosome. The second programmed mechanism, Rho-dependent termination, relies not on an RNA secondary structure but the homohexameric RNA translocase Rho. This enzyme recognizes and binds C-rich RNAs, after which it is activated for translocation of the nascent transcript through ATP hydrolysis. Once Rho makes contact with RNAP, its motor activity works to disengage RNA from RNAP, destabilizing the EC and inducing termination. These two programmed mechanisms are the focus of this dissertation. A third, non-programmed mechanism, Mfd-dependent termination, utilizes a DNA translocase to induce dissociation of intragenically arrested ECs.

Prior to these studies, the role of nucleic acid sequence and structure in the programmed termination mechanisms were well-characterized, but the role of protein conformation changes were largely unknown. Utilizing two variants of a Cysteine-pair *in vitro* transcription assay, I was able to demonstrate how movements in the RNAP clamp module affect the rate of elongation but not the rate of EC inactivation at an intrinsic terminator. I was also able to demonstrate that topological constraint of the clamp had no effects on the rate of RNA release from an intrinsic terminator but had strong effects on the ability of RNAP to release template DNA. Next, I demonstrated how constraining the conformational freedom of the trigger loop, an

active site-proximal motif that enhances the rate of catalysis, was able to slow the dissociation rates of both RNA and DNA from RNAP, thus expanding a multistate-multipath model of termination to include both the inactivation and dissociation processes.

Finally, I resolved an outstanding contradiction by which the universally-conserved transcription elongation factor NusG also acts as a Rho-dependent termination factor. In collaboration with X-ray crystallographers, I determined that after binding C-rich RNAs in its secondary binding site, the Rho ring is capable of making small amino acid rearrangements that induce an open-to-closed ring, even in the absence of its NusG accessory factor. However, at subpar RNAs that Rho cannot effectively bind on its own, the C-terminal domain of NusG can associate in the vicinity of the relevant amino acids switch network to induce the necessary rearrangements. I validated these structural findings through *in vitro* transcription and confirmed the existence of the NusG–Rho–EC regulatory network, thus establishing the mechanism of NusG-dependent Rho termination.

Taken together, the studies contained in this dissertation underscore the importance of protein conformation during the termination process, from small-scale Brownian motion to large-scale movements of protein modules. As we improve our understanding of the termination mechanism, so too will we improve our abilities to design new antibiotics and improve the function of synthetic gene circuits of all types.

Acknowledgements

To my family.

You may not have ever understood what I was talking about, and you may have even once thought I was being literal when I said Rho was “like a machine”, but that’s okay because you’ve always been supportive of me in my zany pursuit of academic truths and scientific awesomeness. I particularly need to thank those of you who subjected yourself to my dissertation defense seminar, because I already know that it will be the most boring hour of your life.

To my co-workers.

I cannot stress how important to it was too me that I was able to work with so many delightful scientists here in Madison. The Landick Group is filled, and has been filled, with some of the most clever, insightful, and all-out fun people I have ever had the joy of meeting, and I have been so fortunate to study transcription here, where I can just as easily ask members of the Gourse and Record Groups for their own expert perspectives. I particularly need to thank Ananya Ray-Soni, upon whose wisdom I depended upon—without you, Ananya, I would not have half the understanding of the termination mechanism I do today. I need to thank Rachel Mooney, both for her scientific prowess and because she so often saved me from the tedium that is molecular cloning. And, of course, I need to thank Professor Robert Landick, my mentor, for all that he has done. His patience with me is nearly as infinite as his ability to think, and if I ever know even half of what he has forgotten since last night I will consider myself a very learned man indeed.

To my friends.

You are legion for you are many. You have supported me in good times and in bad, and I am glad to have known (and drank too much beer) with all of you. Some of you I have known for years, and others are those people I'm glad to have met here in Madison. Many of you were labmates, and others were biochemists and microbiologists in other labs. A good number of you taught me the power of solidarity amongst employees. A couple of you got mad at me when I got a trivia question wrong, but I forgive you anyway because you've put up with me when I was at my worst. One of you was a font of some much-needed joy during the most trying months of my life, and for that you will long have my gratitude. You've all since scattered across the globe, and I don't see any of you nearly as often as I wish anymore. But, hey, that's okay! Now I never have to reserve another hotel room in my life.

To RNA polymerase

May you never fail to give incoming graduate students years of study.

Table of Contents

Abstract.....	i
Acknowledgements	iv
Table of Contents.....	vi
List of Figures	xi
List of Tables.....	xiii
 Chapter 1: All good things must end: an introduction to bacterial transcription	
termination.....	1
1.1. The three mechanisms of bacterial transcription termination.....	2
1.2 Elongation complex stability and its destabilization by terminators	3
1.2.1 EC stability.....	3
1.2.2 EC destabilization by termination and the role of transcriptional pausing.....	6
1.3 Impact of transcription factors on termination	11
1.4 The mechanism of intrinsic termination.....	12
1.4.1 Elemental pausing at the U-tract.....	13
1.4.2 Terminator hairpin nucleation.....	13
1.4.3 Terminator hairpin completion	14
1.4.4 EC inactivation and commitment to termination	18
1.4.5 EC dissociation.....	20
1.4.6 Additional roles of multiple RNAP conformations in termination.....	21
1.5 Rho-dependent transcription termination.....	24
1.5.1 Recognition of Rho-utilization elements.....	24
1.5.2 Translocation towards the elongating EC.....	30

1.5.3	Kinetic coupling of Rho to the EC.....	31
1.5.4	The mechanism of RNA release and EC dissociation.....	32
1.6	NusG acts both as a positive and negative regulator of Rho	34
1.7	Physiological roles for Rho	37
1.7.1	Rho acts to prevent the accumulation of R-loops	37
1.7.2	Nucleoid-associated proteins promote termination of foreign DNA and antisense transcripts	38
1.8	Conclusions.....	40
1.9	Organization of dissertation.....	43
1.10	References.....	44
 Chapter 2: RNA polymerase clamp movement aids dissociation from DNA but is not required for RNA release at intrinsic terminators		
2.1	Abstract.....	57
2.2	Introduction.....	58
3.3	Results	63
3.3.1	Restricting clamp movements by Cys-pair crosslinking alters elongation rate.....	63
2.3.2	Clamp conformation impacts termination commitment through effects on elongation	68
2.3.3	RNA efficiently dissociates from an elongation complex at <i>t_{hisL}</i> regardless of clamp conformation	70
2.3.4	RNAP release from DNA is inhibited when clamp movement is blocked	78
2.3.5	Inhibiting RNAP conformational flexibility with Microcin J25 impairs both RNA and RNAP release	79

2.4	Discussion.....	83
2.4.1	Clamp conformation has stronger effects on elongation than termination.....	84
2.4.2	Upstream-for junction mobility may weaken terminators that require forward translocation for commitment.....	87
2.4.3	Clamp opening appears to be required for RNAP release from DNA at intrinsic terminators.....	88
2.4.4	TL movement is required for release of both RNA and DNA at an intrinsic terminator	89
2.4.5.	Conclusion	90
2.5	Materials and Methods	92
2.5.1	Materials	92
2.5.2	Protein and peptide purifications	92
2.5.3	Producing linear DNA fragments for transcription.....	97
2.5.4	Transcription on fully complementary oligonucleotide scaffolds	98
2.5.5	Dissociation assays.....	99
2.5.6	MccJ25 abortive initiation assays	100
2.5.7	Trigger loop cysteine-pair reporter assays.....	100
2.5.8	Determination of commitment rates and propagation of error	101
2.6.	References.....	101
 Chapter 3: Mechanism for the regulated control of Rho-dependent termination by NusG.....		
		107
3.1	Abstract.....	108
3.2	Introduction.....	108

3.3	Results	112
3.3.1	Structure of a Rho–NusG complex	112
3.3.2	NusG promotes early Rho-dependent termination through its action on Rho ring conformation.....	116
3.4	Discussion.....	118
3.5.	Materials and Methods	126
3.5.1	Cloning and protein purification	126
3.5.2	Protein crystallization and data collection.....	128
3.5.3	Structure solution and refinement	128
3.5.4	<i>In vitro</i> Rho-dependent transcription termination assays	130
3.6.	References.....	130
Chapter 4: Dissertation dissociation: conclusions and future directions.....		135
4.1.	Conclusions.....	136
4.2	Future Directions.....	139
4.2.1	Determining the structures of the committed EC with cryo-EM.....	139
4.2.2	MS-MP energetics and termination commitment during Rho-dependent termination.....	141
4.2.3	Recruitment of Rho to the EC.....	143
4.2.4	The termination mechanism in other species.....	144
4.3	References.....	145
Appendix A: Hairpin-stabilized but not elemental pause durations are affected by clamp topology.....		149
A.1	Introduction.....	150

A.2	Results	151
A.2.1	Topologically constraining the clamp perturbs the duration of the <i>his</i> hairpin-stabilized pause	151
A.2.2	Biasing the clamp into a displaced but not closed conformation perturbs elemental pausing duration.....	155
A.4	Discussion.....	157
A.5	Materials and Methods	159
A.5.1	Materials	159
A.5.2	Protein purifications	159
A.5.3	Transcription on fully complementary oligonucleotide scaffolds.....	163
A.6	References.....	164

List of Figures

Chapter 1: All good things must end: an introduction to bacterial transcription

termination.....	1
Figure 1.1. Structural domains of the elongation complex (EC).....	5
Figure 1.2. Nucleic-acid scaffolds at stages of transcript elongation and termination.....	8
Figure 1.3. The mechanism of intrinsic termination.....	10
Figure 1.4. Structures of Rho and mechanisms of Rho termination.....	25
Figure 1.5. Synergy between Rho and bridged H-NS filaments <i>in vitro</i>	41

Chapter 2: RNA polymerase clamp movement aids dissociation from DNA but is not

required for RNA release at intrinsic terminators	56
Figure 2.1. The mechanism of intrinsic termination.....	59
Figure 2.2. Forming disulfide crosslinks between the RNAP flap and lid.	61
Figure 2.3. The bead-tethered ligation scaffold assay.	65
Figure 2.4. Constraining the clamp perturbs overall elongation rates through effects on relative pause strength.....	67
Figure 2.5. Evaluation of termination commitment rates.	69
Figure 2.6. T-RP forms a disulfide crosslink within the transcription bubble.....	72
Figure 2.7. Elongation to t_{hisL} from A26 with different NTP concentrations under reducing conditions.	73
Figure 2.8. Representative dissociation assay non-reducing gels.	75
Figure 2.9. Evaluation of RNA and RNAP dissociation rates.	76
Figure 2.10. Evaluation of RNA and RNAP dissociation rates in the presence of MccJ25.	80
Figure 2.11. Effects of MccJ25 on RNAP during abortive initiation and elongation.	81

Figure 2.12. Cysteine-pair reporting suggests MccJ25 binds the TL in its unfolded conformation.	82
Figure 2.13. Distance between the TL and the RNA 3'-end as is proposed to occur during termination commitment.	91
Chapter 3: Mechanism for the regulated control of Rho-dependent termination by NusG.....	
NusG.....	107
Figure 3.1. NusG promotes Rho ring closure.....	110
Figure 3.2. Structure of closed-ring Rho in complex with NusG.....	113
Figure 3.3. Representative electron density for the Rho–NusG, rU12, ADP–BeF3 structure.....	114
Figure 3.4. <i>In vitro</i> termination assays confirm the dependence of NusG on switch amino acids in Rho.....	117
Figure 3.5. Model for NusG-mediated promotion of Rho-dependent transcription termination.	120
Figure 3.6. Role of NusG in regulating Rho-dependent termination.	123
Appendix A: Hairpin-stabilized but not elemental pause durations are affected by clamp topology.....	
clamp topology.....	149
Figure A.1. Cys-pair ligation scaffold assay.....	152
Figure A.2. Effects of topologically constraining the clamp on the <i>hisP</i>	154
Figure A.3. Effects of topologically constraining the clamp on two classes of the elemental pause.	156

List of Tables

Chapter 2: RNA polymerase clamp movement aids dissociation from DNA but is not required for RNA release at intrinsic terminators	56
Table 2.1. Oligonucleotides used in this study.....	93
Table 2.2. Plasmids used in this study.....	95
Chapter 3: Mechanism for the regulated control of Rho-dependent termination by NusG.....	107
Table 3.1. Data collection and refinement.	115
Appendix A: Hairpin-stabilized but not elemental pause durations are affected by clamp topology	149
Table A.1. Oligonucleotides used in this study.....	160
Table A.2. Plasmids used in this study.	162

Chapter 1

All good things must end: an introduction to bacterial transcription termination

Adapted from:

Ananya Ray-Soni, Michael J. Bellecourt, and Robert Landick. (2016). Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annual Review of Biochemistry*. **85**: 319–347.

Chapter contributions:

M.J.B., A.R.S., and R.L. wrote the original text. M.J.B. revised the text to reflect the most current published findings. M.J.B. constructed all figures.

1.1. The three mechanisms of bacterial transcription termination

Understanding the mechanism and sequence determinants of transcription termination by bacterial RNA polymerases (RNAP) is a long-standing goal in the study of gene regulation. Termination (*i*) prevents inappropriate expression of downstream genes and interference from antisense transcripts; (*ii*) defines RNA 3' ends for precise RNA structures (*e.g.*, small noncoding RNAs) or for regulation; (*iii*) recycles RNAP for efficient gene expression; and (*iv*) minimizes collisions with replication complexes that cause damaging double-strand breaks in the chromosome.

Transcription termination in bacteria occurs in one of three pathways: intrinsic termination, which requires only RNAP, RNA, and DNA, or the two factor-dependent pathways of Rho- and Mfd-dependent termination, which require the additional action of *trans*-acting proteins that dissociate the elongation complex (EC) [1]. Intrinsic termination occurs when an RNA secondary structure (called a terminator hairpin, T_{hp}) forms in the exit channel of RNAP positioned on an unstable RNA-DNA hybrid, causing EC dissociation. Rho-dependent termination relies on the ATP-dependent RNA-DNA helicase, Rho, which binds to and translocates the nascent RNA until it encounters a paused EC and causes release of RNAP. Rho requires a less specific sequence at which to dissociate than intrinsic termination, which is important for its essential role in suppression of R-loops, horizontally acquired genes, and antisense transcription [2-5]. Together, intrinsic and Rho-dependent termination comprise the two major programmed mechanisms of transcription termination.

Mfd-dependent termination, on the other hand, is not programmed by specific nucleic-acid sequences. Instead, the Superfamily 2 DNA translocase Mfd recognizes stalled ECs to initiate the transcription-coupled repair (TCR) pathway. During TCR, stalled ECs that could otherwise

block DNA replication are removed and the UvrABC repair proteins are recruited. Upon binding to both the stalled RNAP and upstream DNA, Mfd activates its ATP-dependent translocase activity. Relative movement of Mfd in the direction of the stalled EC induces hypertranslocation of RNAP without NTP addition and subsequent bubble rewinding, inducing release of the nascent transcript and dissociation of RNAP [6]. UvrAB is then recruited directly to the damage site via Mfd's UvrB homology domain [7]. Ongoing studies into the mechanism of Mfd-dependent termination are focused on how Mfd discriminates between arrested and paused ECs and on the fraction of TCR that depends upon Mfd action.

In this work I focus on the programmed mechanisms of intrinsic and Rho-dependent termination, the pathways associated with gene expression and regulation rather than DNA repair. Crystallographic and cryo-electron microscopy (cryo-EM) advances over the past two decades have significantly improved our understanding of RNAP and EC architecture and provided an understanding of how the interconnected domains, modules, and motifs of RNAP interplay to regulate the entire transcription cycle (Fig. 1.1). This body of work seeks to further our understanding of how conformation changes modulate the termination processes.

1.2 Elongation complex stability and its destabilization by terminators

1.2.1 EC stability

To understand dissociation of ECs at terminators, one must first understand the determinants of EC stability. RNAP has evolved to form a remarkably stable, yet dynamic network of contacts with the nucleic-acid scaffold in the EC, enabling it to transcribe over 10^4 nt without dissociating from the template. The enzyme protects a 12–14 bp transcription bubble, a 9–10 bp RNA-DNA hybrid, ~5 nt of newly transcribed single-stranded RNA (ssRNA), and ~18 bp of downstream

duplex DNA (Fig. 1.1) [8, 9]. The major sources of EC stability are polar and van der Waals contacts between RNAP and the RNA-DNA hybrid in the main channel of the enzyme, augmented by H-bonds to the ssRNA in the exit channel and long-range electrostatic and van der Waals contacts to the downstream DNA [10]. Comparison of the structures of the EC versus free RNAP reveals that these contacts are established upon closure of the clamp domain and a small rotation of the shelf module relative to the core of RNAP (Fig. 1.1A) [10-15]. These changes are detectable by disulfide bond formation using cysteine pairs strategically engineered into *Escherichia coli* or *Thermus thermophilus* RNAP [16, 17]. The closed-clamp state appears to be stabilized in part by interactions between downstream DNA and the RNAP clamp and jaw, and most significantly by extensive contacts between the folded forms of three switches and the RNA-DNA hybrid and downstream DNA [9, 18-20]. The switches (Sw1-5) are five highly conserved polymorphous elements at the base of the clamp that rearrange to enable a 30° clamp rotation relative to the RNAP core, and are compacted by stabilizing contacts to nucleic-acids in the EC [18]. Specifically, Sw1 makes stabilizing contacts to the template strand in the RNA-DNA hybrid, absolutely conserved Sw2 K334 marks the downstream edge of the RNA-DNA hybrid by stabilizing the 90° kink in the template strand between +1 and +2. In crystal structures of the *T. thermophilus* enzyme, Sw3 forms a hydrophobic binding pocket for the first ssRNA base upstream of the RNA-DNA hybrid [10, 18, 21], while cryo-EM structures of *E. coli* ECs suggests that it, along with the Zn²⁺-binding domain, dock, and the β C-terminus, forms a positively-charged groove that guides the nascent transcript through the RNA exit channel [15, 22, 23]. Whether these distinct apparent roles for Sw3 are due to differences in species or artifacts of crystallization remains unresolved. The upstream edge of the hybrid is further stabilized by stacking interactions with the lid [10, 15]. This network of non-specific interactions

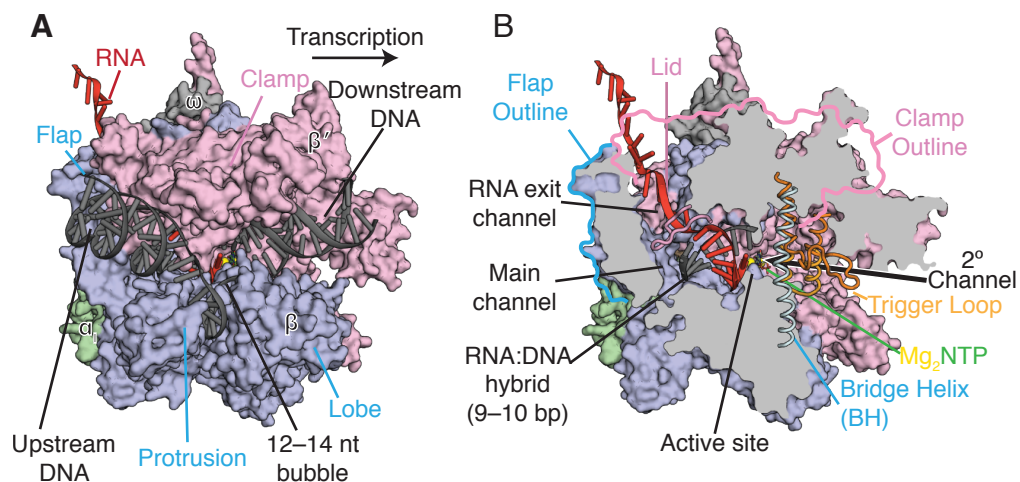


Figure 1.1. Structural domains of the elongation complex (EC). (a) Model of the EC structure based on a crystal structure of the *Thermus thermophilus* EC (PDB 2o5j); subunits and motifs of interest are indicated. (b) Sliced view of the EC, revealing key features of the RNA exit channel, main channel, and active site (PDB 2o5j). Both folded (light orange; PDB 1iw7) and unfolded (dark orange; PDB 2o5j) conformations of the trigger loop (TL) are depicted.

creates an EC that is stable up to 70 °C and at concentrations of more than 500 mM salt [24-26] while still allowing efficient translocation along the DNA.

1.2.2 EC destabilization by termination and the role of transcriptional pausing

The challenge of transcription termination is to destabilize the highly stable EC sufficiently at specific genome locations to enable RNAP dissociation from the DNA and RNA. However, termination can only occur at a given position if RNAP enters the termination pathway faster than nucleotide addition moves it to the next position. Consequently, efficient termination is aided by factors that increase the kinetic window at the termination branchpoint and impeded by factors that increase the elongation rate. Sequence-dependent pauses can reversibly halt ECs along the template from seconds to minutes [27, 28], thereby increasing the kinetic window for entry into the termination pathway. They may also alter the EC conformation to make it susceptible to termination. Pausing is thus a necessary first step for efficient termination [25, 29].

Transcriptional pausing is defined as a kinetic branch from the main elongation pathway at which the EC enters a halted state that inhibits nucleotide addition. Pause escape occurs when the active site rearranges back into an active configuration and resumes nucleotide addition. An EC may isomerize into an offline paused state at any position on a template [30], but these excursions from the nucleotide addition cycle into elemental pauses are preferentially triggered and stabilized by specific sequences in the RNA-DNA hybrid, non-template strand, downstream DNA, and active-site nucleotides [21, 31, 32]. Recent structural studies into the elemental pause suggest that the initial offline state arises during translocation when the incoming template base becomes trapped on the BH, generating an unusual RNA-DNA “half-translocated” hybrid register where the RNA is able to translocate but the DNA template cannot (Fig. 1.2) [22].

Earlier crystallography studies into similar sequences indirectly suggest that the elemental pause is associated with clamp “loosening”; that is, a relaxing of protein-NA contacts that allow transient explorations of the open clamp. This crystal structure revealed RNAP bound to elemental paused sequences with a widened main cleft, the switch regions swung away from the nucleic-acids, breaking their contacts with the RNA-DNA hybrid [21]. Since it is now known that the elemental pause does not drive clamp opening in *E. coli*, these studies suggest that the loosening of protein-NA contacts enabled stabilization and capture of the open-clamp by crystal packing forces. RNA secondary structures in the exit channel can serve to stabilize the loosened clamp into an alternative conformation, but cryo-EM surprisingly reveals that this is through a subtle, 3° swiveling movement parallel to the nucleic-acid scaffold rather than the 30° orthogonal opening anticipated by the crystal structure [16, 17, 22, 23].

These RNA secondary structures in the RNA exit channel extend the lifetime of the elemental pause through a mechanism named hairpin-stabilized pausing. A second mode of pause stabilization occurs through backward translocation of the EC (backtracking). Hairpin-stabilized pausing occurs when complementary sequences in the nascent transcript pair to form a pause hairpin in the RNA exit-channel 11–12 nt from the RNA 3' end and swivels the clamp the concerted motion of the swivel module, comprised of the clamp, dock, shelf, and (in *E. coli*, at least) sequence insertion 3. [16, 17, 22, 23]. Clamp swiveling allosterically inhibits elongation by preventing proper folding of the active site-proximal trigger loop motif into its catalytically-active helical form [22]. Hairpin formation also requires interactions between the pause hairpin and the flap domain (Fig. 1.1A) [33, 34] and can be aided by binding of NusA [35-37], which contacts both the flap-tip and the pause hairpin [38-40].

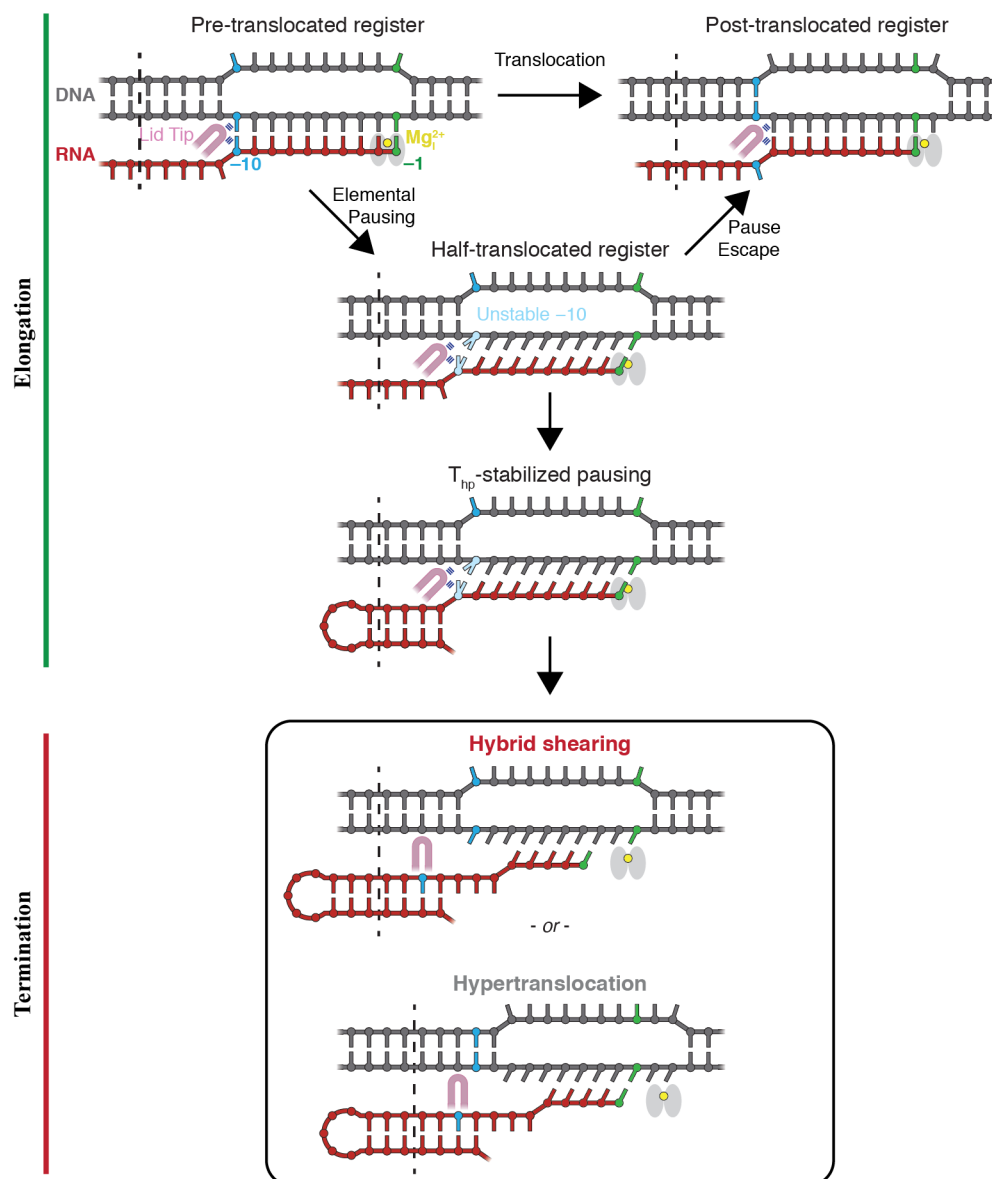


Figure 1.2. Nucleic-acid scaffolds at stages of transcript elongation and termination. Pre-translocated EC, post-translocated EC, half-translocated PEC, after hairpin nucleation, and after inactivation by hybrid-shearing or hypertranslocation. The active site is denoted by two grey ovals and the tip of the lid by a purple loop. Stacking interactions between the lid and the upstream hybrid are depicted by blue dashed lines. The position of the upstream end of RNAP relative to RNA is represented by a black dashed line. Green and blue nucleotides represent the -1 and -10 bp, respectively. The instability of the -10 bp is represented by its lighter coloring. Green and red sidebars: though pausing is an obligate process in the termination pathway, these are considered elongation, not termination processes.

Backtracking is a mechanism by which RNA and DNA reverse translocate through RNAP such that the 3' nascent RNA extrudes into the secondary channel (Fig. 1.1B), consequently removing the RNA 3' end from the active site. It can occur as a proofreading mechanism in response to misincorporated nucleotides, or when the EC is positioned over a weak RNA-DNA hybrid and can gain stability by backtracking onto a stronger hybrid [41-43]. Backtracked pauses can be resolved by forward translocation to restore the 3' end to the active site, by intrinsic transcript cleavage by RNAP to create a new 3' OH, or by GreA/B-stimulated cleavage, in which the Gre transcription factor binds in the secondary channel and aids cleavage of the backtracked transcript in the RNAP active site.

Pausing of the EC is a fundamental step in both the intrinsic and Rho-dependent termination pathways by stalling ECs and either decreasing the rate of pause escape from the termination site (in intrinsic termination) [25], or allowing Rho to “catch up” to the EC by translocating RNA more rapidly than RNAP synthesizes RNA (in Rho-dependent termination) [44-47].

Intrinsic terminators encode a T_{hp} structurally (Fig. 1.3A–B) and functionally similar to a pause hairpin [48]. Intrinsic terminators also encode a weak, U-rich RNA-DNA hybrid that contains many of the sequent determinants of the elemental pause [25]. The weak U-rich hybrid could cause backtracking in the absence of the T_{hp} . However, backtracking is an unlikely predecessor to termination as it would inhibit T_{hp} formation. In fact, one role of the T_{hp} is thought to be the prevention of backtracking to “lock” the EC on the weak hybrid that is favorable for dissociation [26, 49]. The U-rich hybrid at intrinsic terminators induces an elemental pause and may instigate the initial RNAP conformational change necessary for termination, by loosening

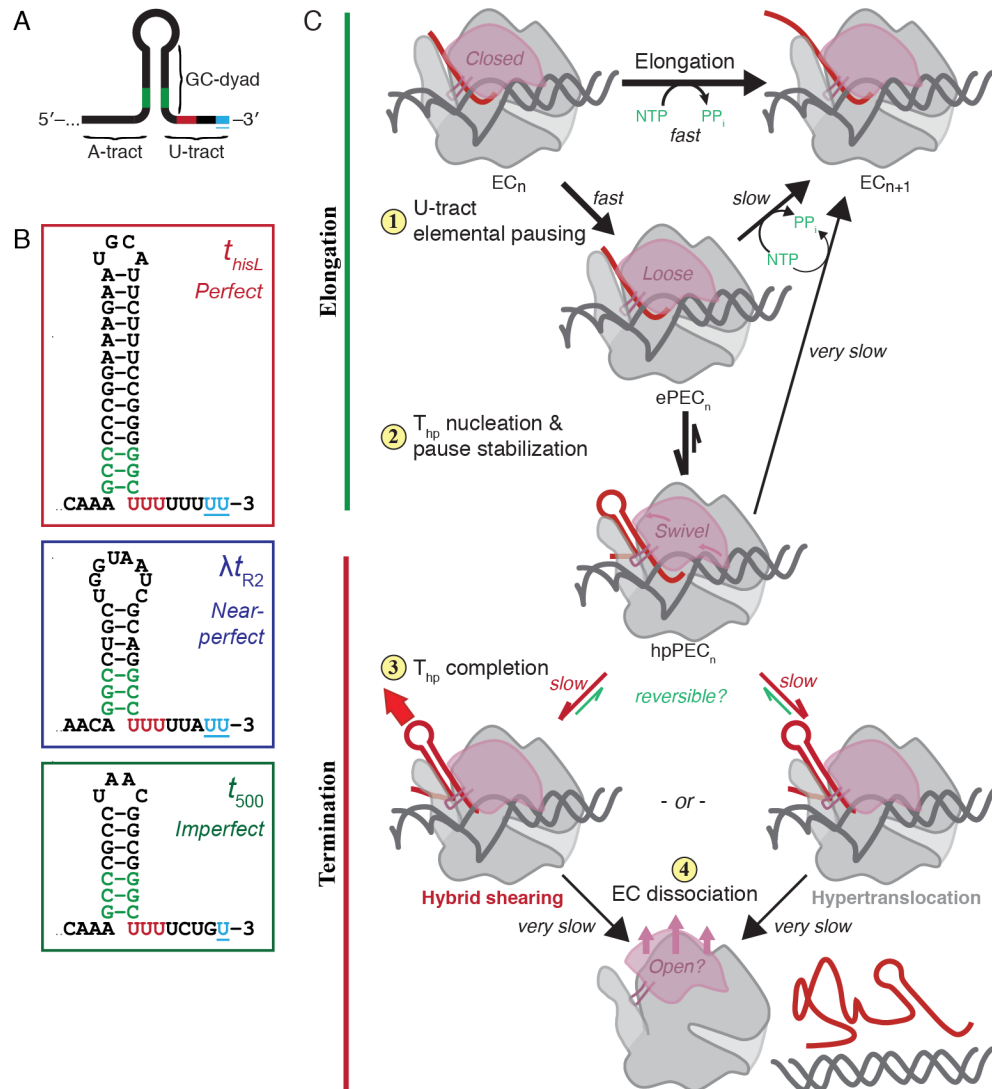


Figure 1.3. The mechanism of intrinsic termination. **(A)** Secondary structure and elements of an intrinsic terminator. Red highly conserved upstream hybrid Us that are melted upon T_{hp} completion; green, GCs base pairs that can only form upon hybrid shortening and T_{hp} completion; blue (underlined), 3' termination sites. **(B)** Representative model terminators. “Perfect” etc. represents the strength of their U-tracts. Bases are colored as in panel A. **(C)** A reaction scheme is illustrated with competing steps in the termination pathway. The EC is depicted schematically in an orientation similar to Fig. 1.1A. RNA (red); DNA (dark gray); mobile clamp domain and lid (pink); secondary channel (light gray). Two alternative mechanisms for T_{hp} completion and EC inactivation are depicted. Red arrows depict the proposed EC inactivation step, which appears to be irreversible *in vitro*; green arrows depict possible EC reactivation, which may occur *in vivo*, possibly facilitated by transcription factors. “n”, “n+1” EC subscripts denote the transcript lengths.

protein-nucleic acid contacts in the EC and making the complexes susceptible to subsequent inactivation and dissociation.

Rho termination efficiencies have been shown to be correlated to pause strength [50]. However, pause stabilization by either a pause hairpin or backtracking impedes Rho termination [51]. Whether the pauses at Rho termination sites are all elemental pauses of varying strengths, or whether they are stabilized by other means requires further characterization. A better description of these early events in both termination pathways is vital to understanding the role of pausing in intrinsic and Rho-dependent termination.

1.3 Impact of transcription factors on termination

Most mechanistic studies have been performed in cell-free environments to leverage the manipulatable nature of *in vitro* experiments. Although this focus enables investigation of the termination mechanism in the absence of extrinsic factors, it also may miss more complex layers of regulation that occur *in vivo*. Multiple factors in the cell may affect termination, including transcriptional factors that can bind ECs and modulate their stability at various steps along the termination pathway.

Transcriptional factors that bind the EC can perturb termination efficiency by modulating elongation rate, pause entry or pause lifetime, nascent RNA structure formation, or EC stability. For example, the widely distributed NusG protein has been shown to promote forward translocation (*i.e.*, elongation) of *E. coli* ECs by binding the clamp near the upstream fork-junction and likely helping upstream DNA reanneal [52, 53]. In other bacterial strains, however, NusG has been shown to stimulate hairpin-stabilized pausing (*Bacillus subtilis*) [54, 55], or to enhance intrinsic termination efficiency (*Mycobacterium bovis*) [56] possibly by effects on

translocation opposite to those seen in *E. coli*. NusG aids Rho-dependent termination at a subset of Rho-terminators [57, 58], but conversely, NusG can also occlude Rho-RNAP interactions by an alternate interaction with the ribosomal subunit, S10, that may facilitate transcriptional-translational coupling [59]. The NusG paralog, RfaH, acts by a distinct mechanism; RfaH stabilizes the closed clamp by binding the β' -clamp helices and the β -gate loop, thereby impeding clamp opening required for both pausing and termination [40, 53, 60-62]. Finally, the exit-channel binding factor, NusA, can stimulate both pausing and termination at weak intrinsic terminators by stimulating hairpin formation and stabilizing its interaction with the flap domain, effectively decreasing the rate of pause escape [23, 33, 36, 37, 39, 40]. However, like NusG's paradoxical roles in Rho-dependent termination, when complexed with other transcription and translation factors, NusA can also act as an antitermination factor to inhibit intrinsic termination (reviewed in [63]).

Gaining a complete picture of termination *in vivo* therefore requires an understanding of the abundance of various transcriptional factors in the cell, and how they impact the termination mechanisms identified in purified ECs.

1.4 The mechanism of intrinsic termination

At intrinsic terminators, DNA and RNA elements at the site of termination direct a series of steps that, after RNAP transcribes kilobases of gene information, cause the enzyme to (i) terminate transcript elongation at a discrete location spanning 2–3 nt, and (ii) release from the chromosome and transcript. The canonical intrinsic terminator (Fig. 1.3A) consists of 2 main elements: (i) a GC-rich dyad that forms the T_{hp} 7–8 nt from the transcript's 3'-end, followed immediately by (ii) a 7–8 nt U-rich tract, of which the first 3 Us (Fig. 1.3A–B, red Us) are the most highly conserved

[64]. Here we describe the steps by which this termination signal directs EC destabilization: (1) pausing at the 3' end of the U-tract, (2) terminator hairpin nucleation, (3) hairpin completion and possible EC inactivation, and (4) EC dissociation (Fig. 1.3C).

1.4.1 Elemental pausing at the U-tract

As RNAP transcribes the final nucleotides of the terminator U-tract, it pauses (Fig. 1.3C, step 1), favoring the termination pathway in the kinetic competition between elongation and termination [25, 29]. Decreased elongation rates favor termination efficiency [65] as pausing creates a window of time in which the T_{hp} can form in the exit channel and initiate the cascade of events that cause termination [25]. Mutation of the 3'-most U in the U-tract of the λ_{R2} terminator that was shown to abrogate pausing at the termination site also eliminated termination [25]; pausing at the U-tract must therefore be crucial for formation of T_{hp} . In addition to the U-tract, nucleic acid sequences in the active site, downstream DNA channel, and RNA exit channel play a role in pause entry, and determine pause efficiency and duration [27, 35, 66, 67]. These sequences contribute to the maximal termination efficiency by determining the proportion of complexes that enter the elemental U-tract pause, and the fraction of those complexes that are able to form the T_{hp} prior to pause escape (Fig. 1.3C).

1.4.2 Terminator hairpin nucleation

Pausing of the EC at the end of the 7–8 nt U-tract stalls RNAP for a sufficient time to allow the T_{hp} to form in the exit channel (Fig. 1.3C, step 2). Two separate studies have shown that a 7 bp T_{hp} does not nucleate until the EC transcribes 7 nt downstream from the T_{hp} -encoding sequence [25, 68]. This is consistent with evidence that RNAP protects 14 nt of RNA in the hybrid and

RNA exit channel [8], and with a preference for $T_{hp}S \geq 7$ bp at canonical terminators [64, 69]. Termination can also occur at low efficiency 8 nt downstream from a 6 bp T_{hp} [36], with hairpin nucleation likely occurring 8 nt downstream from the T_{hp} sequence. Conversely, $T_{hp}S$ with longer stems or with stabilized loops may nucleate before the EC reaches this position so that programmed pausing at the end of the U-tract becomes dispensable for T_{hp} nucleation [70], although this relationship has not been studied systematically. However, in the absence of a U-tract, hairpin formation does not result in efficient termination [70], verifying that both the T_{hp} and U-tract are essential for the termination mechanism. Upon T_{hp} nucleation, all but the bottom 2–3 bp of the T_{hp} rapidly pair [68], aided by the flap domain in the case of weak $T_{hp}S$ [33]. The partially-formed T_{hp} extends the duration of U-tract elemental pause through hairpin-stabilization [48, 68, 71].

Historically, some studies have described effects of sequences far removed from a terminator on termination efficiency, and have considered various explanations for the phenomenon including induction of persistent alternative conformations of the EC that resist termination [72, 73]. However, competing RNA structures formed with the upstream transcript can inhibit T_{hp} formation by sequestering the upstream arm of the T_{hp} prior to hairpin nucleation [68, 74]. These more recent demonstrations of the effects of alternative RNA folding highlight the importance of ensuring that comparisons of terminators are made within a constant sequence context to avoid potentially complicating effects of RNA folding.

1.4.3 Terminator hairpin completion

The structure of the EC suggests that pairing of the bottom 2–3 bp of the T_{hp} stem would require unstacking of the lid from the –10 RNA-DNA base pair (numbered according to the RNA in the pre- and half-translocated states; Fig. 1.2) [10]. Half-translocation of the RNA at a hairpin-

stabilized pause, as occurs during T_{hp} formation, does not appear to be sufficient to disrupt these interactions; although the lid shifts upstream to accommodate the -10 RNA-DNA base-pair, it does so without disruption of the electrostatic and aliphatic interactions that stabilize the EC [22, 23, 48]. Consistent with this idea, T_{hp} formation has been found to occur in two steps for the λt_{R2} T_{hp} [68]. In the first step, all but the bottom 2 bp formed upon T_{hp} nucleation 7 nt downstream from the T_{hp} -forming sequences. In the second step, the bottom 2 bp of the T_{hp} paired when the EC transcribed the 8th nucleotide downstream from the T_{hp} sequence. This energetic barrier to completion of the T_{hp} likely reflects the need to disrupt the lid:hybrid and Sw3: -10 RNA interactions that limit complete hairpin extension upon nucleation of the T_{hp} (Fig. 1.3C, steps 2 & 3).

Termination is frequently observed 7 nt downstream from the T_{hp} on endogenous terminators (Fig. 1.3A–B, blue Us) and requires that T_{hp} nucleation and completion occur at the same template position [25, 75], inconsistent with the findings of the T_{hp} folding study described above [68]. It is possible that the use of stabilized hybrids to enable EC “walking” downstream from the T_{hp} sequence precluded T_{hp} completion until the 8th nucleotide was added [25, 26, 68]. Upstream melting of the stabilized hybrid would be disfavored in this case due to the lack of Us, thereby making the energetic barrier for hairpin completion even higher and preventing T_{hp} completion at this position. However, the two-step folding pathway of T_{hp} nucleation followed by energetically expensive T_{hp} completion likely exists for all cases, independent of terminator structure and the exact position for each step, due to the energetic barrier created by the lid-hybrid and other protein-NA interactions.

T_{hp} completion causes 3–4 bp of the upstream RNA-DNA hybrid to melt (Fig. 1.2 & 1.3B, red Us) [25, 26], thus contributing to destabilization of the EC. This hybrid melting step is

essential for hairpin completion and termination, consistent with the high degree of sequence conservation for Us at the first 3 positions in intrinsic terminator U-tracts [25, 26, 64]. From these observations, it appears that the length of the hairpin stem and the structure of the T_{hp} loop dictate the kinetics and position of T_{hp} nucleation by determining the position at which the loop and GC-rich dyad emerge from the RNA exit channel and the amount of RNA needed to nucleate T_{hp} formation. On the other hand, the template position at which T_{hp} completion can occur is likely determined by the relative strengths of the bottom 2–3 bp of the T_{hp} versus the upstream hybrid, which will dictate the energetic cost of melting the upstream hybrid and breaking RNAP-NA interactions versus pairing of the bottom of the T_{hp} .

Several different models have been proposed to describe how the completed T_{hp} affects the EC. These models involve (i) hypertranslocation of the EC 2–4 bp downstream from the termination point without concomitant nucleotide addition (Fig. 1.2 & Fig. 1.3B) [29, 74, 76], or (ii) shearing or slippage of the RNA from the RNA-DNA hybrid by upward “lifting” or “rotational wrenching” of the hairpin [26, 74]. A single-molecule force-clamp experiment revealed that both mechanisms may be possible, depending on the sequence of the terminator [74]. A terminator with an imperfect U-tract (a U-tract interrupted by one or more G-C bp; e.g., t_{500}) was found to be aided by RNA-pulling, and impeded by hindering force on DNA translocation (consistent with hypertranslocation), whereas two terminators with near-perfect U-tracts (7 or more tandem rU-dA bp and at most one rA-dT bp; e.g., t_{his} and λt_{R2}) were also aided by RNA-pulling, but unaffected by force on the DNA (consistent with hybrid shearing). Taken together, these data suggest that either mechanism can occur, with hypertranslocation favored at imperfect U-tracts and hybrid shearing favored on weaker, near-perfect U-tracts. Consistently, an analysis of a set of 100 *E. coli* terminators found that the downstream DNA is enriched for AT-

rich sequences at positions +10 to +12 after imperfect U-tracts, but not after near-perfect U-tracts [77], indicating a need for melting of downstream DNA in hypertranslocating terminators but not in hybrid-shearing terminators. These results support a composite hybrid-shearing/hypertranslocation model in which the mechanism employed is determined by the relative energetic cost of shearing the U-tract hybrid *versus* melting downstream DNA (Fig. 1.2 & Fig 1.3B, step 3) [74, 78]. Extensive characterization of various terminator structures *in vivo* identified the strongest terminators to have both near-perfect U-tracts and AT-rich downstream DNA [69]. This finding also supports the idea that both hybrid shearing and hypertranslocation can occur, and that termination efficiency is highest if both mechanisms are energetically favorable.

T_{hp} completion thus appears to be the first step that distinguishes paused complexes from termination-competent intermediates. It occurs in a second step separate from T_{hp} nucleation, likely impeded by RNAP-NA contacts such as the lid-hybrid interactions that must be broken to allow pairing of the bottom 2–3 bp of the T_{hp} stem. Concomitant melting of the weak 3–4 rU:dA bp in the upstream end of the hybrid is required to accommodate the completed T_{hp} , either by hypertranslocation of RNAP, or by hybrid shearing of RNA depending on the terminator sequence. The resulting shortened hybrid then favors eventual EC dissociation by partial destabilization of the complex. The combination of these nucleic-acid rearrangements in the EC and likely concerted rearrangements of RNAP is crucial for efficient termination. It is thought to cause irreversible inactivation of the EC, and consequent commitment to the termination pathway.

1.4.4 EC inactivation and commitment to termination

Although several studies model termination efficiency as being determined by direct kinetic competition between elongation/pause escape and EC dissociation [79, 80], there is some evidence (as well as uncertainty) that an irreversible EC inactivation step occurs in the termination pathway prior to EC dissociation. A single-molecule termination assay found that a “terminal dwell” occurred at the t_{hisL} terminator prior to transcript dissociation only for complexes that terminated, but not for complexes that read through the terminator [81]. This result suggests that nucleotide addition is irreversibly halted in a step distinct from dissociation. In contrast, no terminal dwell was observed for the t_{his} , λt_{R2} , or t_{500} terminators in a force-clamp experiment [74]. It is possible to explain this discrepancy if the assisting force used in the force-clamp study increased the rate of EC dissociation, masking a terminal dwell prior to dissociation. The terminal dwell then implies that ECs at the termination site can exist in two distinct states: one that is able to recover from the pre-termination pause and continue elongation past the terminator, and another inactivated state that is incapable of nucleotide addition but remains bound to the hybrid at the termination site until EC dissociation occurs at a slower rate. Notably, the existence of an irreversibly “trapped,” or inactivated, termination complex at the λt_{R2} terminator prior to EC dissociation has been reported [25]. The rate of transcript release *in vitro*, after T_{hp} completion and upstream hybrid melting, occurs on the order of many minutes [26], too slow to compete with pause escape, which can occur on the order of seconds even when stabilized by a pause hairpin [28]. Given this slow rate of EC dissociation, an irreversible inactivation step seems kinetically necessary to prevent pause escape and resumption of elongation. This step, requiring inactivation of the RNAP active site, represents “commitment” of ECs to the termination pathway.

EC inactivation might occur by restructuring of the active site. All proposed models for commitment posit T_{hp} completion as a driving force for downstream hybrid rearrangements that result in EC inactivation [25, 26, 74]. A prominent model for inactivation of ECs is suggested by the composite shearing/hypertranslocation model for T_{hp} accommodation; both hybrid shearing and hypertranslocation would directly inactivate the EC via removal of the RNA 3'-end from the RNAP active site (Fig. 1.2B & 1.3B) [74]. Understanding the mechanism for EC inactivation and whether the nucleic-acid rearrangements described by any of these models are (i) sufficient to prevent nucleotide addition and (ii) irreversible will require further study.

Mechanisms that invoke irreversible commitment to termination require that all regulation of termination efficiency occur at a step prior to inactivation of the EC (because events occurring after irreversible inactivation cannot affect the proportion of ECs that terminate). However, studies into the role of the A-tracts immediately upstream of the T_{hp} (Fig. 1.3A) These A-tracts were found to increase termination efficiency, likely by pairing with the U-tract since complementarity of A-tract and U-tract was required to observe the effect. A:U-tract pairing would extend the T_{hp} duplex and completely eliminate the RNA-DNA hybrid; it must therefore occur subsequent to T_{hp} completion and the presumed EC inactivation. This logic implies that, *in vivo*, T_{hp} completion must be reversible, and that A:U-tract pairing can compete with EC reactivation (Fig. 1.3C, green arrows) by accelerating the dissociation of RNA from the EC (Fig. 1.3C, step 4).

This A-tract-dependent enhancement of termination efficiency has not been observed *in vitro* [36]. One interpretation is that in purified ECs, the rate of EC reactivation after T_{hp} formation is so slow relative to the dissociation rate that it is functionally meaningless, rendering any downstream A:U-tract pairing effect irrelevant to termination efficiency. However, cellular

elongation factors such as NusA and NusG that interact with the EC may be capable of binding the inactivated intermediate, facilitating EC reactivation. This could allow competition between dissociation and EC reactivation (Fig. 1.3C, step 4 *versus* step 3, green arrows), allowing acceleration of dissociation by A–U-tract pairing to affect termination efficiency. In this view, the absence of transcription factors in the *in vitro* study explains the lack of an A-tract effect on termination efficiency [36]. Potential mechanisms of transcription factor-mediated EC reactivation include: (i) NusA-mediated stabilization of the flap- T_{hp} interaction in the inactivated complex [33], or (ii) NusG-mediated stabilization of the clamp-DNA interaction. Either interaction could facilitate EC reactivation by affecting nucleic-acid pairing patterns in the T_{hp} , the upstream hybrid, and the upstream DNA fork-junction. *In vitro* testing of the effect of A-tracts on termination efficiency and EC dissociation in the presence of NusA, NusG, and other EC-stabilizing factors could help determine whether these factors sufficiently stabilize termination intermediates to make an otherwise irreversible EC inactivation step reversible.

1.4.5 EC dissociation

The final step in the termination pathway is dissociation of RNAP from both the DNA template and the RNA transcript. Several factors likely contribute to dissociation. The RNA-DNA hybrid is significantly weakened by the rU–dA bps of the terminator U-tract [82], and destabilized further by shortening of the hybrid after upstream hybrid melting [25, 26]. A 5-bp hybrid, representative of the residual hybrid length retained by the terminator after the upstream hybrid melts (Fig. 1.2), has been demonstrated to almost completely recapitulates the dissociation rate of a termination complex [26]. T_{hp} completion must disrupt the RNAP-NA interactions described above and likely also the H-bonds between ssRNA and the exit channel [10]. It is unclear

whether any of these contacts could be re-established upon hypertranslocation or hybrid shearing and consequent repositioning of the RNA with respect to RNAP.

An open-clamp conformation with loosened contacts between the RNA-DNA hybrid and switch regions also is proposed to favor EC dissociation [13, 14, 18, 21]. Collapse of the transcription bubble and reannealing of the template and non-template strands appears to be incompatible with the narrow width of the main cleft in the closed- and swiveled clamp conformation. Disulfide crosslinking studies that report clamp state indicate that a displacement of the clamp from the closed-clamp state favors duplex formation in the exit channel [16, 17]. This was at first thought to suggest that that pause hairpin formation stabilized the open clamp and weakened contacts between the nucleic-acid scaffold and RNAP. Though the recent discovery of the swiveled-clamp conformation casts doubt on this particular interpretation, [22, 23] the notion that clamp opening would facilitate the termination process remains an appealing hypothesis.

Other RNAP rearrangements have been proposed to aid EC dissociation based on the reduced stabilities of ECs deleted for the lid, rudder, or flap modules [33, 37]. However, the conformational rearrangements of RNAP in a destabilized termination complex remain poorly characterized. Further, the order of RNA and DNA release, and whether the order is obligate or affected by terminator sequences, remains unknown.

1.4.6 Additional roles of multiple RNAP conformations in termination

A surprisingly under-studied aspect of the termination mechanism is the contribution of conformational changes of RNAP to the energetics of each step in the termination pathway, leaving many questions unanswered. Are the steps of termination governed principally by

nucleic-acid folding and translocation dynamics, or are conformational changes of RNAP also necessary to inactivate and dissociate the enzyme? The flap tip has been found to aid nucleation of weak T_{hpS} [33]. Restructuring of the clamp [13, 14, 16, 17] and lid are thought to be involved in EC dissociation [37]. Do other modules also play a role in the pathway? What inactivates the EC active site at terminators? Is removal of the RNA 3' end from the active site (as expected in hypertranslocation and hybrid shearing; Fig. 1.2 & Fig. 1.3C) irreversible on the time scale of dissociation, or is active-site restructuring at the protein level also essential for efficient termination?

As discussed above, the switch regions are thought to couple clamp closure to nucleic-acid scaffold binding, and conversely, could promote efficient termination by destabilizing contacts to the scaffold upon clamp opening [18, 21]. Switch rearrangements are therefore likely to be involved in EC dissociation. Mutational analysis of RNAP identified several mutations in the switch regions with altered termination efficiencies [83, 84], consistent with a role of the switch regions in the termination mechanism (presumably by mediating clamp opening). Such an effect on termination efficiency *in vitro* implies that switch remodeling occurs upstream of EC inactivation, perhaps by clamp loosening that aids both an elemental pause and T_{hp} formation (Fig. 1.3C, steps 1 and 2) [21]. Consistent with this view, programmed transcriptional slippage in ECs positioned on a hybrid with U_5C_4 RNA was recently reported to occur with ~50% efficiency downstream of an RNA hairpin [70]. Sliding of the RNA in the main cleft of the enzyme is presumably facilitated by (i) a shortened RNA-DNA hybrid due to exit-channel duplex formation and resultant melting of the upstream U-rich hybrid, and (ii) loosened interactions between the hybrid and switch/clamp domains.

Roles of the active-site proximal TL (Fig. 1.1B) in termination have also been suggested. In one model for termination it was suggested that folding of the active-site proximal trigger loop was required to form a termination-competent complex, possibly inactivating the active site [85]. A recent kinetic analysis that successfully disentangled the rates of elongation from termination instead found that stabilizing the TL motif in either its folded or unfolded states slowed the rate of termination relative to a wild-type TL that was free to explore multiple distinct conformations [48]. Based on these results, the multistate-multipath (MS-MP) model of termination was proposed as a modification to the original elongation *versus* termination thermodynamic model [86]. Under MS-MP termination, inactivated ECs are free to sample multiple conformational states with fluctuating protein-nucleic acid contacts. Each distinct state has its own free energy barrier to termination, weaker in some states than in others. Variable EC stabilities yields multiple “parallel” conformational paths to termination, meaning that no one specific EC conformation is required. Under MS-MP, the ability to sample multiple conformations through both large-scale movements and small-scale Brownian motion will contribute to termination.

The kinetic analysis discovered MS-MP was exclusively focused on TL, and it remains an outstanding question as to whether other RNAP modules and motifs also promote termination through the extended thermodynamic model. Investigating how conformation changes by the remaining RNAP elements in the inactivated EC is therefore central to completing our understanding of the energetic contributions of RNAP rearrangements to stabilization of the inactivated complex and dissociation.

Recent advances in cryo-EM methods have added to our pre-existing retinue of biochemical tools used to study the termination process, and it is likely that determining the

structures of the family of inactivated ECs will soon be possible. Rather than requiring isolation of metastable complexes, rapid freezing can be used to kinetically trap intermediates and distributions of intermediates can be visualized in single time points [87]. Obtaining such information has the potential to define the identity and stability of each intermediate in the termination pathway and clarify the relative contributions of both RNAP and nucleic-acid scaffold conformations to intrinsic termination.

1.5 Rho-dependent transcription termination

In contrast to intrinsic terminators that function in response to signals directly encoded at the end of a transcription unit, Rho-dependent termination requires the RNA translocase activity of the homohexameric RecA-family helicase Rho (Fig. 1.4), which non-specifically recognizes and binds signals on the nascent RNA transcript. After binding to primary recognition sites, RNA is threaded through Rho's central pore to trigger translocase activity. Dissociation of RNAP from the transcript occurs after Rho translocates all RNA upstream of the EC through the pore. Here we will discuss recent work on Rho-dependent termination that has increased our understanding of the signals and sequences necessary for recruitment and activity of the Rho helicase, how the structure of Rho dictates its activities during termination, and how the structure and composition of the bacterial nucleoid can enhance the efficiency of Rho-dependent termination *in vitro*.

1.5.1 Recognition of Rho-utilization elements

Termination by Rho is regulated by sequences within the template DNA and in the nascent RNA transcript that can affect at least one of three steps: (i) recognition by Rho of substrate RNA; (ii) translocation of RNA through the Rho motor protein; and (iii) entry of the EC into a paused state

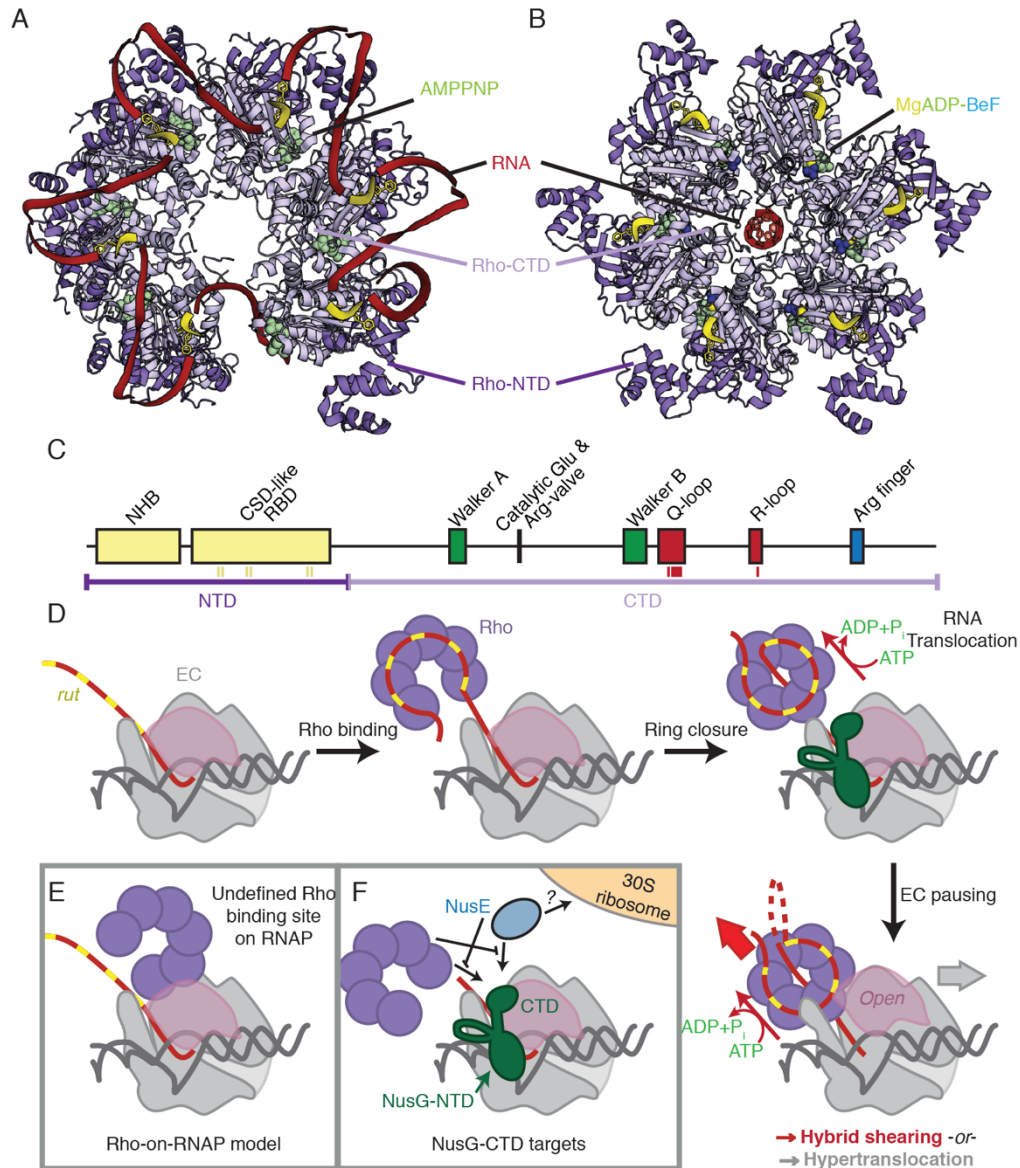


Figure 1.4. Structures of Rho and mechanisms of Rho termination. **(A)** A model of open-ring Rho (derived from PDB 1pvo, provided by James Berger) bound by *rut* RNA at the primary binding sites and the non-hydrolyzable ATP analog AMPPNP. The RNA dinucleotides directly bound to the primary binding sites are highlighted in yellow. **(B)** Structure of closed-ring Rho (PDB 3ice) bound by RNA in the secondary binding site, Mg²⁺ and the ATP transition state analog ADP-BeF₃. Primary binding site dinucleotides are modeled in; *rut* RNA loops are omitted for clarity. **(C)** Domain and module locations in *E. coli* Rho (colored, where relevant, based on interaction targets in panels A and B). NHB, N-terminal helical bundle (yellow, YC-motif in primary RNA site); CSD-like RBD, cold-shock domain-like RNA binding domain (yellow); Walker A&B, ATP-binding motifs (green, ATP/ADP); Q- and R-loops, secondary binding site loops (red, RNA secondary binding site); catalytic Glu, Arg valve, & Arg finger, catalytic residues (blue). Yellow and red lines indicate residues implicated in RNA binding. Dark and light purple NTD and CTD correspond to domains in panels A and B. **(D)** The pathway to Rho-dependent

termination. Upon contact with the EC, Rho extracts the nascent transcript out the RNA exit channel by hybrid-shearing or hypertranslocation (red and gray arrows) and possibly RNAP rearrangement (*e.g.*, clamp opening).

(E) Possible Rho binding to RNAP throughout transcription. Such an initial state would not necessarily alter steps in Rho termination, but could affect RNAP rearrangement. (F) Rho and antitermination factor are both targets of the NusG-CTD. Binding of NusE could facilitate coupling of transcription and translation, which would sequester *rut* sites in the ribosome and inhibit Rho termination.

on which Rho can act. Subsequent to these three steps, Rho induces termination in a fourth step, (4), via a mechanism that involves extraction of the nascent transcript from the RNA exit channel followed by bubble collapse and template DNA release.

The first step requires binding of Rho to the nascent RNA transcript after it emerges from the RNA exit channel of RNAP. Rho preferentially binds to C-rich, G-poor sequences that lack extensive secondary structure, with the highest affinity for poly(C) RNA sequences [88]. Binding of Rho to these sequences is necessary to stimulate Rho's RNA translocase activity.

Endogenous Rho binding elements, referred to as Rho-utilization (*rut*) [89] sites [89], lie upstream of the site of termination and are generally thought to be ~80–90 nt long [90–92]. This length is sufficient to allow for binding of RNA to each of Rho's six monomers: two pyrimidines tightly bound to the enzyme, interspersed by loops predicted to be at least 7–8 nt, followed by a length of RNA sufficiently long to be threaded into the secondary binding site (Fig. 1.4A) [93]. Depletion of guanine within a natural *rut* site minimizes the formation of potentially interfering secondary structures, which are generally inhibitory to Rho binding [3, 94]; *in vitro*, short oligonucleotides complementary to *rut* RNA significantly reduces termination efficiency [95]. Rho has also been shown to be capable of looping out extensive secondary structures so long as they do not sequester the primary *rut* RNA–Rho contacts [96–98].

A consensus full-length *rut* sequence is not thought to exist, as Rho activity can be stimulated on diverse, C-rich RNAs. Indeed, termination still occurs at the model Rho terminator λ_{R1} even after its natural *rut* site has been replaced with an artificial, C-rich sequence [99]. This lack of a consensus sequence hinders bioinformatic identification of *rut* sites, but efforts to identify sites of *E. coli* Rho activity *in vivo* using the Rho-specific inhibitor bicyclomycin have

given insight into Rho action on a genomic scale: >1,250 putative sites of Rho action were identified in this manner [2].

Biochemical studies, near-atomic resolution crystal structures, and EM have identified the sites on Rho to which *rut* RNA binds. Both crystallographic and EM studies suggest that *rut* RNA binds the NTD of Rho in an open-ring state (Fig. 1.4A) [100, 101], which guides RNA into the exposed central pore of the Rho-CTD. Binding of RNA within the ring then isomerizes Rho into a closed-ring state (Fig. 1.4B) that is competent for ATP-hydrolysis and translocation [102, 103]. These structures are consistent with biochemical studies that indicate the Rho-NTD contains the primary RNA-binding site, a name accurate both in that *rut* RNA binds this site first and that it is the source of the high-affinity Rho–RNA interaction [100, 104]. Rho binds C-rich sequences with a $K_d \approx 0.01\text{--}10$ nM, an extensive range dependent upon the sequence of the RNA ligand and experimental conditions [105-110]. The NTD contains a cold shock domain-like RNA-binding domain (CSD-like RBD) whose OB-fold creates a narrow cleft that can only accommodate pyrimidines, with a preference for cytosine (Fig. 1.4C) [100, 111]. A “consensus” of 5'-YC-3' was identified; it is not yet understood whether each of Rho's six monomers must bind the *rut* element or if binding to only a subset is sufficient to induce RNA entry into the pore, ring closure, and translocase activity. In addition to the primary binding site, the Rho-NTD contains a positively-charged N-terminal helix bundle (NHB) that may assist in binding of RNA to the enzyme via weak electrostatic interactions with the negatively-charged phosphate backbone [112].

Rho in many bacterial phyla contain bulky sequence insertions between the NHB and the CSD-like RBD that are not found in *E. coli* Rho. The functions of these NTD insertions, which are poorly conserved in both sequence and length, are not understood [111]. Recent studies on

Mycobacterium tuberculosis Rho have demonstrated that its NTD insertion may enhance the affinity of *Mtu* Rho to *rut* RNA while simultaneously slowing Rho motor activity [113]. Whether all such insertions, which vary among lineages, function in this manner merits serious study.

The Rho-CTD contains the RNA secondary binding site. A high-resolution structure of Rho in the closed-ring conformation (Fig. 1.4B) shows the nascent RNA transcript threaded through the central pore. This structure reveals, in agreement with biochemical evidence, that the secondary binding site is comprised of the conserved Q- and R-loops, which form contacts with 6 nucleotides of the RNA substrate in a “spiral staircase” conformation [102, 114]. These loops are structurally analogous to the L1 and L2 loops found in other RecA-family helicases [115]. Unlike the primary site, which requires a cytidine-rich substrate, the secondary site is capable of binding RNA substrates of any sequence. However, ATPase assays show a clear preference for poly(C) [116]. The structure also explains Rho’s specificity for RNA; whereas the primary binding sites can recognize bases from either DNA or RNA [117], the carbonyl of residue V284 in the Q-loop of *Eco* Rho forms a H-bond with a 2'-OH of the RNA ribose moiety, making binding of DNA to the secondary site unfavorable [102, 118]. In addition to the secondary binding site, the Rho-CTD also contains the Walker A and B motifs (commonly referred to as the P-loop) that form the ATP-binding pocket. Activation of catalytic water and subsequent ATP-hydrolysis is mediated through interactions between the catalytic glutamate, arginine valve, and arginine finger (*Eco* Rho E211, R212, and R366 respectively; Fig. 3C) [102, 119].

A major unresolved question involves the possibility of Rho pre-binding the EC early in the transcription cycle, prior to appearance of *rut*-containing nascent RNA. ChIP-chip data shows a genome-wide correlation between Rho and EC occupancy, with Rho associating with ECs early in the transcription unit without triggering termination [120]. The precise nature of this

interaction is undefined. Is this a strong, direct interaction between RNAP and Rho, or a transient interaction with minimal physiological relevance? If Rho pre-binds the EC, the local concentration of Rho would be significantly increased, poisoning the system for quick capture and binding of a *rut* site upon its appearance in untranslated RNA [121]. If such a Rho–RNAP interaction exists, it should be possible to identify the interaction sites on both RNAP and Rho that, upon alteration for instance by genetic mutation, would perturb Rho termination both *in vivo* and *in vitro*. Despite a long history of Rho and RNAP genetics such mutations have yet to be identified; approaches using crosslink mapping and directed mutagenesis should readily enable the necessary test.

1.5.2 Translocation towards the elongating EC

Binding of Rho to both *rut* RNA and ATP induces ring closure, activating Rho for catalysis (Fig. 1.4D) [102]. Rho translocates the nascent RNA via ATP-hydrolysis; although all four NTPs can support Rho translocation [122], ATP binds and drives translocation most efficiently with a secondary preference for GTP over C/UTP. Structural and biochemical studies suggest that the asymmetrical, spiral-staircase conformation of closed-ring Rho promotes stepping of 1 nt per molecule of ATP hydrolyzed [102, 123]. An asymmetric model of translocation, in which each Rho monomer establishes a different set of protein-RNA contacts throughout the cycle of ATP binding, hydrolysis, and PP_i release, explains how Rho drives translocation of RNA [102, 123]. Typically, the site of termination is within 10–20 nt of the *rut* site [93], and is rarely more than 100 nt downstream. Translocation can be inhibited by structural roadblocks such as RNA hairpins, ribosomes engaged in active translation, and other RNA-binding factors. It has been shown that Rho's helicase and translocase activities are sufficient to displace streptavidin from a

biotinylated RNA. These forces may be sufficient to melt certain interfering RNA secondary structures and displace a subset of RNA-binding proteins [124].

During translocation, *rut* site RNA remains stably bound to the primary binding sites, producing an RNA loop that grows in size between the primary and secondary binding sites [93, 123, 125]. This model of Rho translocation, known as the tethered-tracking model, was proposed as an alternative to one in which these primary RNA-Rho contacts were released during translocation. Initial evidence supporting the tethered-tracking model was obtained when it was observed that active translocation by Rho extended the size of an RNA footprint during nuclease digestions [126]. Further support came from single-molecule force clamp experiments in which force extension curves of nascent RNA bound by active Rho were collected. Upon the addition of Rho, abrupt, discontinuous changes in extension were observed that corresponded to the displacement of Rho from RNA under a mechanical load. Importantly, these abrupt “rips” in extension were proportional in size to the length of the transcribed template, strongly implicating tethered tracking in translocation by Rho [93].

1.5.3 Kinetic coupling of Rho to the EC

Rho is thought to induce termination when it contacts the RNA exit channel. Because the EC continues active elongation even as Rho engages the *rut* site and begins translocation, there is a competition between the translocation rates of the two enzymes. This competition is referred to as kinetic coupling, and the fact that elongation-deficient or pause-susceptible RNAPs show increased Rho termination efficiencies supports the model [50]. It follows that ECs composed of wild-type RNAPs will be most susceptible to Rho at long-lived pause sites. This has been verified for a wide assortment of both natural and artificial pause sites, although not all pauses

support efficient termination by Rho [44-47]. For example, the well-studied hairpin-stabilized *his* pause is resistant to Rho termination [51], likely because the interaction between the hairpin and RNAP prevents Rho from extracting the nascent RNA from the RNA-DNA hybrid. Similarly, backtrack-stabilized pauses are poorly terminated by Rho [51] since reverse displacement of the RNA 3' end from the active site by several nucleotides would require Rho to perform more work to first pull the RNA back into register, and then to extract the transcript from the hybrid. It remains to be determined whether most sites of Rho termination are elemental pauses, stabilized neither by a pause hairpin nor through extensive backtracking. Importantly, a single *rut* site often precedes more than one termination site; for example, $\lambda_{T_{R1}}$ contains three major sites of termination, mapping to three pause sites from which the EC can escape before Rho can dissociate the EC [127].

1.5.4 The mechanism of RNA release and EC dissociation

If Rho's ability to act as an RNA-DNA helicase [128] were sufficient for termination, then any RNA polymerase paused downstream of a *rut* site should be susceptible to dissociation by Rho. In addition to bacterial RNAP, Rho can efficiently terminate T7 RNAP and eukaryotic RNAPII ECs [129, 130]. However, Rho cannot induce dissociation of ECs formed with eukaryotic RNAPI or RNAPIII [130], suggesting that Rho activity depends upon the establishment of specific interactions with certain ECs. Termination by Rho is likely accommodated by structural changes in both the nucleic-acid scaffold and RNAP itself. Two models for termination by Rho, similar to those proposed for intrinsic termination, are feasible: hypertranslocation and hybrid-shearing. Both mechanisms rely upon Rho's ability to exert a strong force via its motor activity [124] in addition to its helicase function. These activities are thought to supplant the role of the

T_{hp} during intrinsic termination and should proceed similarly regardless of whether Rho binds *rut* RNA or RNAP first.

During hypertranslocation, Rho exerts a pushing force that drives RNAP forward on template DNA without nucleotide addition, destabilizing the transcription bubble and driving RNA release and bubble collapse [6]. Hypertranslocation is supported by the finding that Rho termination is inhibited when the upstream fork-junction cannot reanneal [6], a phenomenon that generally impairs forward translocation by an EC. It has also been observed that Rho can promote forward translocation of an EC by an additional 2 bp even in the presence of a DNA-binding protein that functions as a roadblock [6]. However, actual hypertranslocation in which the 3' end of the RNA has been displaced from the active site has yet to be directly observed.

During hybrid-shearing, Rho translocates until it reaches the RNA exit channel at the EC. Continued ATP-hydrolysis by Rho exerts a pulling force on the now-taut transcript, shearing the RNA-DNA hybrid [131]. Rho-driven hybrid-shearing would require strong force since Rho termination most frequently occurs at non-U-tract pauses, where the hybrid will contain stronger base pairing than found at the U-tract pauses of an intrinsic terminator. Although Rho can generate enough force to shear streptavidin from a biotinylated bead [124], as discussed above, it is unknown if this is sufficiently powerful to shear a non-U-tract hybrid stabilized by interactions with RNAP.

Whereas the sequence of the RNA-DNA hybrid has been implicated in dictating which pathway to release occurs during intrinsic termination, it is unresolved whether this is the case for Rho termination. One view could be that, due to the enhanced hybrid strength of the non-U-tract pause, Rho generally drives hypertranslocation of the EC, and that this is usually sufficient to induce termination. However, should a strong DNA roadblock prevent hypertranslocation of

the EC, then Rho's helicase activity may instead drive hybrid-shearing. It will be highly advantageous to identify DNA templates and conditions that allow detailed study of these two mechanisms.

Conformational rearrangements within RNAP itself are also thought to be necessary for efficient Rho-dependent termination [121]. Many of the conformational rearrangements proposed to occur during intrinsic termination would likewise destabilize the EC during Rho-dependent termination. Experiments that assay the positions of RNAP mobile domains such as the clamp, switches, and the active site proximal trigger loop and bridge helix will be required to elucidate the roles of RNAP conformational changes during Rho-dependent termination, and whether or not Rho also acts through a MS-MP thermodynamic mechanism.

1.6 NusG acts both as a positive and negative regulator of Rho

NusG, first identified as a component of the λ N antitermination complex [132], has since been identified as the only universally conserved transcription factor. Spt4/5, its archaeal/eukaryotic paralog, is also known as DSIF in metazoans [133] and is known to act as a positive elongation factor in yeast [133, 134]. In *E. coli*, NusG associates with most ECs *in vivo* [120] and enhances the overall rate of elongation [135]. The NusG-NTD directly binds RNAP through an interaction with the β' clamp helices, near the upstream fork-junction [136]. The NusG-CTD, containing a KOW domain, has two known targets: the 30S ribosomal subunit S10 (also known as NusE) [59] and the Rho hexamer [59, 137]. These distinct and mutually exclusive targets [59] permit NusG to act as both a negative and positive regulator of Rho-dependent termination (Fig. 1.4F).

The NusG-S10 interaction has traditionally been thought to allow for physical coupling of transcription and translation by the leading ribosome in bacteria [59]. During coupled

transcription-translation, ribosomes physically occlude the *rut* site during transcription, preventing Rho binding, activation, and termination activities [138]. Under conditions of active translation, Rho can only bind to *rut* sites upstream of the lagging ribosome, suppressing activity of Rho against the EC. However, when ribosomes become stalled or dissociate from the nascent RNA, *rut* sites will be exposed, allowing for Rho binding and termination. This mechanism explains transcriptional polarity, the decrease in the expression of distal genes in an operon resulting from upstream interruptions to translation. Inefficient translation or nonsense mutations that arise as a result of transcriptional error will inhibit transcriptional/translational coupling and consequently induce release of the ribosome from the EC by breaking of the NusG–S10 interaction. The exposed *rut* site can then be bound by Rho, leading to EC dissociation. Recent structural studies call this specific model for coupling into question. In one study, ribosomes and ECs were reconstituted together and the complex’s structure was resolved through cryo-EM [139]. Surprisingly, S10’s position within the 30S ribosome was far too distant from the known NusG binding site on RNAP to permit simultaneous association. Instead, physical coupling of the two complexes appeared to be mediated directly through RNAP-ribosomal interactions, the tightest of which was between the CTD of RNAP’s α subunit and helix 39 of the 16S rRNA. In a second study, the cryo-EM structure of RNAP bound to the 30S subunit in the absence of nucleic acids determined a wholly different set of contacts between the two complexes but agreed that one NusG protein could not simultaneously bind both RNAP and 30S. *In vivo* data nonetheless demonstrates a strong interaction between NusG and the 70S ribosome [140] and so the role of a NusG–S10 interactions during coupling remains unresolved. One possibility is that NusG serves as a “coupling initiation” factor. By binding RNAP and then S10 (or vice versa), the two

complexes would be brought into close proximity together. Persistent coupling could then be achieved through a rearrangement that involved NusG release from the ribosome.

Despite its potential ability to promote *rut* site sequestration by the ribosome, and its ability to increase the overall elongation rate of the *E. coli* EC (therefore decreasing the kinetic window during which Rho can act at any given pause site), NusG can also act as a positive Rho termination factor. When NusG is present with Rho during in vitro transcription assays, termination occurs much earlier within the transcription unit [57, 58]. The NusG-CTD binds Rho with a $K_d \approx 12$ nM [129] at a much higher affinity than the NusG:S10 interaction ($K_d \approx 50$ μ M) [59]. In fact, many Rho-dependent terminators *in vivo* were shown to be NusG-dependent [2, 57]. When the nucleotide content of these NusG-dependent Rho terminators was analyzed, it was discovered that they contained a lower C/G ratio [2], suggesting that *rut* sites at NusG-dependent Rho terminators may more often form secondary structures. The simplest hypothesis would be that, in ECs uncoupled from the ribosome, exposing *rut* RNA, NusG activates Rho termination by enhancing recruitment of Rho to *rut* sites that are partially occluded by RNA secondary structure formation due to subpar sequences. The finding that NusG does not enhance the apparent rate of ATP-hydrolysis by Rho as measured through an RNA-DNA helicase activity appears to support this proposal [58], though these experiments were not performed in the context of an elongation complex.

However, the hypothesis is contradicted by several previously established biochemical results. Richardson and co-workers found that NusG does not affect the half-maximal concentration of Rho required for termination at a NusG-dependent terminator within the *lacZ* gene [141]. Sen and co-workers have also argued that NusG stimulates Rho at the dissociation step, not the recruitment step [137, 142]. Additionally, if it can be shown that Rho is permanently

associated with the EC, even prior to transcription and extrusion of the *rut* site, a role in NusG-mediated tethering of Rho to the EC would seem unnecessary. It is possible that, at NusG-dependent terminators, NusG stimulates a step subsequent to RNA binding that is dependent upon the strength or structure of the *rut* site in ways not yet understood (*e.g.*, ring closure, dissociation, or even translocation). Biochemical analyses that can monitor individual steps throughout the process of Rho termination on several substrate RNAs will be necessary to define NusG's exact role as a positive regulator of Rho activity.

1.7 Physiological roles for Rho

Compared to intrinsic termination, Rho plays a wider variety of roles *in vivo*, acting as a sort of “cellular housekeeper” of transcription. In *E. coli*, a third of tRNA operons, several sRNAs, and no more than 20% of mRNA-encoded transcription units are terminated by Rho [3]. The number of genes terminated by Rho in the Gram-positive bacterium *B. subtilis*, where Rho is dispensable, is believed to be substantially lower [143]. Stalled ECs, which can form deleterious roadblocks to the replication fork during DNA replication, have also been implicated as a target for Rho-dependent termination [144].

1.7.1 Rho acts to prevent the accumulation of R-loops

A primary function of Rho may be to suppress and remove toxic R-loops that often form upstream of active ECs [5]. These three-stranded structures form when the nascent RNA transcript binds the transiently exposed template strand upstream of ECs to form an RNA-DNA duplex, displacing the non-template strand. R-loops are not thought to be toxic on their own but can be processed into deleterious double-stranded breaks, giving rise to genomic instability [145,

146]. Although it has long been appreciated both that *E. coli* Rho is necessary for cellular viability and that Rho plays a crucial role in suppressing R-loop formation, Gowrishankar and coworkers recently demonstrated that ectopic expression of a phage RNA-DNA helicase can rescue $\Delta\rho$ lethality [5]. This discovery implicates R-loop suppression as the essential Rho activity and opens new doors for *in vivo* research of Rho function.

1.7.2 Nucleoid-associated proteins promote termination of foreign DNA and antisense transcripts

Rho plays an important role in the silencing of foreign transcription units (i.e., cryptic prophages and other horizontally transferred genes) [3, 4]. For example, toxic genes within the *rac* prophage are efficiently silenced via upstream termination by Rho at t_{imm} [4]. Suboptimal codon usage within exogenous DNA may prevent efficient coupling of the EC to the ribosome, exposing *rut* sites for rapid binding by Rho [4]. At other sites within the genome, insertion of foreign DNA has been observed to disrupt intrinsic terminators, requiring the compensatory action of Rho [3].

Antisense transcription has also been identified as a target for silencing by Rho [2, 3]. Tiling microarrays and RNA sequencing analyses were used to map sites of Rho termination across the *E. coli* genome. The results suggested that a major role for Rho was the suppression of both intragenic transcription from antisense promoters within genes and intergenic transcription past the end of one gene into an oppositely-oriented gene immediately downstream. Of the >1,250 Rho terminators identified via these methods, 88% were found to act against antisense transcription units [2].

It has long been understood that the histone-like nucleoid structural protein (H-NS) and its closely related paralogs coat foreign genes [147] as well as sites of antisense and noncoding promoters [2, 148] via nucleation at AT-rich DNA sequences [149], a hallmark of foreign DNA in *E. coli*. In this way, H-NS blocks transcription initiation via physical occlusion of some promoters and trapping of open complexes near others [148]. Genetic synergy between H-NS-family proteins and Rho, however, was recently identified and implicates H-NS in aiding Rho-dependent termination in addition to its effects on initiation [150, 151]. ChIP experiments identified a strong correlation between Rho and regions of the bacterial nucleoid coated with H-NS [2]. Of the Rho-dependent terminators shown to suppress antisense transcripts, >80% were correlated with H-NS, with the highest H-NS occupancy near the actual site of termination. This result suggests that Rho activity is pervasive within H-NS filaments. Intrinsic terminators, on the other hand, were not found to be correlated with H-NS binding to the nucleoid [2].

H-NS forms oligomeric filaments by establishing DNA:H-NS interactions with its DNA-binding domain and H-NS–H-NS interactions with its oligomerization domain [152]. Biochemical work has established that two distinct forms of filaments can be formed *in vitro* [153]. At high concentrations of H-NS, filaments adopt a linear conformation, in which one DNA molecule is coated. At lower concentrations that may closer represent physiological H-NS conditions, a bridged filament is formed in which multiple DNA molecules, or two distal segments of one molecule, are brought together, possibly through flipping out of alternating DNA-binding domains [153]. Surprisingly, ECs were found to be capable of elongating through a linear filament at nearly the same rate as on bare DNA (Fig. 1.5A), but backtracked pausing was preferentially enhanced under conditions that favor bridged-filament formation. Addition of Rho induced significant levels of early termination at many of these bridged H-NS filament-

dependent pause sites (Fig. 1.5B) [153]. Despite backtracked pauses being atypical Rho termination sites, the long-lived nature of H-NS-dependent pauses may act to extend the kinetic window long enough for Rho to act. It will be important to determine to what extent the synergy between H-NS and Rho, now demonstrated *in vitro*, acts to silence transcription of foreign DNA and antisense transcripts *in vivo*.

1.8 Conclusions

The sequence determinants of intrinsic and Rho-dependent termination and the basic steps in the mechanisms are now relatively well understood, including some insight into the rearrangements of nucleic acids and their energetic contributions to termination. However, many aspects of these termination mechanisms remain to be characterized, including the structural rearrangements of RNAP involved in each step of the two termination pathways and the energetic contributions of these rearrangements. In this review, we have highlighted key new insights that have advanced understanding of the termination mechanisms, including evidence for additional layers of regulation *in vivo*.

Multiple lines of evidence indicate that the T_{hp} forms in a two-step mechanism due to an energetic barrier to hairpin completion likely created through RNAP-NA interactions such as between the hybrid and lid [10, 68, 74]. Whether these interactions are responsible for the energetic barrier to T_{hp} completion requires testing. Completion of the T_{hp} causes melting of the upstream RNA-DNA hybrid [25, 26], accompanied by unknown RNAP rearrangements. Together these changes result in inactivation of the EC, which commits the EC to the termination pathway [25, 74, 81].

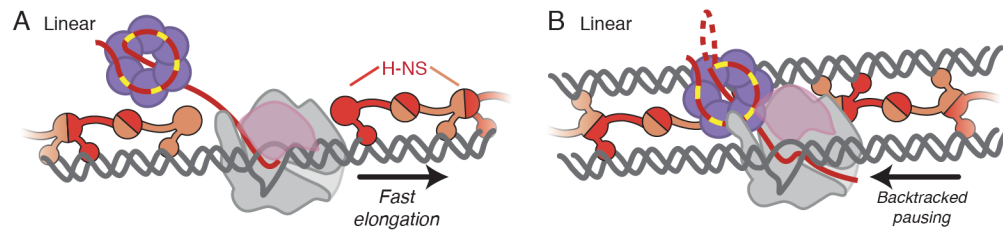


Figure 1.5. Synergy between Rho and bridged H-NS filaments *in vitro*. **(A)** At high concentrations, H-NS oligomerizes into linear filaments on DNA through which ECs can rapidly elongate. **(B)** At lower concentrations, H-NS can form bridged filaments on DNA by bringing two distinct molecules or two distal sites on one molecule together. Topological constraints introduced by DNA bridging enhances pausing by the EC, extending the kinetic window during which Rho can act to induce termination.

In vivo characterization of terminator sequences indicates that A-tracts upstream of the terminator may pair with the U-tract to extend the T_{hp} and increase termination efficiency [69]. An effect of A–U-tract pairing on termination efficiency is in apparent disagreement with the idea of commitment to termination upon T_{hp} completion. However, dissociable transcription factors may bind and alter inactivated ECs, accelerating isomerization back into an active EC and allowing A:U-tract pairing to affect termination efficiency. Significant efforts are being made to characterize the activities of various transcription factors on the termination mechanisms, necessary for gaining a full picture of the regulation of termination *in vivo*.

The discovery of synergy between H-NS and Rho activity *in vitro* similarly suggests another layer of Rho-regulation may exist *in vivo* [2, 150, 153]. Elucidating the mechanistic details of this synergy and its effects on gene expression will help define its physiological significance. A recent *in vivo* study also demonstrated that one of the primary roles of Rho-dependent termination may be to suppress R-loop formation, and prevent consequent DNA damage [5].

Another possible explanation of differences in termination mechanisms defined *in vitro* and *in vivo* may be differences in the solutes present and the effects of solutes on protein-nucleic acid interactions. Most *in vitro* studies of termination mechanisms have used solutions with significant concentrations of Cl^- , yet Cl^- can compete for binding of DNA and RNA to proteins [154] and is not representative of the cellular environment [155]. Use of acetate or glutamate in place of Cl^- creates a better mimic of the ionic conditions in the bacterial cytoplasm [75, 78, 154, 156] and should be adopted in future studies of termination mechanisms where possible.

Substantial strides have been made in elucidating the steps of intrinsic and Rho-dependent termination, particularly the role of nucleic-acid rearrangements *in vitro*. However,

many questions remain, most notably about how RNAP structural rearrangements participate in these termination pathways and about differences between mechanisms operating *in vitro* and *in vivo*. Answering these questions will likely require fully exploiting new structural methods like cryo-EM and high-throughput sequencing-enabled *in vivo* methods like ChIP-seq and NET-seq.

1.9 Organization of dissertation

The research presented in this dissertation exploited biochemical approaches to determine how protein conformation changes influence the two programmed mechanisms of transcription termination in *Escherichia coli*. I defined how clamp movements influence the commitment and dissociation steps in the intrinsic termination mechanism, explored the role of the TL in promoting dissociation, and helped define the mechanism of NusG enhancement of Rho-dependent termination.

In Chapter 2 I explored the role of clamp conformation during intrinsic termination across three model terminators. To do so, I developed a cysteine-pair ligation scaffold that bypasses the initiation mechanism and enables the study of transcription elongation and termination. Using cysteine-pair RNAPs that biased the clamp into particular conformations (*e.g.*, closed or displaced) I determined that while clamp conformation has dramatic effects on termination efficiency, this is through adjustments made to the EC elongation rate, not its commitment rate. I was also able to explore how these same clamp conformations affected the second step of the intrinsic termination mechanism, RNA and DNA dissociation from RNAP. Contrary to its role (or lack thereof) in commitment, inducing a topological constraint on the clamp that prevents displacement slows the rate of DNA dissociation from the main cleft; however, no effect on the apparent rate of RNA release was observed. Finally, I provided further

evidence of the MS-MP thermodynamic model of termination commitment by monitoring the slowed rates of RNA and DNA release in the presence of Microcin J25, an antimicrobial peptide that biases the TL towards an unfolded state and thus reduces its conformational flexibility.

I next describe the fruits of a collaborative project that I began with Michael R. Lawson at a Mechanism and Regulation of Prokaryotic Transcription Science Research Conference. The mechanism by which NusG promotes early Rho-dependent termination has long been a matter of debate and confusion (see Chapter 1.6). Based upon the most recent structural and biochemical data, we proposed that NusG acted upon the open-Rho ring structure to promote ring closure at subpar RNA substrates. M.R.L. performed solved a crystal structure of Rho bound by the NusG-CTD and I performed *in vitro* Rho termination assays to identify an amino acid switch network within the Rho ring that mediates Rho conformation, thus confirming that NusG can promote Rho ring closure at weak *rut* sites.

Finally, in Chapter 4, I briefly summarize the findings of this dissertation and suggest future directions for the upcoming members of the field of bacterial transcription termination.

1.10 References

1. Selby, C.P. and A. Sancar, *Molecular mechanism of transcription-repair coupling*. Science, 1993. **260**(5104): p. 53-8.
2. Peters, J.M., et al., *Rho and NusG suppress pervasive antisense transcription in Escherichia coli*. Genes Dev, 2012. **26**(23): p. 2621-33.
3. Peters, J.M., et al., *Rho directs widespread termination of intragenic and stable RNA transcription*. Proc Natl Acad Sci U S A, 2009. **106**(36): p. 15406-11.
4. Cardinale, C.J., et al., *Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E. coli*. Science, 2008. **320**(5878): p. 935-8.
5. Leela, K.J., et al., *Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in Escherichia coli*. Proc Natl Acad Sci U S A, 2013. **110**(1): p. 258-263.

6. Park, J.S. and J.W. Roberts, *Role of DNA bubble rewinding in enzymatic transcription termination*. Proc Natl Acad Sci U S A, 2006. **103**(13): p. 4870-5.
7. Selby, C.P. and A. Sancar, *Structure and function of Transcription-Repair Coupling Factor*. J Biol Chem, 1995. **270**(9).
8. Komissarova, N. and M. Kashlev, *Functional topography of nascent RNA in elongation intermediates of RNA polymerase*. Proc Natl Acad Sci U S A, 1998. **95**: p. 14699-704.
9. Korzheva, N., et al., *A structural model of transcription elongation*. Science, 2000. **289**(5479): p. 619-25.
10. Vassylyev, D.G., et al., *Structural basis for transcription elongation by bacterial RNA polymerase*. Nature, 2007. **448**(7150): p. 157-62.
11. Darst, S.A., et al., *Conformational flexibility of bacterial RNA polymerase*. Proc. Natl. Acad. Sci. U. S. A., 2002. **99**(7): p. 4296-301.
12. Vassylyev, D.G., et al., *Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution*. Nature, 2002. **417**(6890): p. 712-9.
13. Chakraborty, A., et al., *Opening and closing of the bacterial RNA polymerase clamp*. Science, 2012. **337**(6094): p. 591-5.
14. Tagami, S., et al., *Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein*. Nature, 2010. **468**(7326): p. 978-82.
15. Kang, J.Y., et al., *Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex*. Elife, 2017. **6**.
16. Hein, P.P., et al., *RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement*. Nat Struct Mol Biol, 2014. **21**(9): p. 794-802.
17. Sekine, S., et al., *The ratcheted and ratchetable structural states of RNA polymerase underlie multiple transcriptional functions*. Mol Cell, 2015. **57**(3): p. 408-21.
18. Gnatt, A.L., et al., *Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution*. Science, 2001. **292**(5523): p. 1876-82.
19. Ederth, J., et al., *The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing*. J Biol Chem, 2002. **277**(40): p. 37456-63.
20. Cramer, P., et al., *Architecture of RNA polymerase II and implications for the transcription mechanism*. Science, 2000. **288**(5466): p. 640-9.
21. Weixlbaumer, A., et al., *Structural basis of transcriptional pausing in bacteria*. Cell, 2013. **152**(3): p. 431-41.

22. Kang, J.Y., et al., *RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing*. Mol Cell, 2018. **69**(5): p. 802-815 e1.
23. Guo, X., et al., *Structural Basis for NusA Stabilized Transcriptional Pausing*. Mol Cell, 2018. **69**(5): p. 816-827 e4.
24. Wilson, K.S. and P.H. von Hippel, *Stability of Escherichia coli transcription complexes near an intrinsic terminator*. J Mol Biol, 1994. **244**(1): p. 36-51.
25. Gusarov, I. and E. Nudler, *The mechanism of intrinsic transcription termination*. Mol Cell, 1999. **3**(4): p. 495-504.
26. Komissarova, N., et al., *Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination*. Mol Cell, 2002. **10**(5): p. 1151-62.
27. Lee, D.N., et al., *Transcription pausing by Escherichia coli RNA polymerase is modulated by downstream DNA sequences*. J. Biol. Chem., 1990. **265**: p. 15145-15153.
28. Touloukhonov, I., et al., *A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing*. Mol Cell, 2007. **27**(3): p. 406-19.
29. Yarnell, W.S. and J.W. Roberts, *Mechanism of intrinsic transcription termination and antitermination*. Science, 1999. **284**(5414): p. 611-5.
30. Neuman, K.C., et al., *Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking*. Cell, 2003. **115**(4): p. 437-47.
31. Larson, M.H., et al., *A pause sequence enriched at translation start sites drives transcription dynamics in vivo*. Science, 2014. **344**(6187): p. 1042-7.
32. Vvedenskaya, I.O., et al., *Interactions between RNA polymerase and the "core recognition element" counteract pausing*. Science, 2014. **344**(6189): p. 1285-9.
33. Touloukhonov, I. and R. Landick, *The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination*. Mol Cell, 2003. **12**(5): p. 1125-36.
34. Wang, D., K. Severinov, and R. Landick, *Preferential interaction of the his pause RNA hairpin with RNA polymerase beta subunit residues 904-950 correlates with strong transcriptional pausing*. Proc Natl Acad Sci U S A, 1997. **94**(16): p. 8433-8.
35. Chan, C.L. and R. Landick, *Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin*. J. Mol. Biol., 1993. **233**: p. 25-42.

36. Wilson, K.S. and P.H. von Hippel, *Transcription termination at intrinsic terminators: the role of the RNA hairpin*. Proc Natl Acad Sci U S A, 1995. **92**(19): p. 8793-7.
37. Touloukhonov, I., I. Artsimovitch, and R. Landick, *Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins*. Science, 2001. **292**(5517): p. 730-3.
38. Yang, X., et al., *The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA*. EMBO Rep, 2009. **10**(9): p. 997-1002.
39. Ha, K.S., et al., *The NusA N-terminal domain is necessary and sufficient for enhancement of transcriptional pausing via interaction with the RNA exit channel of RNA polymerase*. J Mol Biol, 2010.
40. Kolb, K.E., P.P. Hein, and R. Landick, *Antisense oligonucleotide-stimulated transcriptional pausing reveals RNA exit channel specificity of RNA polymerase and mechanistic contributions of NusA and RfaH*. J Biol Chem, 2014. **289**(2): p. 1151-63.
41. Nudler, E., et al., *The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase*. Cell, 1997. **89**(1): p. 33-41.
42. Guajardo, R. and R. Sousa, *A model for the mechanism of polymerase translocation*. J Mol Biol, 1997. **265**(1): p. 8-19.
43. Komissarova, N. and M. Kashlev, *RNA polymerase switches between inactivated and activated states By translocating back and forth along the DNA and the RNA*. J Biol Chem, 1997. **272**(24): p. 15329-38.
44. Kassavetis, G.A. and M.J. Chamberlin, *Pausing and termination of transcription within the early region of bacteriophage T7 DNA in vitro*. J Biol Chem, 1981. **256**(6): p. 2777-86.
45. Morgan, W.D., D.G. Bear, and P.H. von Hippel, *Rho-dependent termination of transcription. II. Kinetics of mRNA elongation during transcription from the bacteriophage lambda PR promoter*. J Biol Chem, 1983. **258**(15): p. 9565-74.
46. Lau, L.F. and J.W. Roberts, *Rho-dependent transcription termination at lambda R1 requires upstream sequences*. J Biol Chem, 1985. **260**(1): p. 574-84.
47. Galloway, J.L. and T. Platt, *Signals sufficient for rho-dependent transcription termination at trp t' span a region centered 60 base pairs upstream of the earliest 3' end point*. J Biol Chem, 1988. **263**(4): p. 1761-7.
48. Ray-Soni, A., R.A. Mooney, and R. Landick, *Trigger loop dynamics can explain stimulation of intrinsic termination by bacterial RNA polymerase without terminator hairpin contact*. Proc Natl Acad Sci U S A, 2017. **114**(44): p. E9233-E9242.

49. Komissarova, N. and M. Kashlev, *Transcriptional arrest: Escherichia coli RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded*. Proc Natl Acad Sci U S A, 1997. **94**(5): p. 1755-60.
50. Jin, D.J., et al., *Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho*. Proc Natl Acad Sci U S A, 1992. **89**(4): p. 1453-7.
51. Dutta, D., J. Chalisery, and R. Sen, *Transcription termination factor rho prefers catalytically active elongation complexes for releasing RNA*. J Biol Chem, 2008. **283**(29): p. 20243-51.
52. Herbert, K.M., et al., *E. coli NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase*. J Mol Biol, 2010. **399**(1): p. 17-30.
53. Kang, J.Y., et al., *Structural Basis for Transcript Elongation Control by NusG Family Universal Regulators*. Cell, 2018. **173**(7): p. 1650-1662.e14.
54. Yakhnin, A.V. and P. Babitzke, *Mechanism of NusG-stimulated pausing, hairpin-dependent pause site selection and intrinsic termination at overlapping pause and termination sites in the Bacillus subtilis trp leader*. Mol Microbiol, 2010. **76**(3): p. 690-705.
55. Yakhnin, A.V., H. Yakhnin, and P. Babitzke, *Function of the Bacillus subtilis transcription elongation factor NusG in hairpin-dependent RNA polymerase pausing in the trp leader*. Proc Natl Acad Sci U S A, 2008. **105**(42): p. 16131-6.
56. Czyz, A., et al., *Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators*. MBio, 2014. **5**(2): p. e00931.
57. Sullivan, S.L. and M.E. Gottesman, *Requirement for E. Coli NusG protein in factor-dependent transcription termination*. Cell, 1992. **68**(5): p. 989-994.
58. Nehrke, K.W., F. Zalatan, and T. Platt, *NusG alters Rho-dependent termination of transcription in vitro independent of kinetic coupling*. Gene Expr, 1993. **3**(2): p. 119-33.
59. Burmann, B.M., et al., *A NusE:NusG complex links transcription and translation*. Science, 2010. **328**(5977): p. 501-4.
60. Sevostyanova, A., et al., *The beta subunit gate loop is required for RNA polymerase modification by RfaH and NusG*. Mol Cell, 2011. **43**(2): p. 253-62.
61. Svetlov, V., et al., *Allosteric control of the RNA polymerase by the elongation factor RfaH*. Nucleic Acids Res, 2007. **35**(17): p. 5694-705.

62. Belogurov, G.A., et al., *Structural basis for converting a general transcription factor into an operon-specific virulence regulator*. Mol Cell, 2007. **26**(1): p. 117-29.
63. Santangelo, T.J. and I. Artsimovitch, *Termination and antitermination: RNA polymerase runs a stop sign*. Nat Rev Microbiol, 2011. **9**(5): p. 319-29.
64. d'Aubenton Carafa, Y., E. Brody, and C. Thermes, *Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures*. J Mol Biol, 1990. **216**(4): p. 835-58.
65. McDowell, J.C., et al., *Determination of intrinsic transcription termination efficiency by RNA polymerase elongation rate*. Science, 1994. **266**(5186): p. 822-5.
66. Kireeva, M.L. and M. Kashlev, *Mechanism of sequence-specific pausing of bacterial RNA polymerase*. Proc Natl Acad Sci U S A, 2009. **106**(22): p. 8900-5.
67. Chan, C.L., D. Wang, and R. Landick, *Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways*. J Mol Biol, 1997. **268**(1): p. 54-68.
68. Lubkowska, L., A.S. Maharjan, and N. Komissarova, *RNA folding in transcription elongation complex: implication for transcription termination*. J Biol Chem, 2011. **286**(36): p. 31576-85.
69. Chen, Y.J., et al., *Characterization of 582 natural and synthetic terminators and quantification of their design constraints*. Nat Methods, 2013. **10**(7): p. 659-64.
70. Penno, C., et al., *Productive mRNA stem loop-mediated transcriptional slippage: Crucial features in common with intrinsic terminators*. Proc Natl Acad Sci U S A, 2015. **112**(16): p. E1984-93.
71. Chan, C., D. Wang, and R. Landick, *Spacing from the transcript 3' end determines whether a nascent RNA hairpin interacts with RNA polymerase to prolong pausing or triggers termination*. J. Mol. Biol., 1997. **268**: p. 54-68.
72. Goliger, J.A., et al., *Early transcribed sequences affect termination efficiency of Escherichia coli RNA polymerase*. J Mol Biol, 1989. **205**(2): p. 331-41.
73. Telesnitsky, A.P. and M.J. Chamberlin, *Sequences linked to prokaryotic promoters can affect the efficiency of downstream termination sites*. J Mol Biol, 1989. **205**(2): p. 315-30.
74. Larson, M.H., et al., *Applied force reveals mechanistic and energetic details of transcription termination*. Cell, 2008. **132**(6): p. 971-82.
75. Reynolds, R., R.M. Bermudez-Cruz, and M.J. Chamberlin, *Parameters affecting transcription termination by Escherichia coli RNA polymerase. I. Analysis of 13 rho-independent terminators*. J Mol Biol, 1992. **224**(1): p. 31-51.

76. Santangelo, T.J. and J.W. Roberts, *Forward translocation is the natural pathway of RNA release at an intrinsic terminator*. Molecular Cell, 2004. **14**(1): p. 117-126.
77. Peters, J.M., A.D. Vangeloff, and R. Landick, *Bacterial transcription terminators: the RNA 3'-end chronicles*. J Mol Biol, 2011. **412**(5): p. 793-813.
78. Reynolds, R. and M.J. Chamberlin, *Parameters affecting transcription termination by Escherichia coli RNA. II. Construction and analysis of hybrid terminators*. J Mol Biol, 1992. **224**(1): p. 53-63.
79. von Hippel, P.H., *An integrated model of the transcription complex in elongation, termination, and editing*. Science, 1998. **281**(5377): p. 660-5.
80. von Hippel, P.H. and T.D. Yager, *The elongation-termination decision in transcription*. Science, 1992. **255**(5046): p. 809-12.
81. Yin, H., et al., *Nonequilibrium mechanism of transcription termination from observations of single RNA polymerase molecules*. Proc Natl Acad Sci U S A, 1999. **96**(23): p. 13124-9.
82. Martin, F.H. and I. Tinoco, Jr., *DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription*. Nucleic Acids Res, 1980. **8**(10): p. 2295-9.
83. Weilbaecher, R., et al., *Termination-altering amino acid substitutions in the beta' subunit of Escherichia coli RNA polymerase identify regions involved in RNA chain elongation*. Genes Dev, 1994. **8**(23): p. 2913-27.
84. Cheeran, A., et al., *Escherichia coli RNA polymerase mutations located near the upstream edge of an RNA:DNA hybrid and the beginning of the RNA-exit channel are defective for transcription antitermination by the N protein from lambdaoid phage H-19B*. J Mol Biol, 2005. **352**(1): p. 28-43.
85. Epshtein, V., et al., *An allosteric path to transcription termination*. Mol Cell, 2007. **28**(6): p. 991-1001.
86. Yager, T.D. and P.H. von Hippel, *A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli*. Biochemistry, 1991. **30**: p. 1097-1118.
87. Fischer, N., et al., *Structure of the E. coli ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM*. Nature, 2015. **520**(7548): p. 567-70.
88. Lowery-Goldhammer, C. and J.P. Richardson, *An RNA-dependent nucleoside triphosphate phosphohydrolase (ATPase) associated with rho termination factor*. Proc Natl Acad Sci U S A, 1974. **71**(5): p. 2003-7.
89. Rutherford, S.T., et al., *Allosteric control of Escherichia coli rRNA promoter complexes by DksA*. Genes Dev, 2009. **23**(2): p. 236-48.

90. Chen, C.Y. and J.P. Richardson, *Sequence elements essential for rho-dependent transcription termination at lambda tR1*. J Biol Chem, 1987. **262**(23): p. 11292-9.
91. Zalatan, F. and T. Platt, *Effects of decreased cytosine content on rho interaction with the rho-dependent terminator trp t' in Escherichia coli*. J Biol Chem, 1992. **267**(27): p. 19082-8.
92. McSwiggen, J.A., D.G. Bear, and P.H. von Hippel, *Interactions of Escherichia coli transcription termination factor rho with RNA. I. Binding stoichiometries and free energies*. J Mol Biol, 1988. **199**(4): p. 609-22.
93. Koslover, D.J., et al., *Binding and translocation of termination factor rho studied at the single-molecule level*. J Mol Biol, 2012. **423**(5): p. 664-76.
94. Alifano, P., et al., *A consensus motif common to all Rho-dependent prokaryotic transcription terminators*. Cell, 1991. **64**(3): p. 553-63.
95. Chen, C.Y., G.R. Galluppi, and J.P. Richardson, *Transcription termination at lambda tR1 is mediated by interaction of rho with specific single-stranded domains near the 3' end of cro mRNA*. Cell, 1986. **46**(7): p. 1023-8.
96. Hollands, K., et al., *Riboswitch control of Rho-dependent transcription termination*. Proc Natl Acad Sci U S A, 2012. **109**(14): p. 5376-5381.
97. Hollands, K., A. Sevostiyanova, and E.A. Groisman, *Unusually long-lived pause required for regulation of a Rho-dependent transcription terminator*. Proc Natl Acad Sci U S A, 2014. **111**(19): p. E1999-2007.
98. Schwartz, A., et al., *Noncanonical interactions in the management of RNA structural blocks by the transcription termination rho helicase*. Biochemistry, 2007. **46**(33): p. 9366-79.
99. Hart, C.M. and J.W. Roberts, *Rho-dependent transcription termination. Characterization of the requirement for cytidine in the nascent transcript*. J Biol Chem, 1991. **266**(35): p. 24140-8.
100. Skordalakes, E. and J.M. Berger, *Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading*. Cell, 2003. **114**(1): p. 135-46.
101. Gogol, E.P., S.E. Seifried, and P.H. von Hippel, *Structure and assembly of the Escherichia coli transcription termination factor rho and its interaction with RNA. I. Cryoelectron microscopic studies*. J Mol Biol, 1991. **221**(4): p. 1127-38.
102. Thomsen, N.D. and J.M. Berger, *Running in reverse: the structural basis for translocation polarity in hexameric helicases*. Cell, 2009. **139**(3): p. 523-34.

103. Lawson, M.R., K. Dyer, and J.M. Berger, *Ligand-induced and small-molecule control of substrate loading in a hexameric helicase*. Proc Natl Acad Sci U S A, 2016. **113**(48): p. 13714-13719.
104. Briercheck, D.M., et al., *The NMR structure of the RNA binding domain of E. coli rho factor suggests possible RNA-protein interactions*. Nat Struct Biol, 1998. **5**(5): p. 393-9.
105. Ceruzzi, M.A.F., S.L. Bektesh, and J.P. Richardson, *Interaction of Rho factor with bacteriophage λ cro gene transcripts*. J Biol Chem, 1985. **260**(16): p. 9412-9418.
106. Pereira, S. and T. Platt, *A mutation in the ATP binding domain of Rho alters its RNA binding properties and uncouples ATP hydrolysis from helicase activity*. J Biol Chem, 1995. **270**(51): p. 30401-30407.
107. Gan, E. and J.P. Richardson, *Identification of a structural element that is essential for two functions of transcription factor NusG*. Biochemistry, 1999. **38**(51): p. 16882-16888.
108. Walmacq, C., A.R. Rahmouni, and M. Boudvillain, *Influence of substrate composition on the helicase activity of transcription termination factor Rho: reduced processivity of Rho hexamers during unwinding of RNA-DNA hybrid regions*. J Mol Biol, 2004. **342**(2): p. 403-20.
109. Faus, I. and J.P. Richardson, *Thermodynamic and enzymological characterization of the interaction between transcription termination factor Rho and λ cro mRNA*. Biochemistry, 1989. **28**(8): p. 3510-3517.
110. Steinmetz, E.J. and T. Platt, *Evidence supporting a tethered tracking model for helicase activity of Escherichia coli Rho factor*. Proc Natl Acad Sci U S A, 1994. **91**(4): p. 1401-5.
111. D'Heygere, F., M. Rabhi, and M. Boudvillain, *Phyletic distribution and conservation of the bacterial transcription termination factor Rho*. Microbiology, 2013. **159**(Pt 7): p. 1423-36.
112. Canals, A., I. Uson, and M. Coll, *The structure of RNA-free Rho termination factor indicates a dynamic mechanism of transcript capture*. J Mol Biol, 2010. **400**(1): p. 16-23.
113. D'Heygere, F., et al., *ATP-dependent motor activity of the transcription termination factor Rho from Mycobacterium tuberculosis*. Nucleic Acids Res, 2015. **43**(12): p. 6099-111.
114. Miwa, Y., T. Horiguchi, and K. Shigesada, *Structural and functional dissections of transcription termination factor-Rho by random mutagenesis*. J Mol Biol, 1995. **254**(5): p. 815-837.
115. Story, R., I. Weber, and T. Steitz, *The structure of the E. coli recA protein monomer and polymer*. Nature, 1992. **355**(6358): p. 818-325.

116. Richardson, J.P., *Activation of Rho protein ATPase requires simultaneous interaction at two kinds of nucleic acid-binding sites*. J Biol Chem, 1982. **257**(10): p. 5760-5766.
117. Briercheck, D.M., et al., *¹H, ¹⁵N and ¹³C resonance assignments and secondary structure determination of the RNA-binding domain of E. coli Rho protein*. J Biomol NMR, 1996. **8**(4): p. 429-444.
118. Schwartz, A., et al., *A stepwise 2'-hydroxyl activation mechanism for the bacterial transcription termination factor Rho helicase*. Nat Struct Mol Biol, 2009. **16**(12): p. 1309-16.
119. Adelman, J.L., et al., *Mechanochemistry of transcription termination factor Rho*. Mol Cell, 2006. **22**(5): p. 611-21.
120. Mooney, R.A., et al., *Regulator trafficking on bacterial transcription units in vivo*. Mol Cell, 2009. **33**(1): p. 97-108.
121. Epshtein, V., et al., *An allosteric mechanism of Rho-dependent transcription termination*. Nature, 2010. **463**(7278): p. 245-9.
122. Browne, R.J., E.W. Barr, and B.L. Stitt, *Catalytic cooperativity among subunits of Escherichia coli transcription termination factor Rho. Kinetics and substrate structural requirements*. J Biol Chem, 2005. **280**(14): p. 13292-9.
123. Soares, E., et al., *The RNA-mediated, asymmetric ring regulatory mechanism of the transcription termination Rho helicase decrypted by time-resolved nucleotide analog interference probing (trNAIP)*. Nucleic Acids Res, 2014. **42**(14): p. 9270-84.
124. Schwartz, A., et al., *Transcription termination factor rho can displace streptavidin from biotinylated RNA*. J Biol Chem, 2007. **282**(43): p. 31469-76.
125. Steinmetz, E.J., C.A. Brennan, and T. Platt, *A short intervening structure can block rho factor helicase action at a distance*. J Biol Chem, 1990. **265**(30): p. 18408-13.
126. Galluppi, G.R. and J.P. Richardson, *ATP-induced changes in the binding of RNA synthesis termination protein Rho to RNA*. J Mol Biol, 1980. **138**(3): p. 513-39.
127. Lau, L.F., J.W. Roberts, and R. Wu, *Transcription terminates at lambda tR1 in three clusters*. Proc Natl Acad Sci U S A, 1982. **79**(20): p. 6171-5.
128. Brennan, C.A., A.J. Dombroski, and T. Platt, *Transcription termination factor rho is an RNA-DNA helicase*. Cell, 1987. **48**(6): p. 945-52.
129. Pasman, Z. and P.H. von Hippel, *Regulation of Rho-dependent transcription termination by NusG is specific to the Escherichia coli elongation complex*. Biochemistry, 2000. **39**(18): p. 5573-85.

130. Lang, H., T. Platt, and R. Reeder, *Escherichia coli* Rho factor induces release of yeast RNA polymerase II but not polymerase I or III. *Proc Natl Acad Sci U S A*, 1998. **95**(9): p. 4900-4905.
131. Richardson, J.P., *Rho-dependent termination and ATPases in transcript termination*. *Biochim Biophys Acta*, 2002. **1577**(2): p. 251-260.
132. Downing, W.L., et al., *Sequence and transcriptional pattern of the essential Escherichia coli secE-nusG operon*. *J Bacteriol*, 1990. **172**(3): p. 1621-1627.
133. Martinez-Rucobo, F.W., et al., *Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity*. *EMBO J*, 2011. **30**(7): p. 1302-10.
134. Chen, H., et al., *Repression of RNA polymerase II elongation in vivo is critically dependent on the C-terminus of Spt5*. *PLoS One*, 2009. **4**(9): p. e6918.
135. Burova, E., et al., *Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro*. *J Bacteriol*, 1995. **177**(5): p. 1388-1392.
136. Mooney, R.A., et al., *Two structurally independent domains of E. coli NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators*. *J Mol Biol*, 2009. **391**(2): p. 341-58.
137. Chalissery, J., et al., *Interaction Surface of the Transcription Terminator Rho Required to Form a Complex with the C-Terminal Domain of the Antiterminator NusG*. *J Mol Biol*, 2010.
138. Adhya, S. and M. Gottesman, *Control of transcription termination*. *Annu Rev Biochem*, 1978. **47**: p. 967-96.
139. Kohler, R., et al., *Architecture of a transcribing-translating expressome*. *Science*, 2017. **356**: p. 194-197.
140. Saxena, S., et al., *Escherichia coli transcription factor NusG binds to 70S ribosomes*. *Mol Microbiol*, 2018. **108**(5): p. 495-504.
141. Burns, C.M. and J.P. Richardson, *NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator*. *Proc Natl Acad Sci U S A*, 1995. **92**(11): p. 4738-42.
142. Kalyani, B.S., et al., *Interaction with the nascent RNA is a prerequisite for the recruitment of Rho to the transcription elongation complex in vitro*. *J Mol Biol*, 2011. **413**(3): p. 548-60.
143. de Hoon, M.J., et al., *Prediction of transcriptional terminators in Bacillus subtilis and related species*. *PLoS Comput Biol*, 2005. **1**(3): p. e25.

144. Washburn, R.S. and M.E. Gottesman, *Transcription termination maintains chromosome integrity*. Proc Natl Acad Sci U S A, 2010.
145. Gan, W., et al., *R-loop-mediated genomic instability is caused by impairment of replication fork progression*. Genes Dev, 2011. **25**(19): p. 2041-56.
146. Hamperl, S. and K.A. Cimprich, *The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability*. DNA Repair (Amst), 2014. **19**: p. 84-94.
147. Navarre, W.W., et al., *Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA*. Genes Dev, 2007. **21**(12): p. 1456-71.
148. Singh, S.S., et al., *Widespread suppression of intragenic transcription initiation by H-NS*. Genes Dev, 2014. **28**(3): p. 214-9.
149. Owen-Hughes, T.A., et al., *The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene-expression*. Cell, 1992. **71**(2): p. 255-265.
150. Saxena, S. and J. Gowrishankar, *Modulation of Rho-dependent transcription termination in Escherichia coli by the H-NS family of proteins*. J Bacteriol, 2011. **193**(15): p. 3832-41.
151. Tran, L., et al., *Single-gene deletion mutants of Escherichia coli with altered sensitivity to bicyclomycin, an inhibitor of transcription termination factor Rho*. J Bacteriol, 2011. **193**(9): p. 2229-35.
152. Arold, S.T., et al., *H-NS forms a superhelical protein scaffold for DNA condensation*. Proc Natl Acad Sci U S A, 2010. **107**(36): p. 15728-32.
153. Kotlajich, M.V., et al., *Bridged filaments of histone-like nucleoid structuring protein pause RNA polymerase and aid termination in bacteria*. Elife, 2015. **4**.
154. Leirimo, S., et al., *Replacement of potassium chloride by potassium glutamate dramatically enhances protein-DNA interactions in vitro*. Biochemistry, 1987. **26**(8): p. 2095-101.
155. Cayley, S., et al., *Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity. Implications for protein-DNA interactions in vivo*. J Mol Biol, 1991. **222**(2): p. 281-300.
156. Chan, C.L. and R. Landick, *Effects of neutral salts on RNA chain elongation and pausing by Escherichia coli RNA polymerase*. J Mol Biol, 1997. **268**(1): p. 37-53.

Chapter 2

RNA polymerase clamp movement aids dissociation from DNA but is not required for RNA release at intrinsic terminators

Adapted from:

Michael J. Bellecourt, Ananya Ray-Soni, Alex Harwig, Rachel Mooney, Robert Landick,
Adapted from RNA polymerase clamp movement aids dissociation from DNA but is not
required for RNA release at intrinsic terminators. Article in preparation.

Chapter contributions:

M.J.B. and R.L. wrote the text. M.J.B. performed all experiments. A.R., A.H., & R.A.M. helped
prepare materials for the experiments. M.J.B constructed the figures. M.J.B. and R.L. designed
the experiments.

2.1 Abstract

In bacteria, viability requires efficient disassembly of elongating transcription complexes (ECs) in a 2–3 nucleotide window after transcription of multiple kilobase pairs of DNA. Intrinsic terminator signals utilize pausing over weak RNA-DNA base pairs followed by formation of a strong, GC-rich stem-loop structure in the RNA exit channel of RNA polymerase (RNAP) to inactivate nucleotide addition and induce dissociation of RNA and RNAP from DNA. Although the relative movements of RNA and DNA during intrinsic termination are well-characterized, the effects of RNAP conformational changes are less understood. RNAP contains a clamp domain that closes around the nucleic-acid scaffold during transcription initiation and can be displaced by either swiveling or opening motions. Displacement is proposed to promote termination by releasing RNAP–nucleic acid contacts. We developed a cysteine-crosslinking assay to constrain clamp movements and study effects on the intrinsic termination. We found that biasing the clamp into different conformations perturbed termination efficiency, but that these perturbations were due primarily to changes in elongation rate, not the competing rate at which the EC commits to termination. After commitment, however, inhibiting clamp movements slowed release of DNA but not of RNA from the EC. We also show that restricting movements of the trigger loop (TL) with the RNAP inhibitor microcin J25 prior to commitment inhibits intrinsic termination, in agreement with a recently proposed multistate-multipath model of intrinsic termination. Together our results support views that intrinsic termination and DNA release occur separately and that RNAP might remain associated with DNA after termination.

2.2 Introduction

Termination of transcription by RNA polymerase (RNAP) is the essential process in all organisms that must occur at the end of every functional transcription unit (*i.e.*, a gene or polycistronic operon) to release the newly-synthesized RNA product and dissociate RNAP from the DNA. Failure to terminate efficiently has deleterious effects on the cell, including but not limited to unregulated expression of downstream genes [1] and collisions with replication complexes that can cause double-stranded breaks in genomic DNA [2]. Bacteria have evolved two mechanisms to promote accurate transcription termination at the ends of transcription units. Rho-dependent termination relies upon the RNA helicase Rho to disassemble the elongation complex (EC), whereas intrinsic termination induces EC disassembly via a signal encoded by the DNA and RNA comprised of a palindromic GC-rich dyad immediately followed by an ~8-nt U-tract [3].

In thermodynamic models for intrinsic termination, termination efficiency (TE) is determined by the relative free energy barriers to termination vs. elongation at each template position [4, 5]. At most positions, the EC is strongly biased towards the elongation due to rapid nucleotide addition (up to 100 s⁻¹) and the high stability of the EC resulting from combined van der Waals and polar interactions between RNAP and the nucleic-acid (NA) scaffold. The NA scaffold is comprised of a 9–10-bp RNA-DNA hybrid within a ~12-bp transcription bubble, ~5 nt of single-stranded RNA threading out of the RNA exit channel, and ~18 bp of double-stranded downstream DNA [6-9].

Intrinsic terminators facilitate entry into the termination pathway through the U-tract, which promotes elemental pausing of the EC [10-13], and its surrounding sequences (Fig. 2.1A). The U-tract pause slows elongation and provides time for the GC-dyad to begin folding into a

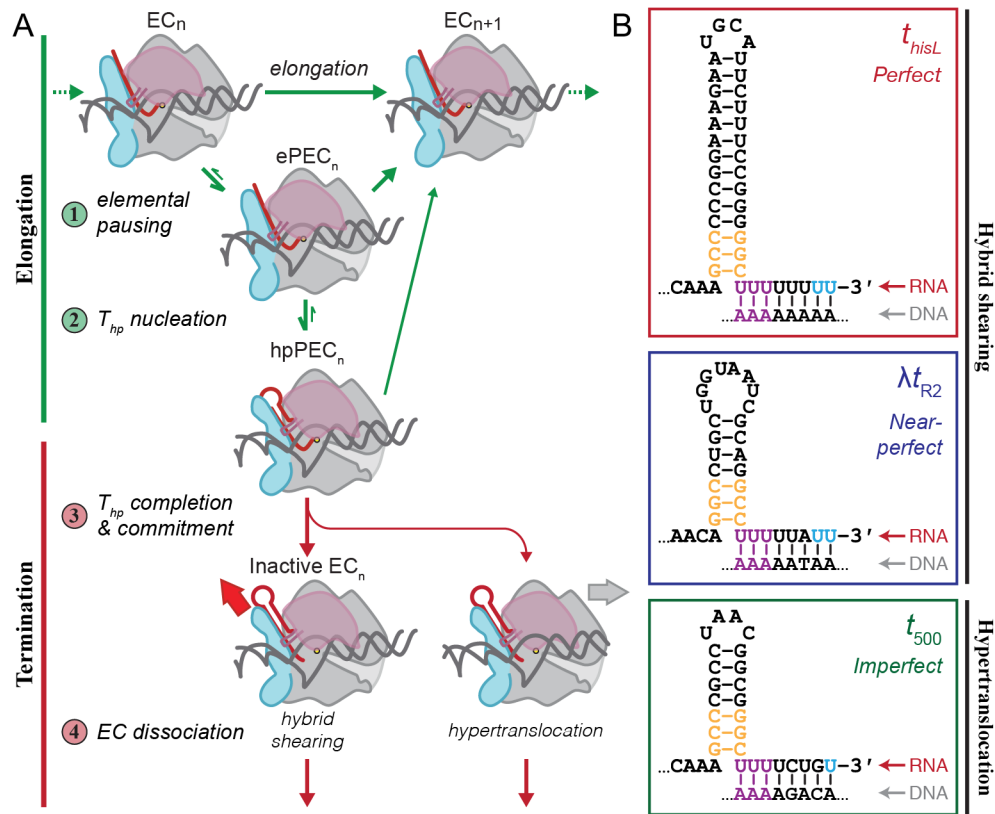


Figure 2.1. The mechanism of intrinsic termination. **(A)** The four major steps of the intrinsic termination pathway include steps that impact the elongation mechanism (green arrows) and those that impact the termination mechanism (red arrows). Each step along the mechanism is modulated by complex interactions between the nucleic-acid scaffold (DNA, grey; RNA, red) and RNAP itself. **(B)** Three model intrinsic terminators used in these studies. The two major components of an intrinsic terminator include a GC-rich dyad that folds into an RNA hairpin and a U-rich tract immediately downstream. The strength of the U-tract, as measured by its uridine-content, determines its class of terminator. Violet uridine bases indicate the most highly-conserved uridines that melt upon pairing of the final 2–3 G–C bases at the stem of the T_{hp} (orange). Blue uridine bases indicate the sites of termination.

terminator hairpin (T_{hp}), possibly inducing a transient hairpin-stabilized pause [14]. Unlike conventional hairpin pause signals, the juxtaposition of the weak rU–dA hybrid [15] and additional strong rG–rC base pairs at a terminator allow extension of the hairpin to within 8 nt of the RNA 3' end. This process, called T_{hp} completion, appears to be rate-limiting for termination because it creates steric clashes in the EC that can only be resolved by partial melting of the RNA-DNA hybrid. At terminators with near-perfect U-tracts like *t_{hisL}* and λt_{R2} (Fig. 2.1B), this melting is achieved by shearing of the hybrid that pulls 2–3 conserved 5' Us of the U-tract into the exit channel, permitting T_{hp} completion [6, 16] and EC destabilization [17]. Shearing is thermodynamically unfavorable at terminators like *t₅₀₀* with “imperfect” U-tracts (Fig. 2.1B). Instead, the T_{hp} exerts a pushing force on RNAP, driving forward translocation and shortening of the hybrid without nucleotide addition [16, 18]. Regardless of mechanism, after T_{hp} completion the RNA 3'-end is displaced from the active site, preventing catalysis and committing the EC to dissociation.

In contrast to those of the NA scaffold, the movements of RNAP that contribute to termination are poorly understood. RNAP is composed of several mobile parts whose coordinated movements catalyze the nucleotide addition cycle. Specific conformational changes are proposed to facilitate termination [19–24]. One hypothesis is that opening of the clamp is necessary for termination [22, 24, 25]. During initiation the clamp moves from an open to a closed state that then persists throughout elongation (Fig. 2.2A) [21, 22, 26, 27]. Clamp closure establishes RNAP–NA interactions that stabilize the EC [21, 26] and prevents transcription bubble collapse because double-stranded DNA cannot fit into the narrowed main cleft [22]. Clamp opening at a terminator could therefore facilitate termination by breaking the stabilizing RNAP–NA interactions, [21, 26] enabling bubble collapse and DNA release [19, 22, 23, 29, 30],

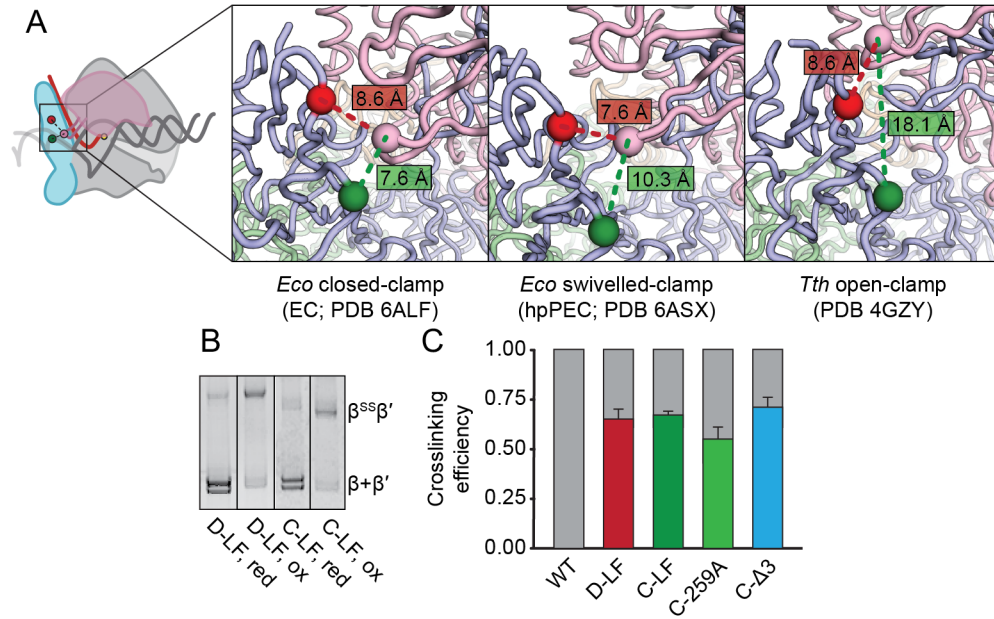


Figure 2.2. Forming disulfide crosslinks between the RNAP flap and lid. **(A)** View of cysteine substitutions and insertions within the flap and lid from the same angle across ECs with three distinct clamp conformations. Crosslinking between β C1044 (green sphere) and β' C258 (pink sphere) biases the clamp towards a closed-conformation, whereas crosslinking between β C843 and β' 258 biases displacement of the clamp into either the swiveled or open conformation. Panels depict all three cysteines sites for comparison of relative distances between C α atom but actual RNAP variants have only two cysteines. Closed-clamp EC, PDB 6alf [27], swiveled-clamp PEC, PDB 6asx [28], open-clamp PEC, PDB 4gzy [21]. **(B)** Representative non-reducing SDS-PAGE gel of D-LF and C-LF under reducing and oxidizing conditions that disfavor and favor disulfide formation, respectively. Crosslink formation causes a supershift of the crosslinked RNAP, allowing for measurement of crosslinking efficiency. **(C)** Crosslinking efficiencies of the five RNAPs used to investigate transcriptional elongation and termination commitment on the ligation scaffold. Error represents the standard deviation at $n \geq 3$.

suggesting that clamp movements occurs during the termination pathway. Studies of the his paused EC (which forms a pause RNA hairpin similar to the incomplete T_{hp}) suggest the hairpin promotes rotation of the clamp toward the RNA exit channel as part of a rigid body rotation of a swivel module that include the clamp [28]. However, it is unknown if clamp opening or swiveling is required for intrinsic termination.

A second hypothesis is that the T_{hp} , once formed, sweeps through the main channel to contact the folded trigger loop (TL), thus disrupting RNAP-scaffold contacts and inducing clamp opening [19]. More recent findings suggest that termination may proceed through multiple, alternative RNAP conformational states in a multistate-multipath (MS-MP) mechanism and that the role of the TL is to aid conformational flexibility required for this mechanism rather than to contact T_{hp} [14]. However, the role of clamp motions required during steps in the MS-MP mechanism remain unexplored.

To investigate the role of RNAP clamp and TL motions during intrinsic termination we developed an experimental approach in which ECs were reconstituted on long DNA templates tethered to paramagnetic beads, followed by treatment to restrict RNAP movements, either by (i) oxidation of a cysteine-pair engineered into variant RNAPs to suppress clamp movements, or (ii) addition of an antimicrobial peptide to block TL movement. Using this experimental approach we were able to test the contributions of RNAP conformational freedom to the termination mechanism.

3.3 Results

3.3.1 Restricting clamp movements by Cys-pair crosslinking alters elongation rate

To explore how clamp conformation impacts the intrinsic termination mechanism, we expanded a set of lid-flap cysteine pair (Cys-pair) variant RNAPs (Fig. 2.2A). The original two RNAP variants, C-LF (closed-clamp, lid-flap; β P1044C β' 258iC) and D-LF (displaced-clamp, lid-flap; β T843C β' 258iC), were used previously to characterize the role of clamp movements at the *his* pause, a hairpin-stabilized pause encoded in the histidine operon leader region [31]. These Cys-pair crosslinks connect the lid, which is a flexible β' module that extrudes from the clamp at the upstream fork junction (Fig. 2.1A) and facilitates separation of the nascent transcript from the template DNA strand [32], to the β flap, which forms one wall of the RNA exit channel. The C-LF crosslink stabilizes the clamp in a closed conformation [26, 27], whereas the D-LF crosslink biases the clamp away from the closed conformation and toward either the open [21] or swiveled conformations [21, 28, 33]. Since either clamp opening or swiveling could occur during termination, we designate D-LF as favoring clamp displacement.

We reasoned that the flexible nature of the lid loop could allow some clamp movement even in the presence of the C-LF crosslink. Thus, we created a C- Δ 3 Cys-pair RNAP with a three amino-acid deletion directly adjacent to the site of cysteine insertion at the tip of the lid loop. One of the deleted residues in this construct, R259, is proposed to contact the -9 base of the template strand in the post-translocated RNA-DNA hybrid [26, 27]. To distinguish whether any differences in elongation and termination behavior by C- Δ 3 were due to the Cys-pair crosslink or the loss of R259-NA contacts, we also constructed a C-259A RNAP (β P1044C β' 258iC R259A) that encodes a β' R259A substitution but is otherwise identical to C-LF. The C- Δ 3, C-259A, C-LF, and D-LF RNAPs all formed Cys-pair crosslinks with $\sim 70\%$ efficiently in an EC

reconstituted on an oligonucleotide scaffold after treatment with the oxidizing agent diamide (Fig. 2.2B–C see also Ref. [31]). C-259A formed crosslinks in $55 \pm 6\%$ of the EC (Fig. 2.2C). Because the Cys-pair crosslinking efficiency can vary dramatically among conditions, we determined the crosslinking efficiency in each experiment (see figure legends).

To analyze the effects of Cys-pair disulfides on termination, we developed an assay that allows crosslink formation in active ECs and subsequent testing of their effects on different terminators ligated downstream of the ECs (Fig. 2.3A) [34, 35]. RNAP was initially reconstituted on a short oligonucleotide scaffold bearing a template-strand 5' phosphate and 4-nt *Sty* I overhang for subsequent ligation. Downstream sequences of interest (*e.g.*, a pause signal or a terminator) were generated by PCR and tethered to paramagnetic streptavidin beads through a 5' biotin moiety on the template-strand PCR primer after *Sty* I-digestion of a site in added in the nontemplate-strand PCR primer. The G17-reconstituted EC and the bead-bound downstream sequences were ligated together followed by EC extension to the A26 position with ATP and [α - ^{32}P] GTP. The A26 ECs were then treated with diamide to generate Cys-pair crosslinks, transcription was restarted by addition of all four NTPs, and its progress along the DNA template was monitored by removing samples at various time points.

The baseline elongation rate of the uncrosslinked RNAPs varied (in reducing conditions; Fig. 2.3B). To facilitate comparisons, we normalized each elongation rate to that of the wild-type EC under reducing conditions. Under these conditions, which disfavor formation of disulfide crosslinks, both D-LF and C-LF RNAPs elongated more rapidly than the wild-type (~ 1.4 - and ~ 1.5 -fold faster, respectively). Treatment with diamide to form Cys-pair crosslinks perturbed the elongation rate of all four Cys-pair RNAPs but also affected the WT (uncrosslinked) RNAP. Under these oxidizing conditions, WT RNAP elongated ~ 1.5 -faster than it did under reducing

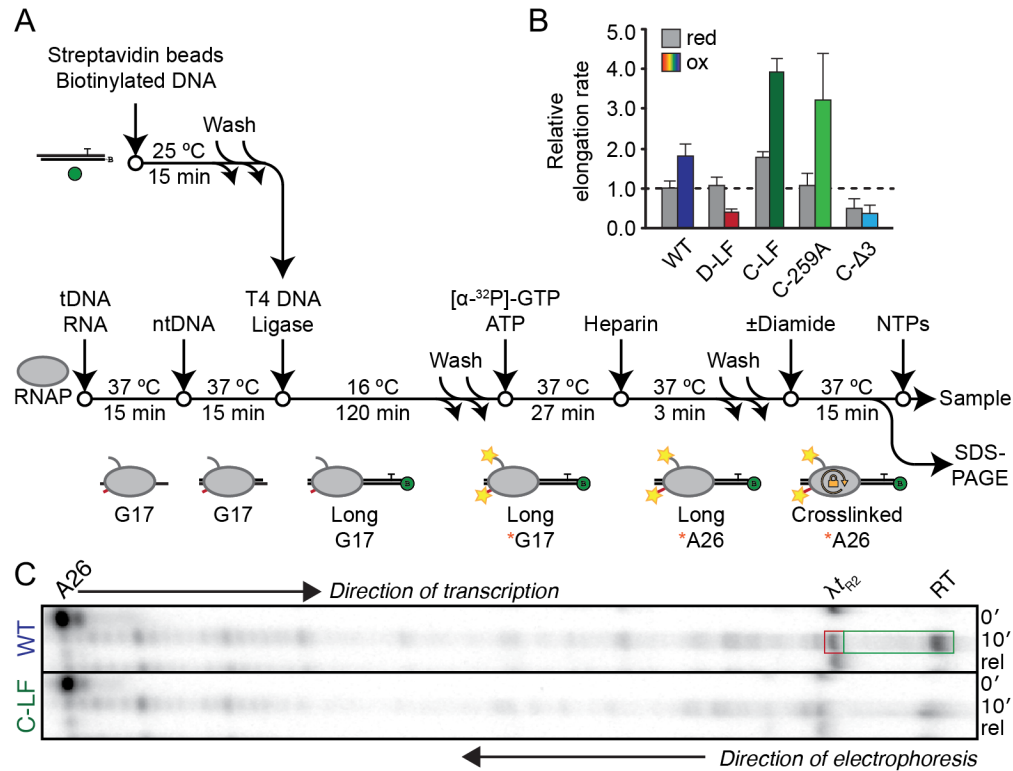


Figure 2.3. The bead-tethered ligation scaffold assay. **(A)** Assay schematic. Manipulations of the oligonucleotide scaffold and the RNAP are exploited to reduce the heterogeneity of complexes. BE, buffer exchange; the magnetic properties of the Streptavidin bead are exploited to eliminate unbound RNAPs and ECs that fail to ligate downstream DNAs of interest. **(B)** Elongation properties of Cys-pair RNAPs under reducing (grey bars) and oxidizing (colored bars) conditions. Relative elongation rates are normalized to that of the wild-type EC under reducing conditions to facilitate direct comparisons of elongation rates and subsequent evaluations of termination processes, as indicated by the dashed black line. Error bars indicate standard error as calculated by KinTek explorer, $n = 3$. **(C)** Representative urea PAGE *in vitro* transcription gels showing WT and C-LF RNAPs termination at λ_{R2}. ECs are radiolabeled and walked to the A26 RNA position. Samples were taken before and after (10 min) addition of NTPs. RT, termination readthrough. To demonstrate that termination was occurring, at the end of the reaction mixtures were also applied to a magnetic rack and RNAs released from the bead-tethered DNA template were collected. TE is calculated as the fraction of radioactive counts enclosed in the red box over the sum of radioactive counts in the red and green readthrough boxes.

conditions, suggesting an undefined role for redox state in the elongation behaviors of RNAP. Oxidation of the D-LF RNAP and biasing the clamp into the displaced conformation slowed elongation to an average rate only 17% of that of the reduced, WT RNAP, and 12% relative to its own rate under reducing conditions. C-LF RNAP, on the other hand, extended at a much faster rate relative to both the wild-type baseline (~ 3.2 -fold faster) and itself under reducing conditions (~ 2.1 -fold faster). In contrast, the C- $\Delta 3$ EC was impaired for elongation under both oxidizing and reducing conditions. Under reducing conditions, it elongated slower by a factor of ~ 1.7 -fold than wild-type. Upon formation of the crosslink, its average elongation rate was further reduced to $\sim 40\%$ that of the baseline wild-type rate, or $\sim 70\%$ of its rate under reducing conditions. C-259A, on the other hand, behaved more similarly to C-LF than to C- $\Delta 3$, elongating at nearly the same average rate as the wild-type under reducing conditions, and ~ 3.4 -fold faster upon formation of the crosslink. These results suggest that shortening of the lid loop of a Cys-pair crosslinked RNAP and not removal of R529 is principally responsible for slow elongation of the C- $\Delta 3$ RNAP.

Examination of the transcriptional behavior of these ECs suggested that changes in elongation rate upon disulfide formation were principally due to changes in pause behavior (Fig. 2.4). For instance, formation of the C-LF crosslink appeared to eliminate most weaker pauses along the DNA template and, in agreement with previous findings [28, 31], reduced the strength of the *his* pause. On the other hand, formation of the D-LF disulfide extended the efficiency and duration of weak pauses and caused the EC to dwell at the *his* pause for tens of minutes.

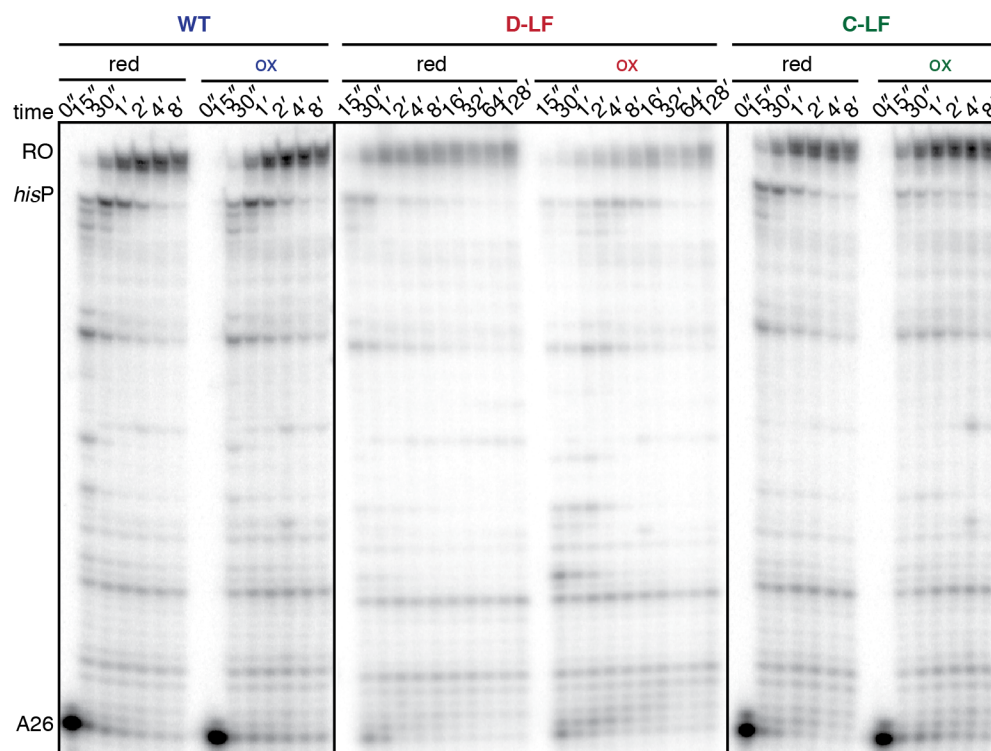


Figure 2.4. Constraining the clamp perturbs overall elongation rates through effects on relative pause strength.

Representative gels showing the transcriptional behavior of WT, D-LF, and C-LF ECs under reducing and oxidizing conditions. RO, template run-off.

2.3.2 Clamp conformation impacts termination commitment through effects on elongation

We next set out to determine the effects of clamp conformation on intrinsic termination efficiency (TE) of ECs that elongate on templates containing three model intrinsic terminators (Fig. 2.3A). By ligating one of three different linear DNA templates to the reconstituted ECs, we could monitor the termination behaviors of the same population of ECs across each of three model intrinsic terminators: t_{hisL} , λt_{R2} , and t_{500} . We chose these terminators because they represent three classes of intrinsic terminators, as categorized by the composition and thus strength of their U-tracts (Fig. 2.1B): perfect (t_{hisL}), near-perfect (λt_{R2}), and GC-containing (t_{500}). Predictably, TEs varied significantly across the three terminators. Under reducing conditions, the wild-type EC exhibited a TE of $31 \pm 3\%$ at t_{500} , $53 \pm 6\%$ at λt_{R2} , and $89 \pm 2\%$ at t_{hisL} (Fig. 2.3C, 4A).

Despite variability in U-tract composition and baseline TE, the general pattern of changes in TE after formation of a crosslink were consistent between λt_{R2} and t_{500} (Fig. 2.5A). Addition of diamide caused small reductions in the TE by the WT EC, despite its lack of crosslinking potential. Disulfide formation by D-LF greatly enhanced TE ($36 \pm 9\%$ to $64 \pm 6\%$ at λt_{R2} ; $23 \pm 6\%$ to $75 \pm 2\%$ at t_{500}). Crosslinking of C-LF, on the other hand, depressed the TE at both terminators ($38 \pm 20\%$ to $16 \pm 3\%$ at λt_{R2} ; $19 \pm 6\%$ to $14 \pm 2\%$ at t_{500}). C-259A again responded to diamide treatment similarly to C-LF at all three terminators (*e.g.*, $41 \pm 1\%$ to $19 \pm 4\%$ at λt_{R2}) whereas the C- $\Delta 3$ EC appeared to behave more similarly to D-LF, exhibiting improved TE at both terminators (*e.g.*, $51 \pm 4\%$ to $63 \pm 4\%$ at λt_{R2}). These same trends were also observed at t_{hisL} , but because the TE was too close to 100% ($\geq 90\%$), the fold effects on TE were too small for robust quantitative analysis.

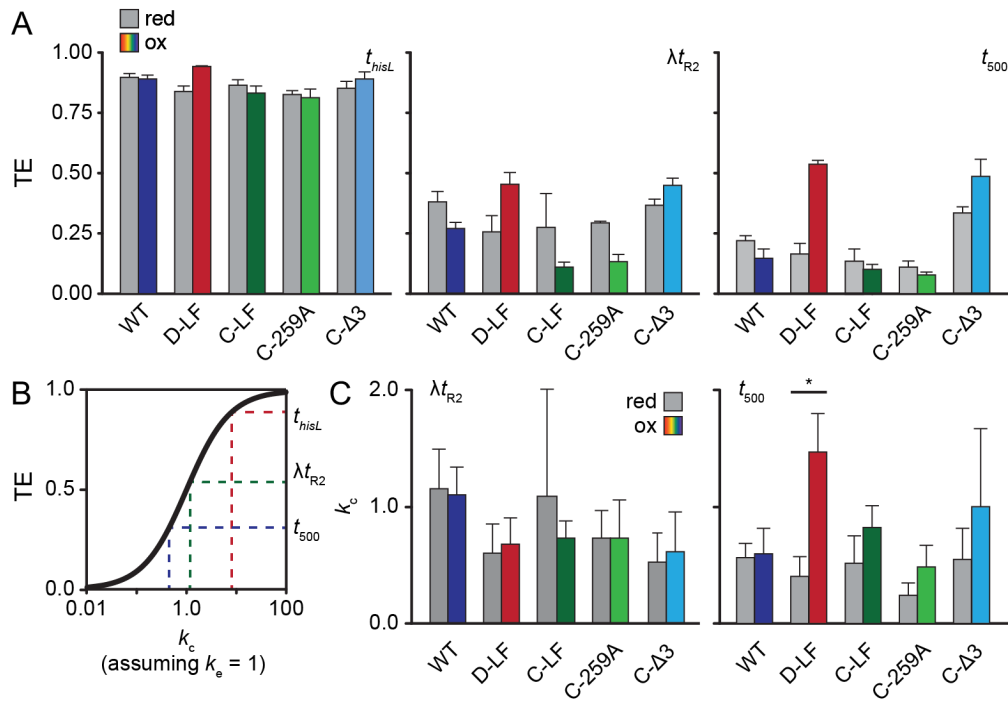


Figure 2.5. Evaluation of termination commitment rates. **(A)** Overall TE of Cys-pair RNAPs under reducing (grey bars) and oxidizing (colored bars) at three model terminators. **(B)** TE is a sigmoidal function of the relative rates of elongation k_e and termination commitment k_c , where $TE = k_c / (k_c + k_e)$. Shown is the effects on k_c where elongation rate $k_e = 1.0$. At low k_c , small increases in rate result in significant jumps in TE whereas at high k_c even large increases will have small effects on TE. The high TE of t_{hisL} precludes accurate calculations of k_c . **(C)** Conversion of TEs for λt_{R2} and t_{500} into k_c , where $k_c = (TE \times k_e) / (1 - TE)$. As both TE and k_e were independent measurements with an associated s.d. ($n = 3$), error bars here represent the propagation of uncertainty (see Methods). *, $p < 0.01$.

Comparison of the elongation rates of each enzyme with the TE revealed a general inverse correlation; faster enzymes terminator less and slower enzymes terminated more. We therefore calculated relative commitment rates for each RNAP at λ_{R2} and t_{500} using the simplifying assumption that commitment to termination is absolute so that TE is hyperbolically related to commitment rate ($TE = k_c / (k_e + k_c)$, where k_e is elongation rate and k_c is commitment to termination rate; Fig. 2.5B). At λ_{R2} , these analyses revealed two results of note (Fig. 2.5C). First, each RNAP exhibited a similar intrinsic commitment rate (k_c). Second, treatment with diamide and formation of a disulfide crosslink did not perturb this commitment rate significantly, suggesting that restricting clamp movements does not impact the commitment process during termination.

However, the results were somewhat different at t_{500} . Although treatment of the WT with diamide had little effect on commitment rate, formation of the D-LF, C-LF, C- $\Delta 3$, and C-259A crosslinks increased commitment rate with the difference becoming statistically significant for D-LF (0.32 ± 0.13 to 1.17 ± 0.26 ; $p > 0.01$ Fig. 2.5C). Taken together, these results suggest that changes in the clamp conformation impact TE but primarily by perturbing the transcription elongation rate, not the termination mechanism *per se*. However, for t_{500} , which is thought to terminate via the forward translocation pathway [16], biasing the clamp into a displaced conformation had an additional effect on the termination commitment rate itself.

2.3.3 RNA efficiently dissociates from an elongation complex at t_{hisL} regardless of clamp conformation

We next sought to explore how biasing clamp conformation with disulfide crosslinks and topologically closing the main cleft or the RNA exit channel affected the second step of the

pathway, dissociation. Because measures of TE can only report on the commitment component of the termination mechanism, we modified the Cys-pair ligation scaffold assay to enable simultaneous study of RNA and RNAP release from bead-tethered DNA templates. As before, we reconstituted RNAP on an oligonucleotide scaffold upstream of a model terminator. We used *t_{hisL}* because its high TE made measurements of release; the lower TEs of λt_{R2} and *t₅₀₀* precluded reliably extracting release data. RNAP was radiolabeled by phosphorylation of a Protein Kinase A/Heart Muscle Kinase (PKA/HMK) tag sequence at the C-terminus of β' by reaction with [γ -³²P] ATP and Protein Kinase A. Disulfide crosslinking was induced as before, after which all four NTPs were added to restart elongation and samples were removed at subsequent time points. At these time points, RNA or RNAP that was released from the paramagnetic bead-tethered DNA was collected after applying the reaction mixture to a magnetic rack for 15 s.

C-LF was the best behaved of the closed-clamp Cys-pair RNAPs and so was chosen for use in the dissociation assay. As an additional probe of release steps, we engineered a Cys-pair RNAP with cysteines in the β' rudder and the β protrusion (T-RP; $\beta A501C$, $\beta'S319C$) that, when crosslinked, topologically traps DNA within the cleft in addition to biasing clamp position (Fig. 2.6). Treatment of T-RP with diamide induces formation of a disulfide within the transcription bubble, forming a steric block to release of DNA from main channel. To promote release and make measurements feasible, we conducted assays in the presence of Cl^- (500 mM).

In these conditions, we achieved crosslinking efficiencies of $81 \pm 6\%$ for C-LF and $40 \pm 2\%$ for T-RP. Under reducing conditions, both WT and C-LF ECs elongated more slowly (Fig. 2.3B) but both RNAPs were able to elongate to *t_{hisL}* and terminate efficiently under these conditions within 20 s at 1 mM NTP (Fig. 2.7). We therefore chose to perform dissociation assays at 1 mM NTP.

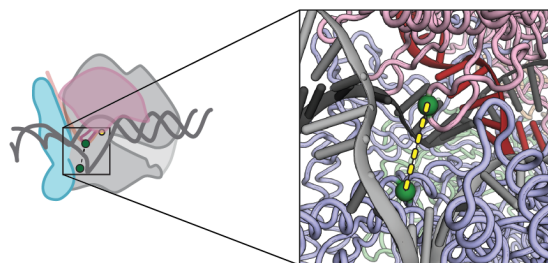


Figure 2.6. T-RP forms a disulfide crosslink within the transcription bubble. Left panel; position of the cysteine substitutions within the context of the entire EC. Right panel: zoomed-in view of the rudder-protrusion crosslink. C α atoms are depicted as green spheres. A full nucleic-acid scaffold is modelled in based upon previous structural predictions [36] (RNA, red; template DNA, black; nontemplate, gray).

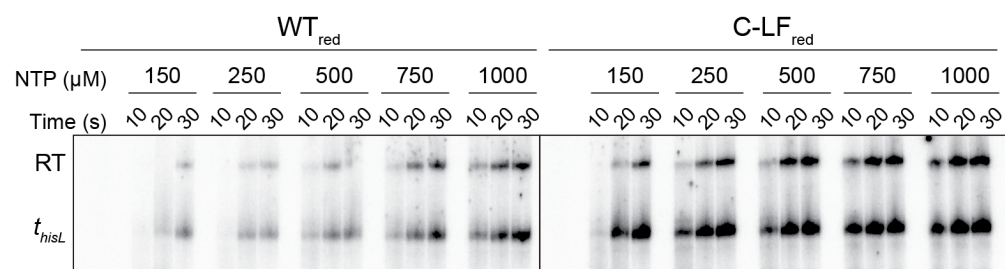


Figure 2.7. Elongation to *t_{hisL}* from A26 with different NTP concentrations under reducing conditions.

By labeling both the nascent transcript and the β' CTD, we could simultaneously monitor both RNA and RNAP dissociation from the bead-tethered DNA template (Fig. 2.8). Under reducing conditions, RNA rapidly dissociated from the wild-type EC at efficiencies of greater than 90% (Fig. 2.9A). When the same assay was repeated under oxidizing conditions, only a slight decrease in the overall transcript release was observed. Because the release values include RNA released from both crosslinked and uncrosslinked fractions and were subject to relatively high error, we used the time at which half of the maximal RNA release occurred as an indicator of effects on release rates rather than calculating release rates *per se*. Treatment of WT ECs with diamide caused slightly decreased RNA release rate by this metric (by a factor of ~ 1.4). For the Cys-pair crosslinks, we calculated the changes in half-maximal RNA release after treatment with diamide and divided these effects by the effect seen for wild-type RNAP to extract the effect of the crosslink. Using this method, the T-RP disulfide was found to release RNA similarly to the WT, while the C-LF disulfide modestly slowed the time of half-maximal RNA release (by a factor of 1.4). Both crosslinks produced only little effects on fraction RNA release. However, the T-RP crosslinked RNAP exhibited a slower phase of RNA release that could reflect inhibition of the rate of RNA release in the $\sim 40\%$ fraction of T-RP bearing the crosslink. Based on these results, we conclude that restricting the clamp in the closed conformation and covalently closing the RNA exit channel has at most a small effect on the rate of RNA release at an intrinsic terminator. In other words, RNA can release out the exit channel even when clamp movement is restricted and the lid is crosslinked in place. A slower rate of release may occur when DNA is topologically trapped in the T-RP crosslinked RNAP, but release still occurred efficiently. These results are consistent with a model in which terminated RNA exits out the RNA exit channel as

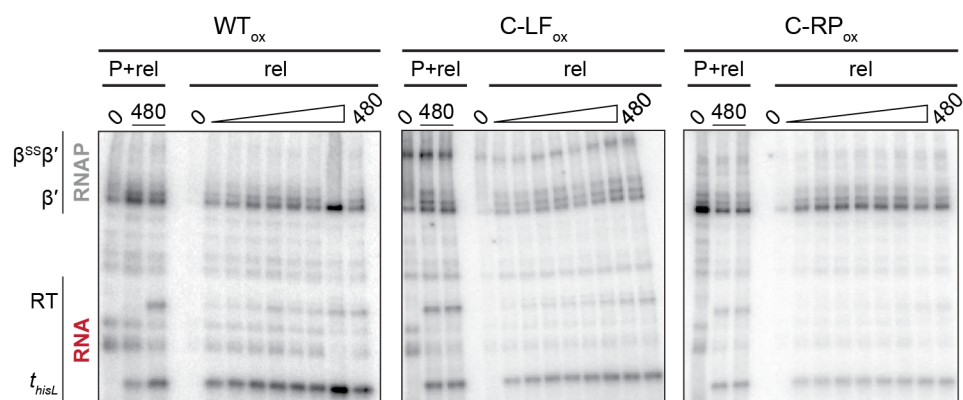


Figure 2.8. Representative dissociation assay non-reducing gels. P, bead-bound pellet; rel, sample released into the reaction supernatant. Release timepoints taken at 0, 20, 30, 40, 50, 60, 120, 240, and 480 s.

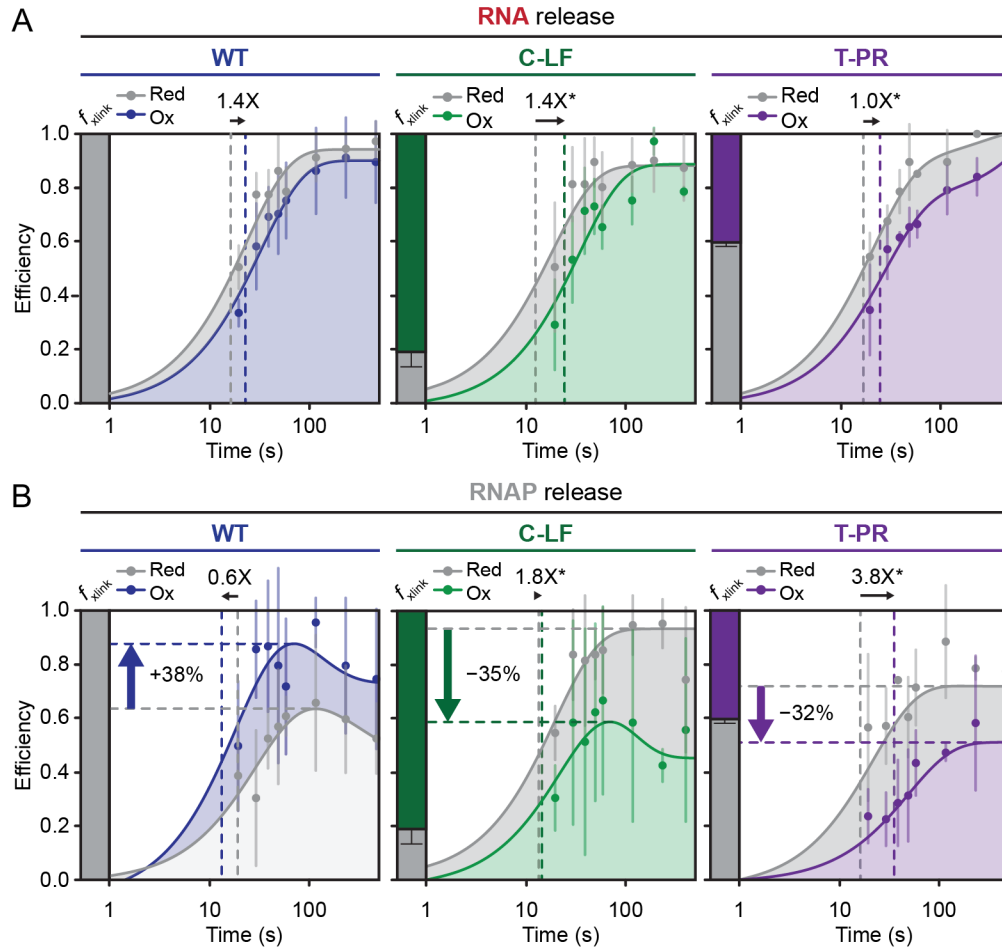


Figure 2.9. Evaluation of RNA and RNAP dissociation rates. **(A)** Semi-log plots of RNA release rates at *thisL*. RNA dissociation was determined as the fraction of RNA released into the reaction supernatant relative to total RNA at the terminator site. Data were fit to either a single exponential [$y = A(1 - e^{-kt})$] or a double exponential [$y = A_1(1 - e^{-k_1t}) + A_2(1 - e^{-k_2t})$] function. Grey curves represent fits for RNA release under reducing conditions; colored curves represent fits under oxidizing conditions. Bars to the left of plots represent crosslinking efficiencies under oxidizing conditions. Dashed lines represent the time at which half of the maximal released RNA was achieved over the time course. Arrows on top of plots represent the fold-change in the time it takes for half-maximal release to occur after treatment with diamide. Asterisks (*) indicate that the fold-changes upon oxidation of C-LF and T-RP have been normalized to the baseline fold-change of oxidation of the WT. Error represents the standard deviations at $n = 3$. **(B)** Semi-log plots of RNAP release at *thisL*. RNAP dissociation was determined as the fraction of RNAP released into the reaction supernatant relative to total RNAP within the reaction mixture multiplied by the TE. Data were fit to the double exponential function $y = y_0 + A_1(1 - e^{-k_1t}) + A_2(1 - e^{-k_2t})$. Plot components are the same as in subpanel A.

Percent changes in maximal release from the uncrosslinked conditions [$\% = 100 \times (rel_{ox} - rel_{red}) / rel_{red}$] are also depicted next to vertical arrows within the plot.

single-stranded RNA after dissociation from the RNA-DNA hybrid without a need for large clamp movements.

2.3.4 RNAP release from DNA is inhibited when clamp movement is blocked

Although clamp crosslinks had minimal effects on the rate of RNA release, it remained difficult to envision how RNAP could dissociate from the DNA template without more dramatic movements based on available EC structures. Initiation is thought to require an open clamp to admit DNA into the main cleft [22, 37-39], making a requirement for clamp opening to release DNA logical.

Despite the challenges in quantification, significant effects of clamp crosslinks on release of RNAP from DNA were obvious (Fig. 2.9B). Using the metric of fold-change in time to half-maximal release, we found that treatment of the wild-type EC with diamide actually *accelerated* the time to half-maximal release 1.7-fold and caused release of a greater fraction of RNAP. This increase in RNAP release is consistent with oxidized RNAP binding the NA scaffold somewhat more weakly, which could also explain its faster elongation rate (see Fig. 2.3B). In contrast, the C-LF and T-RP crosslinks both inhibited RNAP release from DNA. This effect is most obvious from the decrease in total amount of RNAP released, roughly proportional to the amount of crosslink. Further, although wild-type RNAP released faster after diamide treatment, the crosslinked RNAPs released more slowly (Fig. 2.9B). These results strongly suggest that crosslinking the clamp to either the flap or the protrusion strongly inhibits release of RNAP from DNA even though the RNA transcript can still release efficiently. Thus, RNA release without DNA release is clearly possible at an intrinsic terminator. This result has important implications for both the mechanism of termination and the fate of RNAP after termination (see Discussion).

2.3.5 Inhibiting RNAP conformational flexibility with Microcin J25 impairs both RNA and RNAP release

Kinetic analyses led to the development of a multistate-multipath (MS-MP) model of termination commitment in which restricting movements of the polymorphous TL motif slows the rate of termination by reducing the probability that ECs positioned at a terminator fluctuate by thermal motion into a state committed to termination [14]. However, experiments to date only probed effects of TL restriction on the commitment step and not the release step of termination. Our current clamp crosslinking results suggest that large motions of the clamp are not required for either irreversible commitment to termination or RNA release. To ask if TL-mediated fluctuations of the EC conformation other than large clamp movements are required for release, we used the transcription inhibitor microcin J25 (MccJ25). MccJ25 is a 21-residue antimicrobial lariat peptide [40] that binds and blocks the secondary channel, an element of RNAP thought to act as an “access tunnel” for incoming NTPs into the enzyme’s active site and as a regulator binding site (Fig. 2.10A). Based on the location of resistance mutations, MccJ25 is also thought to contact the TL, blocking catalysis by inhibiting TL motion [41-43]. We generated a fresh preparation of MccJ25 and used an abortive initiation assay (Fig. 2.11) to verify its activity. Under our experimental conditions, we determined that a fresh preparation of MccJ25 inhibited abortive initiation at *PlacUV5* with an IC_{50} of $0.63 \pm 0.03 \mu\text{M}$ (Fig. 2.11A), well within the order of magnitude of the published K_i of $1.4 \mu\text{M}$ [41].

We next used an established disulfide reporter assay that detects different TL conformations [44, 45] to determine whether MccJ25 restricts TL conformations (Fig. 2.12A).

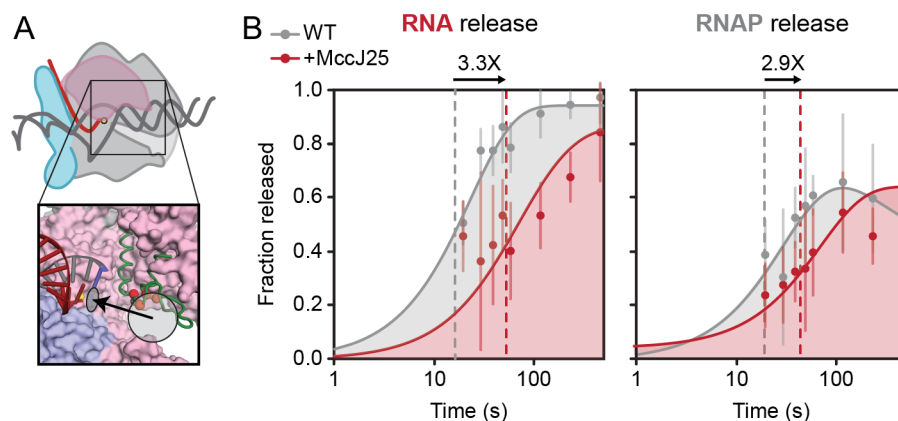


Figure 2.10. Evaluation of RNA and RNAP dissociation rates in the presence of MccJ25. **(A)** MccJ25 bind the TL and the RNAP secondary channel. Red spheres indicate the C α atoms of TL residues that confer MccJ25 resistance in mutational screens. The grey sphere indicates the C α atom of β' T931. The light grey circle indicates the location of the RNAP secondary channel and the black arrow indicates the path of NTP into the active site, light grey ellipse. The RNA-DNA hybrid is shown; other nucleic-acid components are omitted for clarity (red, RNA; grey, DNA; blue, template DNA base positioned in the RNAP active site). Core RNAP components are derived from PDB 6alf [27]; green, unfolded TL, PDB 2o5j. **(B)** Semi-log plots of RNA and RNAP release rates at *t_{hisL}*. When relevant, MccJ25 is added at 5 s after addition of chase NTPs. RNA dissociation was determined as the fraction of RNA released into the reaction supernatant relative to total RNA at the terminator site; RNAP dissociation was determined as the fraction of RNAP released into the reaction supernatant relative to total RNAP within the reaction mixture multiplied by the TE. Grey curves are the same as those in Fig. 2.9B–C; red curves represent RNA and RNAP release from WT RNAP in the presence of 100 μ M MccJ25. Dashed lines represent the time at which half of the maximal released RNA was achieved over the time course. Arrows on top of plots represent the fold-change in the time it takes for half-maximal release to occur upon treatment with MccJ25. Error represents the standard deviations at $n = 3$.

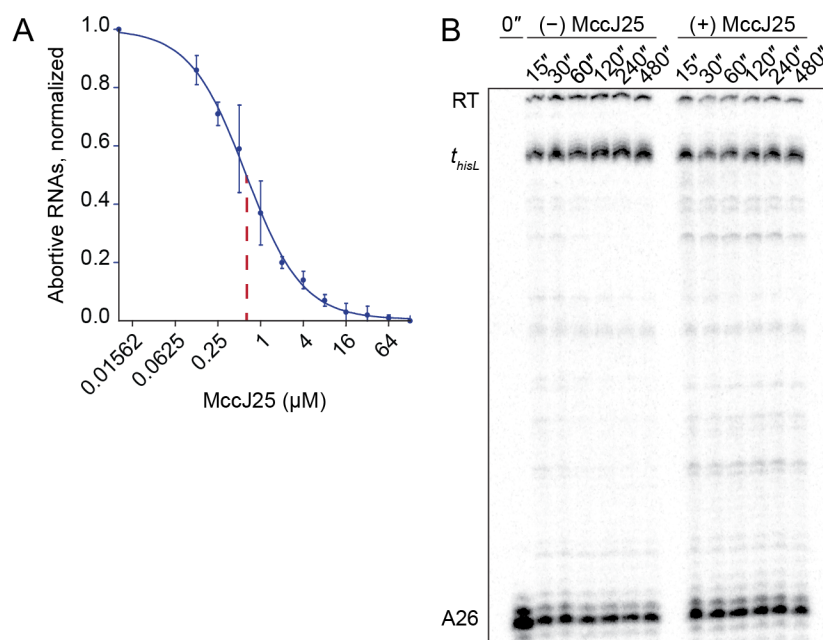


Figure 2.11. Effects of MccJ25 on RNAP during abortive initiation and elongation. **(A)** Abortive initiation by WT RNAP at *lacUV5* is inhibited by MccJ25. Abortive products were normalized to the maximum signal in the absence of MccJ25. **(B)** Representative gel showing that the addition of MccJ25 to *in vitro* transcription assays at 5 s is sufficient time to allow elongation by the majority of ECs to t_{hisL} under conditions of saturating NTP concentration. The presence and persistence of pause bands in the +MccJ25 samples indicates successful arrest of lagging ECs by the inhibitor.

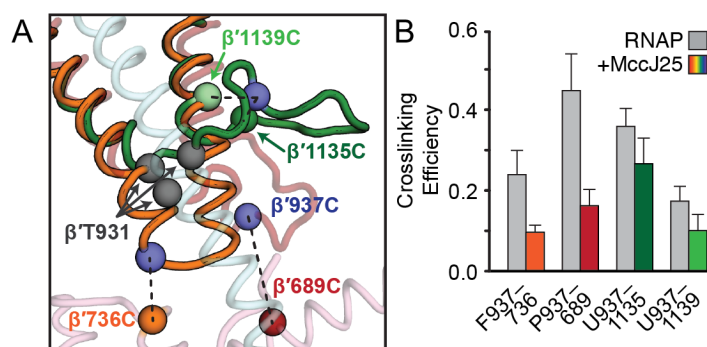


Figure 2.12. Cysteine-pair reporting suggests MccJ25 binds the TL in its unfolded conformation. **(A)** Location of TL cysteine-pairs. Three TL conformations observed in crystal structures were aligned to a cryo-EM structure of the *E. coli* EC [27]: orange, folded trigger helices, PDB 2o5j [8]; red, partially-folded TL, PDB 2nvq [46]; green, unfolded TL, PDB 1iw7 [38]. The bridge helix is depicted in cyan. Spheres depict the Ca atoms of relevant β' residues. Grey spheres indicate the position of β'T931, a residue that when mutated (T931I) confers full resistance to MccJ25 and is thus likely a critical MccJ25 binding determinant. **(B)** Crosslinking efficiencies of cysteine-pair reporters in the absence (grey bars) and presence (colored bars) of MccJ25. Error represents the standard deviation (n=3).

Because genetically determined binding determinants for MccJ25 overlap the TL, including the location of strong resistance mutations at TL residue T931, we anticipated that MccJ25 would limit TL movement and thus reduce crosslinking of the disulfide reporters or TL conformation. As expected, incubation of RNAP with MccJ25 severely reduced the crosslinking efficiency of Cys-pairs that detect the fully-folded trigger helices and a partially-folded conformation (Fig. 2.12B). MccJ25 effects on two Cys-pairs that report the unfolded conformation, however, were smaller, suggesting that MccJ25 positions the TL nearer to the unfolded conformation. The reduction in crosslinking by all four reporters of TL conformation confirms that MccJ25 strongly restricts TL conformational fluctuations.

Having established that MccJ25 blocks TL motion, we next asked if inhibiting TL motions would affect termination would affect the rate of RNA or RNAP release from DNA at *t_{hisL}*. We evaluated the effects of MccJ25 on dissociation at *t_{hisL}* using the same half-maximal dissociation metric used with the Cys-pair assay by careful addition after ECs had elongated to the termination site (Fig. 2.11B). MccJ25 slowed the rate at which half-maximal release of RNAP from template DNA was achieved more dramatically than did the closed-clamp crosslink, by a factor of 2.9. Even more strikingly, addition of MccJ25 also slowed the release of the nascent transcript by a factor of 3.3 (Fig. 2.10B). We conclude that TL motions increase the rate of RNA and RNAP release at an intrinsic terminator in addition to their roles in steps leading to commitment.

2.4 Discussion

Our study provides new understanding for the role of RNAP conformational changes in the mechanism of intrinsic termination. By developing a bead-tethered ligation scaffold assay, we

disentangled the steps of transcription elongation, commitment, and dissociation, which are conflated in the traditional measurement of TE, and gained key insights. First, although clamp conformation affects TE, these effects are due primarily to changes in the elongation rate rather than changes in the commitment rate except at CG-containing U-tract terminators (*e.g.*, t_{500}) that release RNA by forward EC translocation rather than hybrid shearing. Second, we found that clamp movements are much more important for dissociation of RNAP from DNA than for dissociation of RNA from the EC, which is the defining step in termination. Finally, we found that restricting movements of the TL dramatically slowed the rates of both RNA and RNAP release, consistent with a role for TL-mediated conformational fluctuations in the termination process [14].

2.4.1 Clamp conformation has stronger effects on elongation than termination

Structural and biochemical analyses of ECs suggest roles of the clamp in promoting EC stability *via* numerous non-specific interactions with the NA scaffold [26], in initiation by opening and closing to admit DNA to the cleft [22, 39], and in hairpin-stabilized pausing by aiding swiveling that inhibits TL folding and thus catalysis [28, 33]. Here we find that clamp conformation affects the efficiency of intrinsic termination, but primarily through effects on the competing rate of elongation rather than in the termination pathway *per se*. A disulfide bond that biased the clamp toward the displaced (swiveled) conformation observed in paused ECs or a small deletion in the lid slowed elongation, whereas stabilizing the clamp in the closed conformation increased elongation. These effects on elongation, all of which caused corresponding changes in TE, are likely caused primarily by changes in pausing, which was evident in transcription elongation

patterns (Fig. 2.4). However, we cannot exclude effects on on-pathway elongation (*e.g.*, translocation), which will require more detailed studies of elongation to discern.

The key point is that clamp conformation affects pausing/elongation much more than it affects the pathway of RNA release at an intrinsic terminator, which has important implications. Bacterial intrinsic terminators have apparently evolved to be robust relative to EC conformation. Whereas pausing and the overall rate of elongation may be readily modulated by clamp-binding elongation factors (*e.g.*, RfaH, NusG, Spt4/5, hepatitis δ antigen; [47-49]), the resistance of canonical bacterial terminators to clamp conformation may enable intrinsic terminators to halt transcription irrespective of EC composition. A case in point could be ribosomal RNA operons, which are transcribed by antitermination-modified ECs [50, 51] but use intrinsic terminators to define the 3' ends of the transcription units. Conversely, the effect of clamp conformation on forward-translocation-dependent terminators may enable regulation of termination at these such sites by regulators that affect clamp conformation.

The lack of effect of clamp and lid crosslinking on RNA release also provides further reason to doubt the hairpin-TL contact model on intrinsic termination [19]. This model, which also has been questioned on other grounds [14, 17, 52], requires the terminator hairpin break RNAP interactions by sweeping through the cleft, starting by breaking the lid-flap interaction. Since RNA release is unaffected when the lid is crosslinked to the flap, this model seems implausible.

Deletion of three residues from the tip of the lid loop slowed elongation/increased pausing even under reducing conditions that disfavored a disulfide that constrained the clamp. A close inspection of recent structures of the *E. coli* EC suggest that this effect may be due to remodeling of the upstream fork junction, where the RNA-DNA hybrid is separated, single-

stranded RNA is threaded into the RNA exit channel, and template DNA reanneals with the nontemplate strand [27]. The C- Δ 3 RNAP was constructed by removing β' L255, D256, and R259, all of which are conserved across diverse species. L255 interacts with the -9 base pair of the hybrid, D256 forms a salt bridge with β R841 in the flap to seal the RNA exit channel, and R259 forms a base stack with the -10 base of the template strand. Each of these interactions plays a role in forming the proper geometry of the upstream fork junction and they have been observed to shift upstream with the NA scaffold in a translocation intermediate [28]. The C-259A variant RNAP demonstrated that a single substitution in this region did not significantly perturb elongation, but it is likely that deletion or mutation of all three of these conserved residues would have additive or synergistic effects on the architecture of the fork junction. The deleted residues may ordinarily guide strand exchange and thus translocation at the fork junction. Thus, the C- Δ 3 deletions may slow elongation or stimulate pausing by inhibiting translocation.

We also provide evidence that the general redox state of the EC impacts its elongation rate, irrespective of the presence of a disulfide crosslink. Treatment of the wild-type EC with diamide increased its overall elongation rate nearly two-fold. RNAP contains 28 cysteine residues, eight of which play a role in binding Zn^{2+} ions, and one of which is in close proximity to the active site. It is possible that chemical modification of any of these cysteines could impact the elongation rate of the EC by altering RNAP conformation in a way to stimulate catalysis (*e.g.*, promoting TL folding) or by indirectly affecting NA interactions. These results highlight the need for in-depth investigation into the role of redox state on RNAP.

2.4.2 Upstream-for junction mobility may weaken terminators that require forward translocation for commitment

Although EC inactivation/commitment at terminators for which hybrid shearing is a viable route (t_{hisL} and λ_{R2} ; [16]) was largely unaffected by biasing clamp conformation with disulfide crosslinks, this process was affected by crosslinks for a terminator (t_{500}) that uses forward translocation without nucleotide addition as the path to EC inactivation/commitment (Fig. 2.5B). This result suggests that hybrid shearing can occur without major changes in clamp conformation and that for these hybrid-shearing terminators any role of a hairpin-stabilized pause intermediate, for which crosslinks disfavoring clamp swiveling is inhibitory [28], must be simply to increase the time window for shearing to occur and not be required for the inactivation process itself. That a forward-translocation terminator exhibits different responses to the disulfide crosslinks further underscores the existence of this alternative route to termination at terminators with GC-containing U-tracts [18].

The effect of disulfide crosslinks at this class of terminators with GC-containing U-tracts appears more related to restricting mobility of the lid than for a role of any particular clamp conformation. We draw this conclusion because crosslinks that bias the clamp toward either displaced (D-LF) or closed (C-LF) conformations both increase the rate of commitment at t_{500} (Fig. 2.5C), indicating that no particular clamp conformation favors or disfavors forward translocation and instead that restricting lid mobility in any state increases termination. Additionally, the non-crosslinked form of the lid Cys insertion combined with R259A decreased termination significantly at t_{500} , further indicating that the lid may influence this class of terminators. In support of this hypothesis, the lid appears to have pronounced effects on forward translocation of RNAP adjacent to a terminator hairpin [53]. Based on these observations, we

speculate that restricting lid mobility with disulfide crosslinks may destabilize the upstream fork-junction because lid mobility stabilizes the hybrid as it shifts translocation register. For terminators like t_{500} that require forward translocation unlinked to nucleotide addition to commit to termination, this hybrid destabilization may increase the overall rate of commitment.

2.4.3 Clamp opening appears to be required for RNAP release from DNA at intrinsic terminators

It is structurally consistent for an EC to become catalytically inactivated without dramatic movements of the clamp. However, it is less easy to understand how RNAP could dissociate from template DNA without clamp movements. Reannealing of the transcription bubble would generate a DNA duplex too large for accommodation within the main cleft of RNAP in either the closed or swiveled conformation; thus, current structural understanding of the EC requires large movements of the clamp along the axis perpendicular to the hybrid to allow dissociation of RNAP from the template [22]. However, the same structural data suggest that such large-scale clamp motions should not be necessary for release of the nascent transcript, which can release out the exit channel once the hybrid is shortened to an unstable length.

Our studies of the dissociation step in termination support these views. Formation of a crosslink that stabilized a closed clamp had little or no effects upon the rate or efficiency of RNA release from an efficient terminator. However, the rates and efficiencies at which RNAP dissociated from the template were both reduced under conditions that disfavor clamp displacement, suggesting a need for orthogonal opening of the clamp during dissociation. These effects were most dramatic when a rudder-protrusion crosslink was formed within the transcription bubble to induce a steric block to complete reannealing of duplex DNA (Fig. 2.9B).

Taken together, these results indicate that whereas RNA can easily shear out of the RNA exit channel at t_{hisL} in all tested conditions, RNAP release from template DNA depends both upon the clamp's ability to move orthogonally from the main cleft (*i.e.*, “open”) and upon the template strand's ability to reanneal with the non-template strand in order to induce collapse of the transcription bubble [54].

A further implication of our finding that RNA release at a terminator can occur without RNAP release from DNA is that RNAP may remain bound to DNA after termination in some cases. If an unreleased RNAP were still associated with a σ initiation factor, as has been observed to occur in some situations [55-57], then it may be able to directly enter a promoter search mode without release from the DNA.

2.4.4 TL movement is required for release of both RNA and DNA at an intrinsic terminator

Previous studies of the role of the TL in intrinsic termination mechanism led to a model in which termination was promoted not by a TL contact to the terminator hairpin but by conformational fluctuations of the TL that allowed an EC located at a terminator to sample multiple conformations [14]. This multistate-multipath (MS-MP) mechanism was proposed as an extension to the classical thermodynamic model of termination as a competition between elongation *vs.* termination [5]. Our findings that restricting the movements of the TL with the antimicrobial peptide MccJ25 slows the dissociation rates from template DNA are consistent with the MS-MP model, although they do not distinguish effects on commitment *vs.* dissociation. Because we added MccJ25 to ECs already located at a terminator but not yet dissociated, because MccJ25 would inhibit subsequent elongation, and because our readout was release of

RNA or RNAP, it is possible that the effects of restricting TL movements with MccJ25 could be manifest at either the commitment or dissociation steps in the mechanism. Regardless, these effects are striking and suggest that restricting TL movement stabilizes ECs located at a terminator, disfavoring termination. During termination commitment, the 3'-end of the RNA transcript should be displaced from the active site of RNAP by at least 2–3 base pairs [17], positioning the last base several Ångstroms from the TL in either its folded or unfolded conformation (Fig. 2.13). Possibly, restricting TL movements slows this displacement, as the folded form of the TL is thought to contact or at least stabilize RNA in the pretranslocated register.

Alternatively or additionally, limiting TL movement may simply reduce thermal motions of the complex set of interconnected modules in RNAP whose normal fluctuations aid release of RNA and DNA from RNAP (the core idea in the MS-MP model). These limitations appear to be much greater when MccJ25 is bound than when the clamp is restricted by a disulfide crosslink since the former but not the latter slows RNA release.

2.4.5. Conclusion

Our findings confirm that the release step of termination mirrors initiation in requiring clamp movement to allow DNA escape from main channel of RNAP just as it is required for DNA entry during initiation. The fact that RNA can release and transcription can terminate even when RNAP remains bound to the DNA opens the possibility that RNAP could enter a nonspecific sliding mode on DNA without fully dissociating at a terminator. Such an alternative to the conventional view that RNAP re-associates with initiation factors after release at a terminator

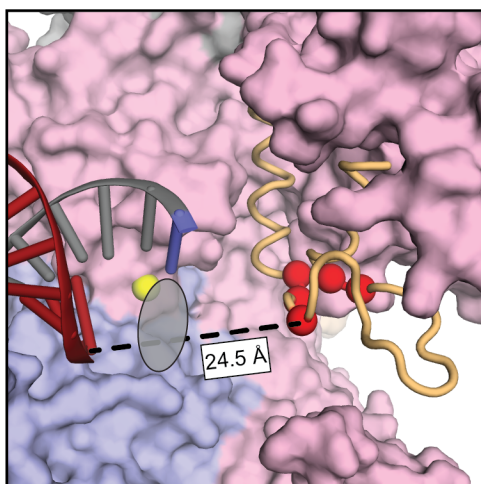


Figure 2.13. Distance between the TL and the RNA 3'-end as is proposed to occur during termination commitment.

View is the same view as depicted in the left panel of Fig. 2.10A rotated 10° along the y-axis. The -1 and -2 RNA bases have been hidden to model the proposed inactivated EC conformation. Distances are measured from the phosphate atom of the “-1” RNA base of the committed EC (-3 base in the actual structure) and the arbitrarily chosen C α atom of β 'R1239.

and then enters the promoter search could have important implications for transcriptional regulation. For example, direct transfer of RNAPII from terminators to promoters in “gene loops” has been proposed to occur in yeast [58]. Retention of RNAP on DNA after termination could help enable this type of localized transcription mechanism.

2.5 Materials and Methods

2.5.1 Materials

All DNA and RNA oligonucleotides were from IDT (Table 1). Oligonucleotides used for reconstitution of elongation complexes were gel purified by denaturing PAGE. HPLC-purified NTPs were from Promega Corp. [α - ^{32}P] GTP and [γ - ^{32}P] ATP were from Perkin-Elmer. T4 *Sty* I-HF and T4 DNA Ligase were from NEB, Ipswich, MA.

2.5.2 Protein and peptide purifications

Wild-type and variant RNAPs were overexpressed from plasmids (Table 2). The parent plasmid pRM843 is a T7 RNAP-based overexpression plasmid encoding *E. coli* *rpoA*, *rpoZ*, *rpoB* and *rpoC*, yielding core RNAP. The N-terminus of β contains a His₁₀ tag followed by a PreScission protease cleavage recognition signal. The C-terminus of β' contains a *Strep*-tag[®] preceded by a Protein Kinase A (PKA) recognition signal. RNAP and σ^{70} were purified from *E. coli* BLR DE3 transformed as previously described [31, 59] and dialyzed into a high [DTT] RNAP storage buffer (10 mM Tris-Cl, pH 7.9, 20% [v/v] glycerol, 100 mM EDTA, 10 mM DTT, 100 mM NaCl, 1 mM MgCl₂, 20 μ M ZnCl) to ensure the integrity of the cysteine sulfhydryl groups. Aliquots were stored at -80°C . RNA polymerase holoenzymes used in MccJ25 abortive

Table 2.1. Oligonucleotides used in this study

Oligonucleotide name	Sequence	Purpose
RNA-8855	5' -UUUUUUUGAGCUAGAGG-3'	Reconstitution scaffold, RNA component
tDNA-10160	5' -/5Phos/CAAGGGACTGGTC TGAATCGGGAATACGGAACGGAA GATCTGAGTTCTCTTCCCCTCTA GCTCAGGACGTACTGACC-3'	Reconstitution scaffold, template DNA component
ntDNA-10161	5' -GTCAGTACGTCCTGAGCTAG AGGGGAAGAGAACTCAGATCTTC CGTTCCGTATTCCCGATTTCAGAC CAGTCC-3'	Reconstitution scaffold, nontemplate DNA component
asDNA-12001	5' -AACAGGTCGATCGGAAGTTG GTAACGTCTTCCATACAACCTCC TACTACATGCAACC-3'	Antisense DNA oligo used to enhance TE at λ_{R2} and t_{500}
asDNA-12952	5' -TTGTCCAGCTAGCCTTCAAC CATTGCAGAAGGTATGTTGGAGG AATGATGTACGTTGG-3'	Antisense DNA oligo used to enhance TE at t_{hisL}
dsDNA-11033	5' -/5Phos/CTTGCCTGACTAG TCTTTCAGGCGATGTGTGCTGGA AGAGATTTCAGAGTCTTCCAGTGG -3'	Template-strand dsDNA used to determine EC elongation rates
dsDNA-11034	5' -/5Biosg/CCACTGGAAGAC TCTGAATCTCTTCCAGCACACAT CGCCTGAAAGACTAGTC AGG-3'	Nontemplate-strand dsDNA used to determine EC elongation rates

primer-10401	5'-GGTGGTCCTTGGCTCTGGCG GTGATAATGGTT-3'	Forward primer for amplification of λ_{tR2} and t_{500}
primer-10402	5'-/5Biosg/CGCTTCTGCGTT CTGATTTA-3'	Reverse primer with 10401
primer-10517	5'-GTTGTTTCCTTGGAACACCA CCATCATCACCA-3'	Forward primer for amplification of t_{hisL}
primer-10518	5'-/5Biosg/CGGCCTGATTTC TGAATAGC-3'	Reverse primer with 10518
primer-12285	5'-CCCGAAAAGTGCCACCTGAC G-3'	Forward primer for amplification of $PlacUV5$
primer-12286	5'-CAGTTCCTACTCTCGCATG G-3'	Reverse primer with 12285

Table 2.2. Plasmids used in this study

Plasmid name	Source	Purpose
pRM843	[31]	Overexpression of wild-type RNAP (<i>rpoA</i> , <i>rpoZ</i> His ₁₀ :PPX: <i>rpoB</i> , <i>rpoC</i> :PKA: <i>strep</i>)
pRM890	[31]	pRM843 derivative. Overexpression of D-LF RNAP (β T843C β' 258iC)
pRM918	[31]	pRM843 derivative. Overexpression of C-LF RNAP (β P1044C β' 258iC)
pRM1100	This paper	pRM843 derivative. Overexpression of C- Δ 3 RNAP (β P1044C β' Δ 255–256,258iC, Δ 259)
pRM1177	This paper	pRM843 derivative. Overexpression of C-259A RNAP (β P1044C β' 258iC R259A)
pRM1001	This paper	pRM843 derivative. Overexpression of T-RP RNAP (β A501C β' S319C)
pDJ115	[44]	Overexpression of P937-687 (β' I937C P1135C :: PKA :: His ₁₀)

pDJ116	[44]	Overexpression of P937-687 (β' I937C T1135C :: PKA :: His ₁₀)
pDJ124	[44]	Overexpression of F937-746 (β' Q736C I937C :: PKA :: His ₁₀)
pDJ129	[44]	Overexpression of P937-687 (β' G687C I937C :: PKA :: His ₁₀)
pET-28 derivative	Gift from R.B.	Overexpression of σ^{70} (His ₆ : <i>rpoD</i>)
pJP73	Gift from A.J.L.	Overexpression of MccJ25
pIA171	[60]	PCR template for amplification of <i>thisL</i>
pRM1065	This paper	PCR template for amplification of λ_{tr2}
pRM1066	This paper	PCR template for amplification of t_{500}

initiation control assays were formed by incubating 2 μ M purified core enzyme with 4 μ M σ^{70} in RNAP storage buffer. Mixtures were incubated at 30 °C for 30 minutes and stored at –20 °C.

MccJ25 was purified from the pJP73 overexpression plasmid (a gift from A. James Link, Princeton University, NJ) by a protocol adapted from [61]. Briefly, DH10B cells containing the pJP73 plasmid were grown up in LB overnight at 37 °C at 250 rpm. Cultures were diluted to a starting OD₆₀₀ of ~0.02 in M9 salts supplemented with thiamine, glucose and all 20 amino acids, and grown at 37 °C at 250 rpm up to a final OD₆₀₀ of 0.5. The *mcjABCD* operon was then induced with 1 mM IPTG and grown overnight for ~20 hrs at 37 °C at 250 rpm. Cells were pelleted and MccJ25-containing supernatant collected. MccJ25 was purified from the supernatant on BondElut C8 cartridges (from Varian, now acquired by Agilent Technologies, Santa Clara, CA) following manufacturer instructions. The elutions were dried by rotovap, followed by speedvac, and dissolved in water. Purity was checked by resolving products on a homogenous 20% denaturing PhastGel (GE Healthcare), and Coomassie brilliant blue R-250 staining.

2.5.3 Producing linear DNA fragments for transcription

Linear DNA templates were generated from plasmids *via* PCR (Table 1 & 2) using OneTaq[®] DNA polymerase and purified *via* spermine precipitation. PCR products were incubated with 10 mM spermine at 4 °C for 15 minutes followed by high-speed centrifugation and removal of the supernatant. Samples were incubated with spermine exchange buffer [75% (v/v) ethanol, 300 mM sodium acetate, 10 mM Mg(OAc)₂] at 4 °C for 1 hr followed by another high-speed centrifugation for 5 min. The supernatant was removed and the pellet was washed with 70% (v/v) ethanol and centrifuged as before. Ethanol was evaporated off *via* speed-vac and the pellet was dissolved in TE. Sample concentration was measured using a NanoDrop Spectrophotometer

(Thermo Fisher Scientific Inc., Waltham, MA). DNA fragments were processed with *Sty* I-HF restriction endonuclease following the manufacturer's protocol and cleaned through a second round of spermine precipitations.

2.5.4 Transcription on fully complementary oligonucleotide scaffolds

ECs were reconstituted step-wise. First, excess template strand (10 μ M) was annealed to 5 μ M RNA in reconstitution buffer [10 mM Tris-OAc, pH 7.9, 40 mM KOAc, 5 mM Mg(OAc)₂] by heating to 75 °C for 2 min, rapidly cooling to 45 °C, and then cooling to room temperature in 2 °C/2-min steps. This initial scaffold was diluted to 500 nM by addition of elongation buffer (20 mM Tris-OAc, pH. 8.0, 40 mM KOAc, 5 mM Mg(OAc)₂, 0.1 mM EDTA, 1 mM DTT) followed by addition of 2.5 μ M RNAP and incubation at 37 °C for 15 min. Non-template strand was then added to 1.5 μ M and the sample was again incubated at 37 °C for 15 min, yielding an EC formed with the molar ratio 1 RNA:2 tDNA:3 ntDNA:5 RNAP. PCR fragments were pre-bound to paramagnetic beads through a 5' biotin moiety on the template strand. 150 nM bead-bound fragments were mixed with 50 nM ECs, a 3:1 ratio, and ligated together with 2,000 U T4 DNA ligase and 1 mM ATP in EB at 16 °C for 2 hr. Non-ligated ECs and excess RNAP were removed from the mixtures through two EB washes on a magnetic rack. RNAs were labeled by addition of 0.2 mCi [α -³²P] GTP and a heparin competitor was added to 50 μ g/ml. Disulfide crosslinking of the Cys-pairs was initiated by adding diamide to 15 mM and incubating ECs at 37 °C for 15 min. To determine crosslinking efficiencies, aliquots were removed, quenched with 15 mM iodoacetamide, and separated through SDS-PAGE on 8% Bolt gels in 1X Bolt running buffer, and quantitated *via* PhosphorImaging. The remaining sample was diluted in EB $\pm$ oxidant to a working concentration of 5 nM, and a DNA oligonucleotide antisense to upstream sequences was

added to 1 mM to prevent formation of secondary structures that could block T_{hp} formation. Transcription was initiated by saturation with 150 μ M ATP, UTP, and CTP and 30 μ M GTP. During endpoint assays, reactions were stopped at 10 min by addition of a urea stop buffer (final concentration 4 M urea, 44.5 mM Tris-borate, pH 8.3, 50 mM EDTA, 0.01% [v/v] xylene cyanol, and 0.01% [v/v] bromophenol blue). In timepoint assays, reactions were likewise stopped at 15, 30, 60, 120, 240, 480, 960, and 1920 s. Samples were separated by electrophoresis in a denaturing 8% (v/v) polyacrylamide gel and quantitated *via* PhosphorImaging.

2.5.5 Dissociation assays

Except where here noted, release assays were performed as above. To promote dissociation, RNA and template-strand DNA were annealed in Cl^- reconstitution buffer (10 mM Tris-Cl, pH 7.9, 40 mM KCl, 5 mM $MgCl_2$) and all subsequent steps were performed in Cl^- elongation buffer (20 mM Tris-Cl, pH. 8.0, 40 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT). The Protein Kinase A (PKA) tag on the β' CTD was radiolabeled by addition of 7.5 U PKA and 0.5 mCi [γ - ^{32}P] ATP/ μ l simultaneously with incorporation labeling of the RNA. This step was performed after reconstitution to block access of PKA to the rifampicin-binding pocket, which can serve as a second labeling site on RNAP. To reduce the possibility of aggregation of released RNAPs, 1 μ M double-stranded DNA was used as a competitor instead of heparin. Antisense DNAs were not added in these assays. Transcription was restarted by addition of 500 mM NaCl and all four NTPs to 1 mM. Reactions were stopped with an equal volume urea/SDS stop buffer (final concentration 4 M urea, 2% [w/v] SDS, 44.5 mM Tris-borate, pH 8.3, 50 mM EDTA, 0.01% [v/v] xylene cyanol, and 0.01% [v/v] bromophenol blue) and separated in denaturing 8% SDS Bolt gels and quantitated *via* PhosphorImaging as before.

2.5.6 MccJ25 abortive initiation assays

In vitro abortive initiation transcription assays were performed by incubating 150 nM wild-type RNAP σ^{70} holoenzyme with 75 nM *lacUV5* promoter-containing DNA template in Cl^- elongation buffer at 37 °C for 15 min to form 75 nM RP_c . MccJ25 was titrated into aliquots followed by incubation for an additional 5 min at 37 °C. Finally, an NTP mix was added to each aliquot for a final concentration of 25 nM RP_{ITC} , 0–128 μM MccJ25, 500 μM ApA, 50 μM UTP, and 0.5 mCi [α - ^{32}P] UTP/ μl . Reactions were stopped after 10 min by addition of an equal volume of urea stop buffer, separated by electrophoresis in a denaturing 23% (v/v) polyacrylamide gel, and quantitated *via* PhosphorImaging. Abortive RNAs were mapped as in Borowiec and Galla [62].

2.5.7 Trigger loop cysteine-pair reporter assays

MccJ25 (100 μM) was added to 50 nM RNAP and incubated for 5 min at 37 °C. Disulfide formation was initiated by treatment with 0.25 mM DTT and 10 mM CSSC ($E_h = -0.16$ mV) for 60 min at 37 °C. Reactions were quenched with 50 mM iodoacetamide and separated by non-reducing SDS-PAGE using 8% Bolt Bis-Tris Plus gels over two hours. Gels were stained with Krypton Fluorescent Protein Stain (Thermo Fisher Scientific) according to the manufacturer's protocol and imaged with a 520-nm excitation filter 580-nm emission filter,. Crosslinking efficiency was determined using ImageQuant software. Experimental error was determined as the standard deviation of measurements from three independent replicates.

2.5.8 Determination of commitment rates and propagation of error

TE was calculated as in Reynolds and Chamberlin [63] and redefined in terms of rate constants as in von Hippel and Yager [64] such that:

$$(1) \quad TE = \frac{k_c}{k_c + k_e}$$

and

$$(2) \quad k_c = \frac{TE \times k_e}{1 - TE}$$

where k_c represents the termination commitment rate and k_e represents the elongation rate. To determine k_c , we first extracted overall elongation rates using KinTek Explorer and next normalized them such that under reducing conditions the WT RNAP had the arbitrary $k_e = 1$ (see Fig. 2.3B). Using Eq. 2 we then determined k_c for each RNAP under both reducing and oxidizing conditions.

Because TE and k_e were separate measurements with associated error, determination of k_c required propagation of uncertainty, which was performed as follows:

$$(3) \quad \sigma_{k_c} = \sqrt{\frac{\sigma_{TE}^2 \times k_e^2 + (TE - 1)^2 \times TE^2 \times \sigma_{k_e}^2}{(TE - 1)^4}}$$

2.6. References

1. Peters, J.M., et al., *Rho and NusG suppress pervasive antisense transcription in Escherichia coli*. Genes Dev, 2012. **26**(23): p. 2621-33.
2. Helmrich, A., et al., *Transcription-replication encounters, consequences and genomic instability*. Nat Struct Mol Biol, 2013. **20**(4): p. 412-8.

3. d'Aubenton Carafa, Y., E. Brody, and C. Thermes, *Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures*. J Mol Biol, 1990. **216**(4): p. 835-58.
4. Wilson, K.S. and P.H. von Hippel, *Transcription termination at intrinsic terminators: the role of the RNA hairpin*. Proc Natl Acad Sci U S A, 1995. **92**(19): p. 8793-7.
5. Yager, T.D. and P.H. von Hippel, *A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli*. Biochemistry, 1991. **30**: p. 1097-1118.
6. Komissarova, N., et al., *Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination*. Mol Cell, 2002. **10**(5): p. 1151-62.
7. Korzheva, N., et al., *A structural model of transcription elongation*. Science, 2000. **289**(5479): p. 619-25.
8. Vassylyev, D., et al., *Structural basis for substrate loading in bacterial RNA polymerase*. Nature 2007. **448**(7150): p. 163-168.
9. Nudler, E., et al., *Transcription processivity: protein-DNA interactions holding together the elongation complex*. Science, 1996. **273**(5272): p. 211-7.
10. Gusarov, I. and E. Nudler, *The mechanism of intrinsic transcription termination*. Mol Cell, 1999. **3**(4): p. 495-504.
11. Yarnell, W.S. and J.W. Roberts, *Mechanism of intrinsic transcription termination and antitermination*. Science, 1999. **284**(5414): p. 611-5.
12. Larson, M.H., et al., *A pause sequence enriched at translation start sites drives transcription dynamics in vivo*. Science, 2014. **344**(6187): p. 1042-7.
13. Touloukhonov, I., et al., *A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing*. Mol Cell, 2007. **27**(3): p. 406-19.
14. Ray-Soni, A., R.A. Mooney, and R. Landick, *Trigger loop dynamics can explain stimulation of intrinsic termination by bacterial RNA polymerase without terminator hairpin contact*. Proc Natl Acad Sci U S A, 2017. **114**(44): p. E9233-E9242.
15. Martin, F.H. and I. Tinoco, Jr., *DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription*. Nucleic Acids Res, 1980. **8**(10): p. 2295-9.
16. Larson, M.H., et al., *Applied force reveals mechanistic and energetic details of transcription termination*. Cell, 2008. **132**(6): p. 971-82.
17. Kashlev, M. and N. Komissarova, *Transcription termination: primary intermediates and secondary adducts*. J Biol Chem, 2002. **277**(17): p. 14501-8.

18. Santangelo, T.J. and J.W. Roberts, *Forward translocation is the natural pathway of RNA release at an intrinsic terminator*. Molecular Cell, 2004. **14**(1): p. 117-126.
19. Epshtein, V., et al., *An allosteric path to transcription termination*. Mol Cell, 2007. **28**(6): p. 991-1001.
20. Epshtein, V., et al., *An allosteric mechanism of Rho-dependent transcription termination*. Nature, 2010. **463**(7278): p. 245-9.
21. Weixlbaumer, A., et al., *Structural basis of transcriptional pausing in bacteria*. Cell, 2013. **152**(3): p. 431-41.
22. Chakraborty, A., et al., *Opening and closing of the bacterial RNA polymerase clamp*. Science, 2012. **337**(6094): p. 591-5.
23. Gnatt, A.L., et al., *Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution*. Science, 2001. **292**(5523): p. 1876-82.
24. Sekine, S., et al., *The ratcheted and ratchetable structural states of RNA polymerase underlie multiple transcriptional functions*. Mol Cell, 2015. **57**(3): p. 408-21.
25. Tagami, S., et al., *Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein*. Nature, 2010. **468**(7326): p. 978-82.
26. Vassylyev, D.G., et al., *Structural basis for transcription elongation by bacterial RNA polymerase*. Nature, 2007. **448**(7150): p. 157-62.
27. Kang, J.Y., et al., *Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex*. Elife, 2017. **6**.
28. Kang, J.Y., et al., *RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing*. Mol Cell, 2018. **69**(5): p. 802-815 e1.
29. Weilbaecher, R., et al., *Termination-altering amino acid substitutions in the beta' subunit of Escherichia coli RNA polymerase identify regions involved in RNA chain elongation*. Genes Dev, 1994. **8**(23): p. 2913-27.
30. Cheeran, A., et al., *Escherichia coli RNA polymerase mutations located near the upstream edge of an RNA:DNA hybrid and the beginning of the RNA-exit channel are defective for transcription antitermination by the N protein from lambdoid phage H-19B*. J Mol Biol, 2005. **352**(1): p. 28-43.
31. Hein, P.P., et al., *RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement*. Nat Struct Mol Biol, 2014. **21**(9): p. 794-802.
32. Touloukhonov, I. and R. Landick, *The role of the lid element in transcription by E. coli RNA polymerase*. J Mol Biol, 2006. **361**(4): p. 644-58.

33. Guo, X., et al., *Structural Basis for NusA Stabilized Transcriptional Pausing*. Mol Cell, 2018. **69**(5): p. 816-827 e4.
34. Palangat, M., et al., *Efficient reconstitution of transcription elongation complexes for single-molecule studies of eukaryotic RNA polymerase II*. Transcription, 2012. **3**(3): p. 146-53.
35. Mishanina, T.V., et al., *Trigger loop of RNA polymerase is a positional, not acid-base, catalyst for both transcription and proofreading*. Proc Natl Acad Sci U S A, 2017. **114**(26): p. E5103-E5112.
36. Opalka, N., et al., *Complete structural model of Escherichia coli RNA polymerase from a hybrid approach*. PLoS Biol, 2010. **8**(9).
37. Cramer, P., D. Bushnell, and R. Kornberg, *Structural basis of transcription: RNA polymerase II at 2.8 Å resolution*. Science, 2001. **292**(5523): p. 1863-76.
38. Vassylyev, D.G., et al., *Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution*. Nature, 2002. **417**(6890): p. 712-9.
39. Felklistov, A., et al., *RNA polymerase motions during promoter melting*. Science, 2017. **356**: p. 863-866.
40. Bayro, M.J., et al., *Structure of antibacterial peptide microcin J25: a 21-residue lariat protoknot*. J Am Chem Soc, 2003. **125**(41): p. 12382-12383.
41. Mukhopadhyay, J., et al., *Antibacterial peptide microcin J25 inhibits transcription by binding within and obstructing the RNA polymerase secondary channel*. Mol Cell, 2004. **14**(6): p. 739-51.
42. Yuzenkova, J., et al., *Mutations of bacterial RNA polymerase leading to resistance to microcin j25*. J Biol Chem, 2002. **277**(52): p. 50867-75.
43. Delgado, M.A., et al., *Escherichia coli RNA polymerase is the target of the cyclopeptide antibiotic microcin J25*. J Bacteriol, 2001. **183**(15): p. 4543-50.
44. Nayak, D., et al., *Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase*. Mol Cell, 2013. **50**(6): p. 882-93.
45. Windgassen, T.A., et al., *Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase*. Nucleic Acids Res, 2014. **42**(20): p. 12707-21.
46. Wang, D., et al., *Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis*. Cell, 2006. **127**(5): p. 941-54.
47. Kang, J.Y., et al., *Structural Basis for Transcript Elongation Control by NusG Family Universal Regulators*. Cell, 2018. **173**(7): p. 1650-1662.e14.

48. Shetty, A., et al., *Spt5 Plays Vital Roles in the Control of Sense and Antisense Transcription Elongation*. Mol Cell, 2017. **66**(1): p. 77-88 e5.
49. Yamaguchi, Y., et al., *Hepatitis delta antigen binds to the clamp of RNA polymerase II and affects transcriptional fidelity*. Genes Cells, 2007. **12**(7): p. 863-75.
50. Singh, N., et al., *SuhB Associates with Nus Factors To Facilitate 30S Ribosome Biogenesis in Escherichia coli*. MBio, 2016. **7**(2): p. e00114.
51. Arnvig, K.B., et al., *Evolutionary comparison of ribosomal operon antitermination function*. J Bacteriol, 2008. **190**(21): p. 7251-7.
52. Peters, J.M., A.D. Vangeloff, and R. Landick, *Bacterial transcription terminators: the RNA 3'-end chronicles*. J Mol Biol, 2011. **412**(5): p. 793-813.
53. Kireeva, M., et al., *RNA-DNA and DNA-DNA base-pairing at the upstream edge of the transcription bubble regulate translocation of RNA polymerase and transcription rate*. Nucleic Acids Res, 2018. **46**(11): p. 5764-5775.
54. Ryder, A.M. and J.W. Roberts, *Role of the Non-template Strand of the Elongation Bubble in Intrinsic Transcription Termination*. Journal of Molecular Biology, 2003. **334**(2): p. 205-213.
55. Harden, T.T., et al., *Bacterial RNA polymerase can retain sigma70 throughout transcription*. Proc Natl Acad Sci U S A, 2016. **113**(3): p. 602-7.
56. Deighan, P., et al., *Initial transcribed region sequences influence the composition and functional properties of the bacterial elongation complex*. Genes Dev, 2011. **25**(1): p. 77-88.
57. Bar-Nahum, G. and E. Nudler, *Isolation and Characterization of σ 70-Retaining Transcription Elongation Complexes from Escherichia coli*. Cell, 2001. **106**: p. 443-451.
58. Hampsey, M., *A new direction for gene loops*. Science, 2012. **338**(6107): p. 624-5.
59. Gribskov, M. and R.R. Burgess, *Overexpression and purification of the sigma subunit of Eschevichia coli RNA polymerase*. Gene, 1985. **26**(1983): p. 109-118.
60. Ederth, J., et al., *The downstream DNA jaw of bacterial RNA polymerase facilitates both transcriptional initiation and pausing*. J Biol Chem, 2002. **277**(40): p. 37456-63.
61. Pan, S.J., W.L. Cheung, and A.J. Link, *Engineered gene clusters for the production of the antimicrobial peptide microcin J25*. Protein Expr Purif, 2010. **71**(2): p. 200-6.
62. Borowiec, J.A. and J.D. Gralla, *Supercoiling response of the lac ps promoter in vitro*. J Mol Biol, 1985. **184**(4): p. 597-598.

63. Reynolds, R., R.M. Bermudez-Cruz, and M.J. Chamberlin, *Parameters affecting transcription termination by Escherichia coli RNA polymerase. I. Analysis of 13 rho-independent terminators*. J Mol Biol, 1992. **224**(1): p. 31-51.
64. von Hippel, P.H. and T.D. Yager, *Transcript elongation and termination are competitive kinetic processes*. Proc Natl Acad Sci U S A, 1991. **88**(6): p. 2307-11.

Chapter 3

Mechanism for the regulated control of Rho-dependent termination by NusG

Adapted from:

Michael R. Lawson, Wen Ma , Michael J. Bellecourt, Irina Artsimovitch, Andreas Martin, Robert Landick, Klaus Schulten, James M. Berger. (2018). Mechanism for the regulated control of bacterial transcription termination by a universal adapter protein. *Mol. Cell*. Article in press.

Chapter contributions:

Chapter consists of the subset of data from the manuscript that M.J.B. contributed to. M.R.L. & M.J.B. conceived the initial experimental hypothesis. M.R.L., M.J.B., and J.M.B. wrote the text. M.R.L. performed the X-ray crystallography and M.J.B. & I.A. performed *in vitro* termination. M.R.L. and M.J.B constructed the figures from the manuscript, and M.J.B. generated new panels for the dissertation chapter. I.A. and R.L. revised the text.

3.1 Abstract

NusG/Spt5 proteins are the only transcription factors utilized by all cellular organisms. In enterobacteria, NusG antagonizes the transcription termination activity of Rho, a hexameric helicase, during the transcription of ribosomal and actively-translated mRNAs. Paradoxically, NusG helps Rho act on untranslated transcripts, including non-canonical antisense RNAs and those arising from translational stress; how NusG fulfills these disparate functions is unknown. Here, we demonstrate that NusG activates Rho by assisting helicase isomerization from an open-ring, RNA-loading state to a closed-ring, catalytically-active translocase. A crystal structure of closed-ring Rho in complex with NusG reveals the physical basis for this activation, and further explains how Rho is excluded from translationally-competent RNAs. This study demonstrates how a universally conserved transcription factor acts to modulate the activity of a ring-shaped ATPase motor, and establishes how the innate sequence-bias of a termination factor can be modulated to silence pervasive, aberrant transcription.

3.2 Introduction

The proper regulation of gene expression relies on the tight control of all stages of transcription, from initiation to termination. The bacterial Rho factor, an archetype of factor-dependent termination, is an ATP-dependent, RecA-family hexameric helicase required for terminating approximately 20% of steady-state mRNA transcription events in *Escherichia coli* [1]. Rho-dependent termination is typically initiated after the transcription and exposure of an unstructured and cytosine-rich Rho utilization (*rut*) RNA sequence. Rho uses an oligonucleotide/oligosaccharide binding (OB) fold located in the N-terminus of each of its six subunits (its so-called primary RNA binding site) to recognize dipyrimidine motifs within a *rut*

site (Fig. 3.1A, left) [2-4]. Prior to *rut* recognition, Rho exists in an open-ring configuration [5, 6], and *rut* binding helps guide the nascent RNA transcript into the interior of the Rho ring (known as the secondary RNA binding site [7]). Once RNA is bound, ATP is used to drive the formation a closed-ring form, trapping RNA in the central pore of the hexamer (Fig. 3.1A, right panel) [8, 9]. ATP turnover then promotes the 5'→3' translocation of Rho along RNA until it encounters RNA polymerase (RNAP) and terminates transcription [10-12].

Although the ability of Rho to terminate the synthesis of transcripts with canonical *rut* sites is well established, this mechanism fails to account for how Rho terminates the synthesis of transcripts that lack strong *rut* sequences. For example, Rho is responsible for resolving R-loops [13], and assists with the repression of horizontally transferred genes [14]. Rho also acts at many cryptic intragenic sites under stress conditions such as amino acid starvation [15] and also suppresses antisense [16]. Curiously, many non-canonical transcripts targeted by Rho are well established responses to translational stress [17].

Rho's ability to function in an apparently sequence-independent manner is strongly influenced by the transcription factor NusG [16, 18]. The NusG and Spt5-family of proteins (which are bacterial and archaeal or eukaryotic homologs, respectively; Fig. 3.1B) are the only transcription elongation factors known to be conserved across all organisms [19]. The N-terminal domains of NusG (herein referenced as NusG-NTD) and Spt5 factors bind to RNAP and reduce transcriptional pausing [20]. All NusG/Spt5-family proteins also contain at least one flexibly attached, C-terminal Kyrpides-Ouzounis-Woese ("KOW," herein referred to as "NusG-CTD") domain [20], which serves as a recruitment platform for accessory factors, including Rho [21]. While existing evidence suggests that NusG accelerates the rate of at least one sub-step of Rho's catalytic cycle [22-25], its ability to compensate for insufficient cytosine content in transcripts is

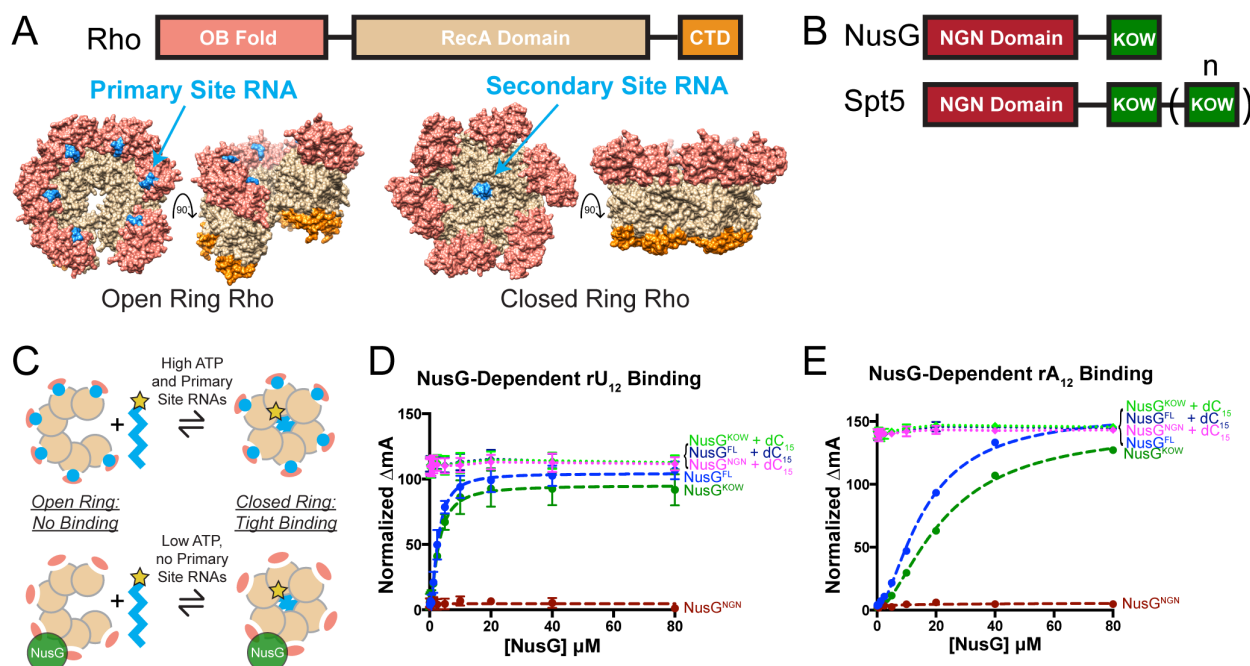


Figure 3.1. NusG promotes Rho ring closure. **(A)** Domain organization of Rho overlaid on open-ring and closed-ring structures shown in surface representation (PDB 1pv4 and 3ice, respectively; [5, 8]). **(B)** Domain organization of NusG/Spt5-like proteins. **(C)** Schematic of ring closure assay used to determine the impact of association of NusG on Rho ring closure. The fluorescently-labeled RNA (blue and yellow) binds preferentially to Rho's secondary site. In the presence of saturating nucleotide concentrations, this RNA binding event triggers ring closure [9]; NusG (green) promotes ring closure at concentrations of nucleotide insufficient to produce this change. **(D–E)** Full-length NusG and the NusG-CTD (but not NusG-NTD domain) favor rU₁₂ (D) and rA₁₂ (E) binding to Rho. The effect of NusG on Rho ring closure is masked by the inclusion of primary site ligands, which strongly promote secondary-site RNA binding and ring closure on their own [26, 27]. The FA data shown represent averages from three experiments; error bars indicate the standard deviations of these averages.

not understood.

In considering the stimulatory effect of NusG on Rho [16], it is notable that NusG was initially identified as an antitermination factor that opposes Rho action [28]. This antagonism arises in part due to an ability of NusG to directly potentiate transcription elongation by RNAP [29]. However, NusG also interacts with the dual-purpose transcription/translation factor NusE (also known as ribosomal protein S10); NusE not only prevents Rho from binding to NusG, but also serves as an anchor point for the binding of coupled ribosomes and additional antitermination factors that further deter Rho function [21, 30]. NusG can be displaced from transcription complexes by other factors, such as RfaH, which also prevent polymerase pausing and aid ribosome recruitment. Intriguingly, RfaH is a NusG paralog, but despite high structural similarity, it does not associate with Rho.

A cohesive understanding of how the regulated exchange of distinct factors on bacterial RNAP elicits the appropriate termination of transcription has been hampered by a lack of knowledge about how NusG positively and negatively regulates Rho. Here, we show that NusG, specifically the NusG-CTD domain, helps Rho encircle RNA to promote transcription termination. Structural, computational, and biochemical analyses of Rho and NusG reveal that NusG manipulates an allosteric network to promote the closure and activation of Rho, and explain how antitermination factors, NusG paralogs, and coupled ribosomes prevent Rho from associating with transcription complexes. Collectively, our work demonstrates how a universally conserved transcription factor can repurpose an otherwise sequence-specific transcription terminator to halt the synthesis of unnecessary transcripts for aiding transcriptional homeostasis and cellular response to stresses.

3.3 Results

3.3.1 Structure of a Rho–NusG complex

To better understand how NusG acts on Rho, we set out to determine a structure of the complex in the presence of RNA (rU₁₂) and a non-hydrolyzable nucleotide analog (ADP•BeF₃) using X-ray crystallography. Molecular replacement was used to phase single-wavelength diffraction data collected from crystals of the complex, which belong to the space group P1 and contain two equivalent Rho hexamers per asymmetric unit. Density was evident for RNA in the secondary site of each hexamer and for one ADP•BeF₃ moiety associated with each Rho protomer (Fig. 3.2 & 3.3A–C). Following building and refinement, the model converged to an R_{work} and R_{free} of 28.7% and 30.0%, respectively (Table 3.1).

In the 3.7 Å structure of the complex, six NusG-CTD domains are seen to associate with Rho, which exists in a closed-ring state (Fig. 3.2A). The initial placement of the NusG-CTD domain relative to Rho was achieved by a 6D search of SA-composite omit maps and validated by single wavelength anomalous dispersion (SAD) experiments using selenomethionine-derivatized NusG (Figs. 3.2A & 3.3C–E, Table 3.1). Several lines of evidence indicate that the complex crystallized in an apparently three-fold symmetric, “trimer-of-dimers” configuration arrayed in an alternating pattern of strong and weak NusG sites around each hexamer (Fig. 3.3). A similar pattern of alternately high and low B-factors is also apparent for the Rho subunits and their associated ADP-BeF₃ moieties. Inspection of the Rho–NusG interface shows that NusG engages the last three α -helices of Rho’s motor domain (the Rho-CTD) using two loops that protrude from the NusG-CTD β -barrel (residues 139–144, referred to as “L1,” and amino acids 164–167, termed “L2”; Fig. 3.2A, inset).

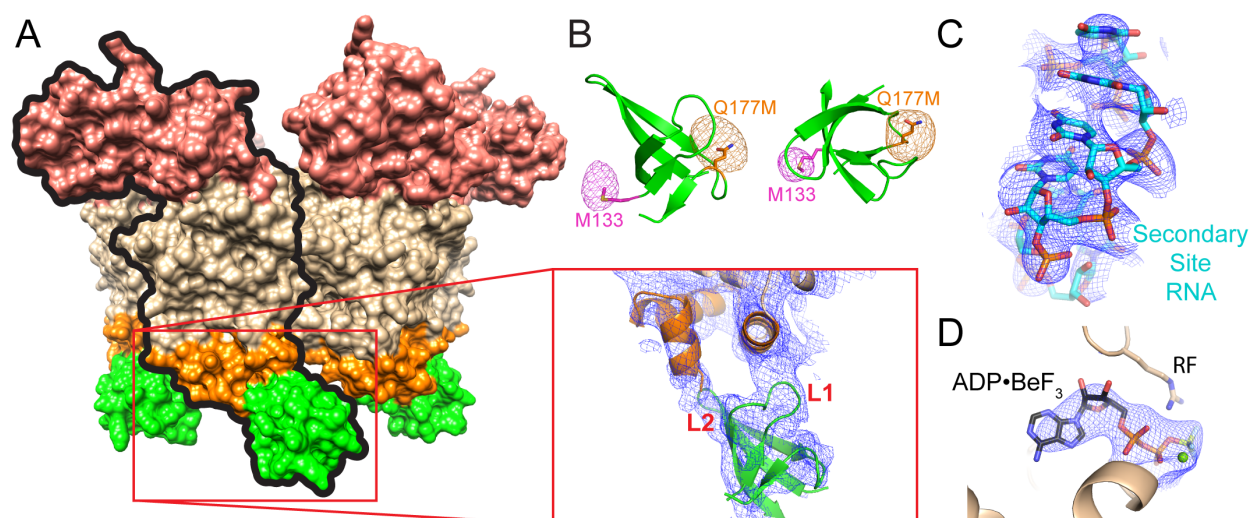


Figure 3.2. Structure of closed-ring Rho in complex with NusG. **(A)** Crystal structure of Rho–NusG-CTD bound by rU₁₂ and ADP-BeF₃ reveals one NusG-CTD and one ADP-BeF₃ moiety associated with every Rho subunit. A single NusG-bound Rho subunit is highlighted by a black outline. (Inset) Representative 2F_o-F_c density of the Rho–NusG-CTD interface. Rho-binding NusG loops are denoted “L1” and “L2”. **(B)** LLG anomalous difference maps obtained from crystals grown with a selenomethionine-substituted mutant of NusG-CTD KOW (Q177M) display two clear peaks, one close the native (M133) site and another at the engineered (Q177M) site. **(C–D)** Representative 2F_o-F_c density of a secondary-site RNA (C) and ADP•BeF₃ (D), shown at 1 and 3 σ contour (respectively). Rho’s arginine finger, a residue required for ATPase activity, is denoted “RF”.

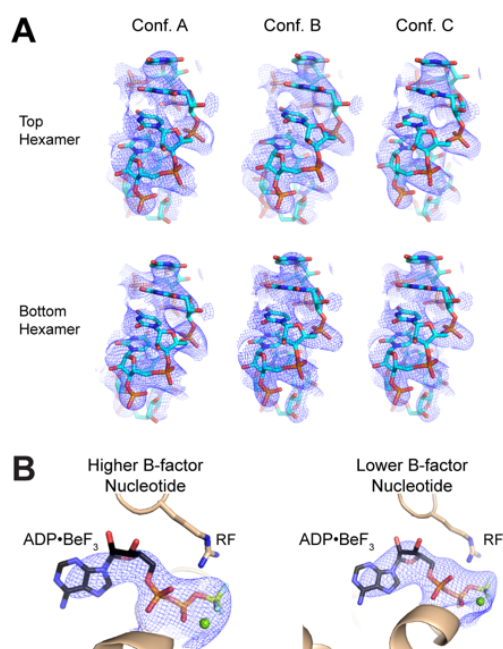


Figure 3.3. Representative electron density for the Rho–NusG, rU12, ADP–BeF₃ structure. **(A)** 2F_o–F_c density around secondary-site RNAs, which were modeled in three alternate positions for each hexamer within the asymmetric unit. **(B)** 2F_o–F_c density (3σ contour) for ADP–BeF₃ moieties from higher (left) and lower (right) B-factor sites. Rho’s arginine finger is denoted as “RF”.

Table 3.1. Data collection and refinement.

	Native	SeMet wild-type NusG	SeMet Q177M NusG
Data collection			
Space group	P1	P1	P1
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	126.35, 126.32, 166.82	126.10, 125.91, 165.64	126.06, 125.97, 166.31
α , β , γ (°)	89.942, 89.899, 60.011	89.732, 89.800, 60.032	90.206, 90.082, 60.109
Wavelength (Å)	1.116	0.980	0.980
Resolution (Å)	40-3.68 (3.79-3.68)	38.9-3.99 (4.13-3.99)	38.97-4.04 (4.17-4.04)
No. of reflections	94,588 (8,258)	144,097 (9,636)	140,670 (13,969)
Completeness (%)	97.8 (97.6)	95.9 (85.8)	96.4 (96.6)
Redundancy	2.2 (2.2)	2.0 (1.9)	2.0 (2.0)
$\langle I \rangle / \langle \sigma I \rangle$	8.9 (1.03)	5.54 (0.63)	8.2 (1.09)
<i>R</i> _{meas}	9.7 (119.9)	13.5 (180.5)	10.3 (115.1)
<i>R</i> _{pim}	6.5 (81.5)	9.5 (118.1)	7.3 (82.6)
<i>CC</i> _{1/2}	99.9 (30.4)	99.8 (22.1)	99.8 (37.5)
Refinement			
Resolution (Å)	40-3.70		
No. reflections	93084		
<i>R</i> _{work} / <i>R</i> _{free} (%)	28.7 / 30.0		
No. atoms	43612		
Protein	43228		
Ligand or ion	384		
B-factors	201.80		
Protein	202.15		
Ligand or ion	162.80		
R.M.S deviations			
Bond lengths (Å)	0.008		
Bond angles (°)	0.772		
Ramachandran Plot			
Favored (%)	98.2		
Allowed (%)	1.8		
Outlier (%)	0.0		

Values in parentheses correspond to the highest resolution bin

3.3.2 NusG promotes early Rho-dependent termination through its action on Rho ring conformation

Molecular dynamic analyses of the Rho–NusG crystal structure implicated a number of residues in NusG-mediated closure of the Rho ring (M.R.L. personal communication, Fig. 3.3A). Three distinct pairs of residues on the NusG-CTD were observed to engage the Rho-CTD throughout the course of the MD simulations. These include: (i) P140 and F141 of NusG, which sit proximal to W381 and I382 of Rho, (ii) F144 and the carbonyl group of D143 of NusG, which interact with I382/K417 and K417/R418 of Rho, respectively, and (iii) I164 and F165 of NusG, which nestle into a hydrophobic pocket composed of I382, I386, F398, M415 and M416 of Rho.

We sought to assess the three identified interaction networks roles during transcription elongation using an *in vitro* Rho-dependent termination assay using alanine mutants of NusG F165, Rho W381, and Rho K417. Transcription was initiated on a DNA template encoding bacteriophage λ_{R1} , an archetypical Rho-dependent terminator (Fig. 3.4B). In the absence of other protein factors, RNAP elongated to the downstream end of this linear template, generating predominantly read-through (RT) transcripts (Fig. 3.4C–D). By contrast, Rho promoted efficient termination at several promoter-distal sites, dramatically reducing RT transcription. The addition of full-length NusG further enhanced termination and stimulated early RNA release at promoter-proximal sites by ~3-fold, whereas the addition of the NusG-NTD domain did not, consistent with prior studies by others [20, 31].

We next tested the effects of selected NusG and Rho mutants on termination. NusG F165A, a mutant that was strongly impaired for binding to Rho and promoting ring closure (Fig. 3.4A), proved unable to drive promoter-proximal termination by wild-type Rho when used at standard concentrations in the assay (150 nM), and produced only a 1.7-fold increase in early

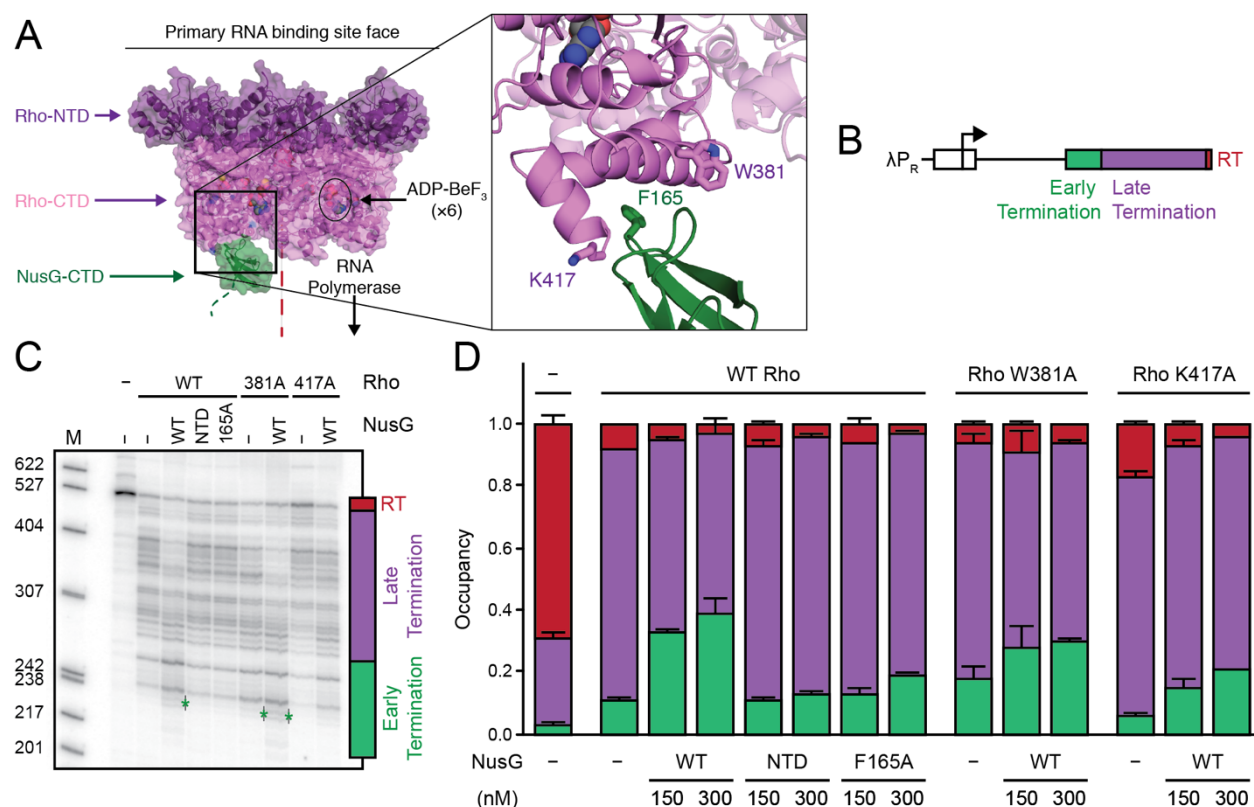


Figure 3.4. *In vitro* termination assays confirm the dependence of NusG on switch amino acids in Rho. **(A)** Location of tested mutations on the Rho–NusG interface. Only one NusG-CTD is shown; the other five in the crystal structure are omitted for clarity. The ATP transition state analog ADP-BeF₃ and Mg²⁺ are shown in spheres, and the path of RNA out of the secondary channel towards RNAP is indicated with a red dashed line. **(B)** A 568-bp DNA template encodes the λP_R terminator downstream of the λP_R promoter and a C-less cassette (white) that allows for the formation of synchronized, radiolabeled elongation complexes halted 26 nucleotides downstream of the transcription start site (arrow). Promoter-proximal (“Early termination”, green) and distal (“Late termination”, purple) regions, as well as the readthrough site (RT, red), are indicated. **(C)** Representative single-round transcription reactions. Elongation complexes halted at position A26 were incubated with Rho and/or NusG (where indicated), restarted by addition of NTPs, and subsequently quenched. Positions of the full-length RT RNA, and both early- and late-termination regions are indicated. Green asterisks mark the positions of the earliest termination sites observed under optimal ring closure conditions. **(D)** Densitometric quantitation of terminated and RT RNA products. ‘Occupancy’ is calculated as a fraction of a given RNA product versus the sum of all RNA products. The data shown represent averages from three experiments; error bars indicate the standard deviations of these averages.

termination when present at double the concentration (300 nM, Fig. 3.4D). The K417A variant of Rho terminated transcription less efficiently than wild-type Rho at both proximal and distal sites and, consistent with its partial response to NusG-stimulated ring closure (M.R.L., personal communication), showed only a slightly reduced response to NusG in early termination (~2.5-fold vs ~3-fold enhancement observed with wild-type Rho (Fig. 3.4D).

Finally, we assessed the contribution of Rho's autoinhibitory network to termination. The W381A substitution disrupted autoinhibition and strongly promoted ring closure, mimicking the effects of NusG (M.R.L., personal communication). Consistently, the W381A Rho mutant exhibited ~1.6-fold increase in early termination in the absence of NusG (Fig. 3.4D). Interestingly, W381A was modestly activated by NusG, suggesting that other elements of Rho's autoinhibitory network (such as the E368/R384 salt bridge) remain partially capable of resisting ring closure in the absence of NusG. Taken together, the *in vitro* transcription termination data support the idea that NusG remodels an autoinhibitory network within Rho to aid ring closure and facilitate termination at sub-optimal sites.

3.4 Discussion

Rho is a textbook exemplar of protein-mediated transcription termination [32-34]. Acting as a hexameric RNA translocase, Rho terminates the synthesis of transcripts downstream from cytosine-rich *rut* sites in bacteria [14, 35]. However, Rho also can function in other regulatory contexts, such as the suppression of antisense transcription [16], the transcriptional repression of horizontally transferred genes [14], and at intragenic termination sites under stress conditions [15]. The diverse roles fulfilled by Rho in addition to normative termination suggest that the enzyme is a global repressor of transcriptional events deemed to be aberrant by a bacterial cell,

particularly when transcripts are devoid of a closely-associated ribosome. These expanded responsibilities necessitate that, in addition to working on classical *rut* elements, Rho must also be able to act in a manner that is largely insensitive to the sequence of the transcript when the translational machinery is absent.

Although it has been shown that NusG can stimulate Rho to terminate transcripts lacking canonical *rut* sequences [16], the mechanistic basis for this activity has not been resolved. We discovered that NusG facilitates the encirclement of Rho around substrate RNAs, even non-ideal purine-rich sequences, in the absence of primary site RNA ligands (Figs. 3.1D–E & 3.5A). This action, which has been suggested by Sen and colleagues to account for the phenotypes of Rho mutants that bypass a requirement for NusG [24], provides a molecular rationale for how NusG promotes efficient Rho-dependent termination irrespective of transcript cytosine content (see Fig. 3.6A for a comparison of the Rho–NusG structure to the model proposed by Sen and colleagues). Our results further explain how NusG helps Rho act at weak *rut* sites *in vitro* [23]. Although a previous observation that NusG does not impact Rho ATPase rate may appear to contradict our findings [22], we discovered that the saturation of Rho’s primary sites—which is typically employed in prior studies by others—can mask the effect of NusG on Rho ring closure (Fig. 3.1D–E). We reconcile these reports by showing a direct impact of NusG on Rho activity *in vitro* in the absence of RNAP under conditions where the primary sites remain unoccupied (Fig. 3.1).

The structure of the Rho–NusG-CTD complex enabled the identification of a network of conserved amino acids in Rho whose relative arrangement can be allosterically remodeled by NusG to foster the formation of a catalytically active, closed-ring state (Fig. 3.6B). This regulatory network straddles an intersubunit interface, suggesting that the binding of a single

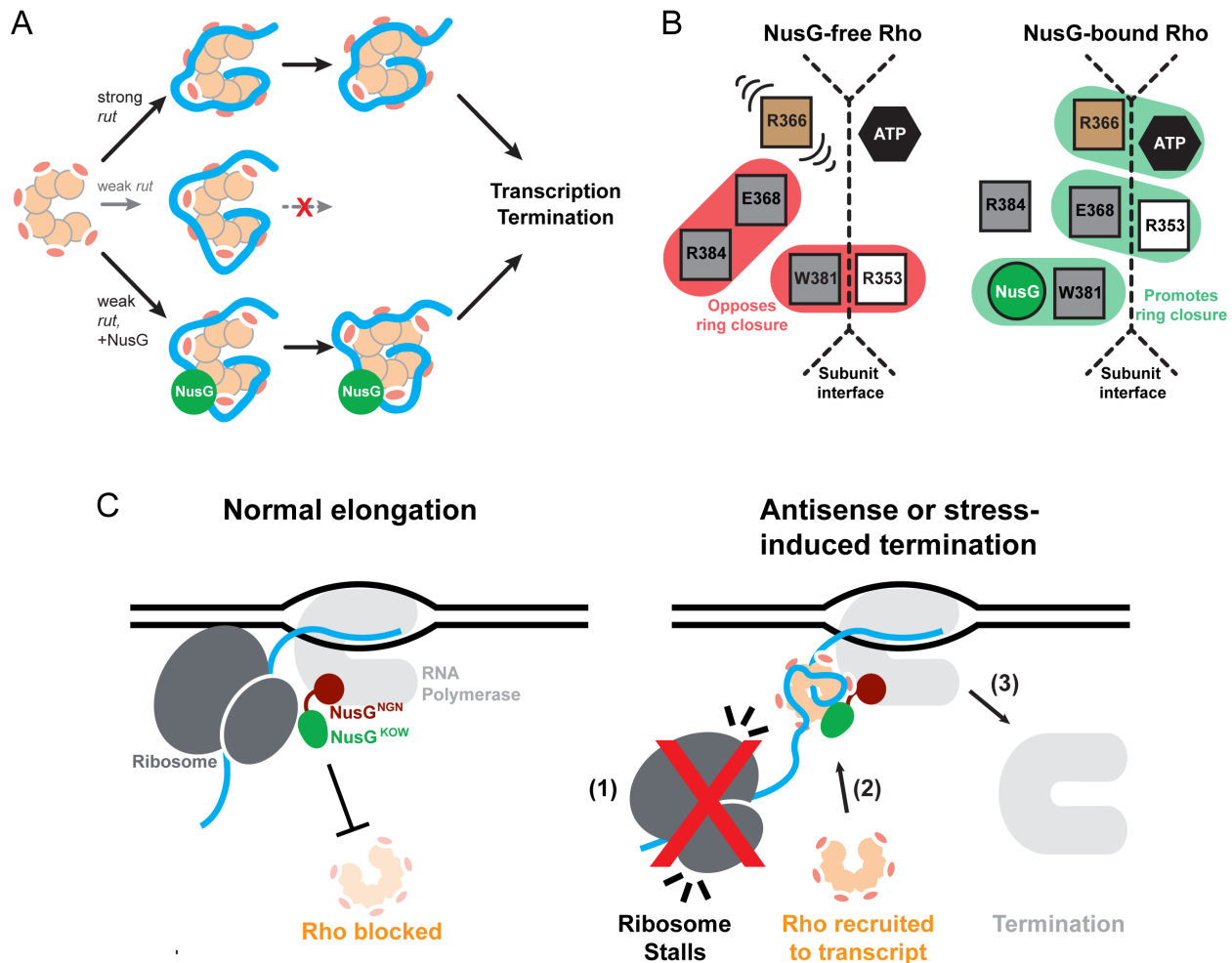


Figure 3.5. Model for NusG-mediated promotion of Rho-dependent transcription termination. **(A)** Schematic showing how Rho ring status is impacted by RNA and NusG. *(Top)* High levels of primary site occupancy by RNA (e.g., due to a preponderance of pyrimidines) promote ring closure and efficient transcription termination. *(Middle)* Purine-rich RNAs allow for binding to only a few primary sites, and hence are unable to trigger ring closure and Rho-dependent termination (denoted as a red “X”) [26, 27]. *(Bottom)* NusG can override low-primary site occupancy to favor Rho activity. **(B)** Schematic depicting how NusG remodels a molecular switch in the Rho-CTD to drive ring closure. In the absence of NusG, two separate sets of bonding elements (shown in red) resist Rho ring closure, and prevent a productive interaction between R366 and ATP. NusG binding to Rho favors the formation of bonds that promote ring closure (green) and stabilize interaction between R366 (the arginine-finger) and ATP. **(C)** Schematic showing how NusG may inhibit and promote Rho-dependent termination. During translation-coupled elongation, Rho action is inhibited either by NusE (S10) bound to the NusG-CTD domain or by

direct RNAP/ribosome interactions shielding the transcript (see also Fig. 3.6). When ribosomes disengage from RNAP or are simply not present (*e.g.*, during translational stress or antisense transcription, denoted by a red “X”), Rho is recruited to the exposed RNA through the NusG-CTD domain, bypassing the need for a *rut*-element.

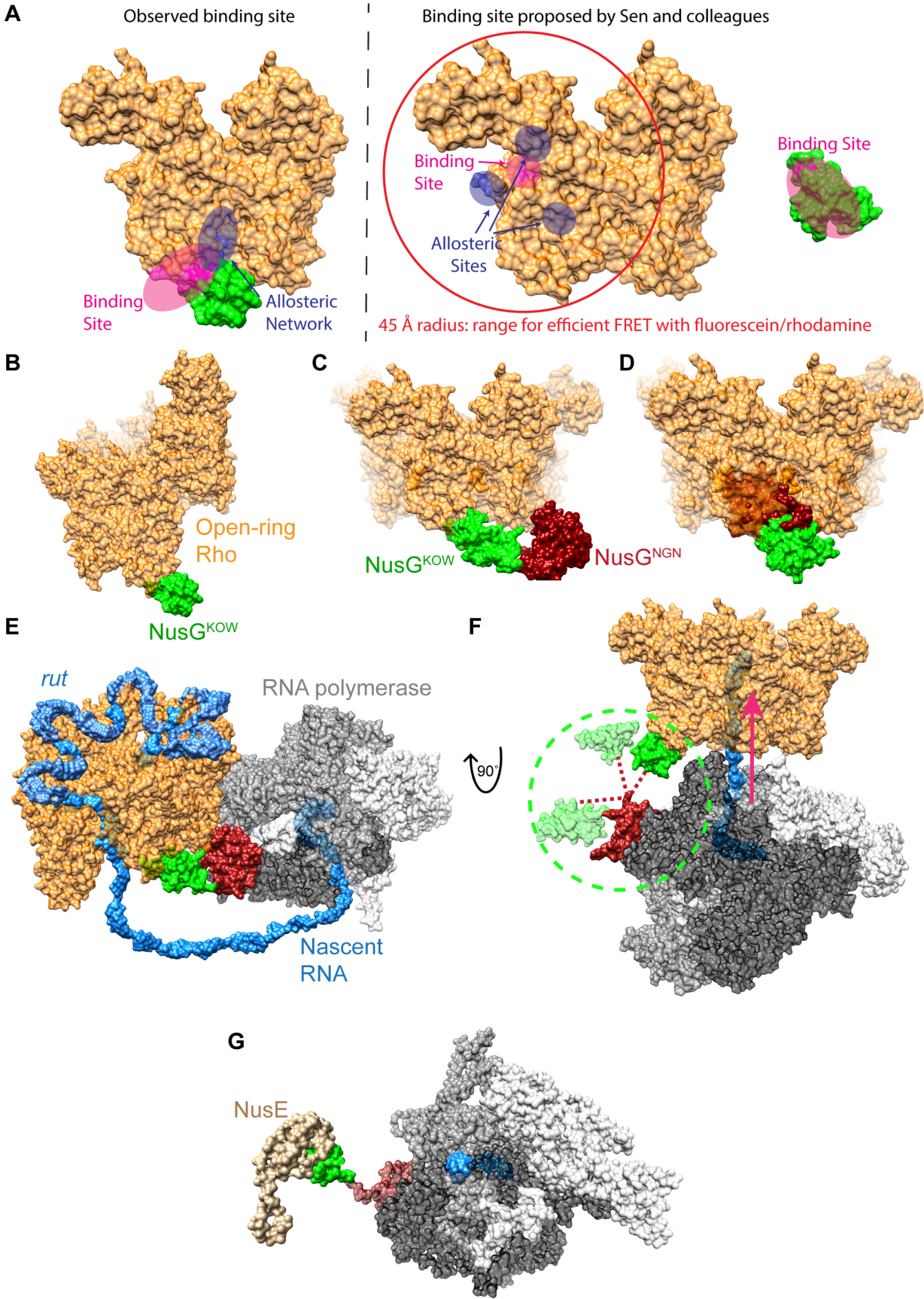


Figure 3.6. Role of NusG in regulating Rho-dependent termination. **(A)** Analysis of the Rho–NusG structure determined here rectifies prior predictions concerning the site of interaction for these two proteins [24, 36]. With one exception (I164A), NusG mutants previously found to interfere with Rho binding and NusG function [36] all result from changes of hydrophobic amino acids to polar residues within the core of the CTD β -barrel. The location and nature of these alterations suggest that they result in non-functional NusG proteins due to improper folding or aggregation. A NusG-binding site on Rho distinct from the interface identified in our work was in turn suggested based by disulfide crosslinking efficiencies [36], but these sites are likewise mostly buried in the Rho protomer; FRET data used for validating the crosslinking turn out to be consistent with the interactions observed in the Rho–NusG structure due to the Förster radius of the fluorophores used in the study. Thus, although prior mutational data did not precisely map the point of Rho–NusG contacts, they helped show that the CTD is necessary for the action of NusG on Rho, and the observations reported here further extend upon this finding to explain how NusG exerts its myriad effects. Interestingly, in follow-up work based on a genetic screen that identified Rho mutants capable of functioning independently of NusG [24], chemical probing was used to reveal that certain alterations can change the conformational state of the helicase. This finding led to the suggestion that NusG might help with Rho ring closure; however, the mutations found in this screen turn out not to map the NusG binding site of Rho but instead are scattered about the RecA domain of the helicase. There are numerous alternative allosteric networks the control Rho’s ring status independently of NusG, such as Q-loop and R-loop communication with the ATPase site [8], and the occupancy of the primary RNA binding site [27, 37]; the mutations identified by Valabhoju *et al.* likely impact one of these networks [24]. **(B)** Model of an open-ring Rho–NusG–CTD complex (PDB 1pv4 [5]) reveals no steric clashes that would prevent NusG from binding to this form of Rho. **(C)** Model of closed-ring Rho in complex an elongated NusG state (PDB 1npp [38]) shows no steric clashes that might impede the association of the two factors (a “mini” domain in *Aquifex aeolicus* NusG, residues 63–128, constitutes an element not present in *E. coli* and is not shown). **(D)** Model of closed-ring Rho with an autoinhibited NusG state (PDB ID 2lq8 [39]) reveals clear clashes between the Rho-CTD and NusG-NTD, suggesting that this form of NusG will not bind to the helicase. **(E)** Model of an open-ring Rho–NusG complex bound to an elongation complex. RNA is shown in blue and RNA polymerase (PDB ID 6alf [40]) in light gray, dark gray (β subunit), and slate (β'). The interaction between the NusG-NTD domain and RNA polymerase is based on superposing an Spt5/RNA polymerase clamp domain structure (PDB ID 3qqc, [41]) with an RNA polymerase elongation complex (PDB ID 6alf [40]) and with an elongated NusG state (PDB ID 1npp [38]). **(F)** Model of a closed-ring Rho–NusG complex bound to an elongation complex. Different NusG-CTD positions observed in various structures with respect to NusG-NTD (PDB IDs 1npp and 5ms0 [30, 38]) are shown in light green, with connecting linkers depicted as red dotted lines. RNA polymerase, a nascent RNA, and NusG-NTD are placed as described for (D). Closed-ring Rho in complex with NusG can be oriented within a spherical volume of space allowed by the flexible linker connecting NusG NTD and CTD (dashed green circle). The direction of RNA synthesis and translocation (by RNA polymerase and Rho, respectively) is shown as a pink arrow. **(G)** The NusE–NusG interaction visualized in λ N antitermination structure (PDB ID 5ms0 [30]) demonstrates that NusE blocks the Rho binding site on NusG-CTD^{KOW}. Other protein subunits and nucleic acids in the complex are hidden for clarity.

NusG may be sufficient to lower the barrier for ring closure when a transcript lacks the pyrimidines needed to saturate Rho's primary binding sites. These findings facilitated the rational design of a hyperactive Rho W381A mutant, which ablates an inhibitory cation- π interaction and is capable of eliciting early transcription termination in the absence of NusG (Fig. 3.4).

To explore whether the Rho–NusG interactions we observe can accommodate the association of Rho with a nascent transcript, we generated a model of a Rho–NusG complex bound to a transcription elongation complex (Fig. 3.6B–F). The NusG NTD and CTD are connected by a flexible linker region [20], which would allow Rho to access a range of orientations relative to RNAP. The NusG-CTD domain has previously been proposed to sample a sphere of allowed positions with respect to the RNAP clamp region [41], the radius of which is dictated by the length of the linker connecting the KOW and NGN domains (Fig. 3.6F). The association of Rho with a transcription elongation complex through the NusG-CTD domain would presumably allow Rho (in either an open- or closed-ring configuration) to access a similar spherical region around the NusG-NTD domain, but with a substantially larger radius owing to the size of the Rho hexamer. Interestingly, within this larger sphere of movement, Rho can spatially access the region around the RNA exit channel of RNAP. This accessibility may help drive Rho's association with weak *rut* sites by increasing the effective local concentration of the terminator and the nascent transcript (see Fig. 3.5F). Although six NusG-CTD domains were seen to bind to a Rho hexamer in our crystal structure, this stoichiometry is likely due to the nucleotide and RNA-binding conditions used in the crystallization buffer, which favor Rho ring closure, and the elevated NusG concentrations necessary to promote crystal formation. We envision that once a single, RNAP-tethered NusG molecule binds to Rho and allosterically

toggles its associated active site to form a catalytically competent subunit/subunit conformation, the other subunits in the helicase cooperatively respond in turn to form a closed ring.

In thinking about how Rho may be regulated, it also is important to consider the physical coupling between transcriptional and translational machineries. The ribosome has recently been observed to directly contact RNAP [42], forming a super-complex that may feed RNA from the polymerase exit channel directly into the ribosome [42] and sequester the transcript away from Rho (Fig. 3.5C, left). However, the disruption of this ribosome-polymerase interaction (*e.g.*, due to a stop codon or ribosomal stalling) would unmask a nascent RNA to facilitate Rho loading and translocation, with NusG playing a stimulatory role if no optimal *rut* sequence is available (Fig. 3.5C, right). Alternatively, ribosome-RNAP coupling may be mediated through NusE (also known as ribosomal protein S10), during translation initiation and/or throughout protein synthesis. The NusG-binding region of NusE is surface-accessible in the context of the ribosome [43], suggesting that a NusE–NusG-CTD interaction may form a physical bridge between RNAP and a trailing ribosome [21]. In this scenario, the status of the NusG-CTD domain could provide a physical readout of translational fidelity, with the loss of this interaction due to ribosome stalling or absence (such as during antisense transcription) revealing a docking site for Rho recruitment and promoting sequence-independent transcription termination (Fig. 3.5C, right).

If Rho and NusG can collaborate to terminate synthesis of transcripts devoid of a trailing ribosome, why does this not happen to all transcripts that lack coding potential (*e.g.*, rRNAs)? To this end, NusG can paradoxically act as a member of sequence-specific antitermination complexes that prevent, rather than stimulate, the Rho-dependent termination of some transcripts [28]. In a recently-visualized antitermination complex consisting of λ N, NusA, NusB, NusE, NusG, RNA, and RNAP [30], the L1 and L2 loops of the NusG-CTD domain are stably

associated with NusE. Comparison of this structure to our work reveals that this inhibitory action results from direct competition between NusE and Rho for the same binding site on NusG (Fig. 3.6G) and implicates the accessibility of the NusG-CTD domain as a critical determinant of whether NusG will act to promote or prevent Rho-dependent termination. This finding provides a unified model to reconcile the disparate actions that Rho and NusG fulfill in both transcription and termination (Fig. 3.5C).

It is notable that the eukaryotic termination factor Sen1 directs the sequence-specific termination of transcription in a manner that bears many analogies to bacterial Rho [44-47]. This striking case of convergent evolution—the two proteins are unrelated, except for their shared ASCE ATPase fold [45]—suggests there is a general utility in harnessing the activity of a nucleotide-directed RNA translocase in regulating transcription termination. Critically, bacterial NusGs, together with their Spt5 counterparts in eukaryotes, represent the only class of transcription factors that are conserved across all three domains of life. Although the Spt5-NTD and NusG-NTD bind RNAP similarly [48], the Spt5-CTD domains participate in interactions that appear to be even more diverse and complex than their bacterial counterparts [49, 50]. It will be revealing to learn whether other cellular organisms also utilize the CTD to choreograph the association of different transcription termination factors with RNAP.

3.5. Materials and Methods

3.5.1 Cloning and protein purification

Escherichia coli Rho and all Rho mutants were expressed and purified as described previously [8]. Coding sequences for *E. coli* NusG-NTD (residues 1–119), NusG-CTD (125–181), native full-length NusG (1–181), and RfaH-CTD (112–162) were each cloned into a “1C” LIC vector,

behind a N-terminal and TEV-cleavable His₆-MBP tag (UC Berkeley MacroLab). NusG and RfaH mutants were constructed from parent plasmids using Quikchange or Round-the-horn site-directed mutagenesis. All expression plasmids were verified by DNA sequencing. For the expression of NusG and RfaH constructs, overproduction plasmids were transformed into BL21 (DE3) Codon+ RIL *E. coli* cells (Agilent). Cells were grown at 37 °C in 2XYT media to an OD₆₀₀ of 0.7–0.8, whereupon expression was induced with 0.5 mM IPTG and cells were left shaking overnight at 18 °C. Induced cells were harvested by centrifugation, lysed by sonication, resuspended in 50 ml of Ni-A buffer (500 mM KCl, 50 mM Tris-HCl, 30 mM imidazole, 10% (v/v) glycerol, 1 mM PMSF, 1 µg leupeptin/ml, 2 µg pepstatin-A/ml, all adjusted to pH 7.5), and purified by Ni²⁺-affinity chromatography using a 5 ml HisTrap HP column (GE Healthcare). Protein was eluted using Ni-B buffer (500 mM KCl, 50 mM Tris-HCl, 300 mM imidazole, 10% (v/v) glycerol, all adjusted to pH 7.5). Following concentration to 50 mg/mL in an Amicon Ultra 30K MWCO concentrator (Millipore), protein was combined with 5 mg of TEV protease, inserted into a 15 ml Slide-a-lyzer 3.5K MWCO dialysis cassette (Thermo Scientific), and allowed to cleave overnight while dialyzing into Ni-A buffer. Protein was then subjected to a secondary Ni²⁺-affinity chromatography step, and the flow-through was collected, concentrated, and further purified by size exclusion chromatography (16/60 Sephacryl S-200 HR, GE Healthcare) in SEC buffer (500 mM KCl, 50 mM Tris-HCl, 10% (v/v) glycerol, 1 mM DTT, all adjusted to pH 7.5). Peak fractions were pooled, concentrated to 10–30 mg/mL, snap frozen in liquid nitrogen, and stored at –80° C until use. Protein concentration was determined based on A₂₈₀ measurements using the following extinction coefficients: 95580 M⁻¹ cm⁻¹ for Rho hexamer (62580 M⁻¹ cm⁻¹ for W381A Rho); 15470 M⁻¹ cm⁻¹ for native full-length NusG; 13980 M⁻¹ cm⁻¹ for NusG-NTD; 1490 M⁻¹ cm⁻¹ for NusG-CTD and RfaH-CTD. Wild-type core

RNAP and σ^{70} were purified and reconstituted as σ^{70} RNAP holoenzyme as previously described [51].

3.5.2 Protein crystallization and data collection

Closed-ring Rho was assembled as described previously [8], and subsequently mixed with a three-fold molar excess of NusG-CTD per Rho monomer to generate Rho–NusG bound by rU₁₂ and ADP-BeF₃ (final concentrations of 14.2 mg Rho/ml, 50 μ M rU₁₂, 3 mM ADP-BeF₃, and 5.8 mg NusG-CTD/ml). Crystals were grown at 291 K in Qiagen 15-well easyXtal plates by combining 2 μ l of the Rho-NusG-rU₁₂-ADP-BeF₃ mixture with 2 μ l of well solution (95 mM potassium glutamate, 50 mM Tris-HCl, pH 8.0, 1 mM spermine, 0.5% (v/v) PEG 8K). Crystals were visible after an hour and grew to full size within one week. Harvesting of crystals was carried out using Hampton CryoLoops after gradually introducing (through numerous, small pipetting steps) 30% 1,2-butanediol into the drop (as cryoprotectant), and then flash-freezing in liquid nitrogen. X-ray diffraction data were collected at 90K at beamline 8.3.1 of the LBNL Advanced Light Source.

3.5.3 Structure solution and refinement

Diffraction data were processed using the XDS/XSCALE package [52]. There was some initial ambiguity in assigning a space group to the data, as indexing and scaling in space groups P1, P3, and P6 all yielded reasonable merging statistics. This ambiguity was resolved during refinement, where attempts to refine the data in P3 or P6 led to substantially higher R factors relative to P1, leading us to assign a triclinic lattice to the data. Flags for R-free were set using thin resolution shells [53]. Molecular replacement (MR) was conducted using PHENIX with an RNA- and

nucleotide-free monomer from the Rho^{PolyU-P21} structure as the search model [9], which resulted in the placement of 12 Rho protomers arranged as two hexameric rings in the asymmetric unit.

Electron density maps derived from MR phasing displayed clear density for RNA within Rho's central pore and for ADP-BeF₃ in each Rho ATPase active site. Interestingly, individual subunits in the Rho hexamers were not arrayed asymmetrically, as seen previous RNA and ADP-BeF₃ bound structures [8, 9]. Instead, during the course of refinement, it became clear from the distribution of global subunit B-factors (which alternated between higher and lower values), and from the relative orientation of each subunit with respect to its partner, that the individual hexamers were pseudo-three-fold symmetric. This configuration likely accounts for the reasonable merging statistics from the data in the P3 and P6 space groups; however, as a consequence, the RNA bound to the central channel of Rho was not uniquely ordered, due to rotational averaging about central axis of the helicase ring. The rU₁₂ RNA was therefore modeled as per a prior model of Rho bound to RNA and ADP-BeF₃ (PDB ID 5jji, [9]), using three separate, overlapping chains of 1/3 occupancy each per Rho hexamer. Nucleotide was modeled with custom bond length/angle restraints defined for the BeF₃ moiety.

Upon confirmation of the NusG-CTD orientation, the polyalanine model was replaced with a Rosetta energy-minimized version of a published NMR model of the NusG-CTD domain (generously provided by Roland Pache and Tanja Kortemme, UCSF) that contained side chains [20]. The final model of the complex, which contains 12 copies of Rho, 12 copies of NusG-CTD, six U₁₂ RNA chains and 12 ADP-BeF₃ moieties, was refined to an R_{work}/R_{free} of 28.7/30.0%. Ramachandran statistics for the model, as reported by MOLPROBITY [54], show 98.2% of the residues in favored conformations, with 1.8% in allowed conformations and zero outliers.

3.5.4 *In vitro* Rho-dependent transcription termination assays

λ_{R1} DNA template was generated by PCR amplification and purified by PCR cleanup (Qiagen). Halted A26 elongation complexes were formed at 37 °C for 13 min by mixing 50 nM RNAP with 30 nM λ_{R1} DNA template, 100 μ M ApU, 10 μ M ATP and UTP, 2 μ M GTP, and 5 Ci/mmol [α^{32} P] GTP in Rho termination buffer (40 mM Tris-Cl, 50 mM KCl, 5 mM MgCl₂, 0.1 mM DTT, 3% (v/v) glycerol, pH 7.9). WT and variant Rho proteins were added to 20 nM, WT and variant NusGs to 150 or 300 nM, and the reactions incubated for an additional 3 min at 37 °C. Transcription was restarted by addition of a pre-warmed mixture containing all four NTPs and rifampentine to a final concentration of 200 μ M and 50 mg/ml, respectively, and incubated at 37 °C for 6 min. Reactions were quenched by addition of an equal volume of stop buffer [45 mM Tris-borate, 10 M urea, 20 mM EDTA, 0.2% (w/v) xylene cyanol, 0.2% (w/v) bromophenol blue, pH 8.3] and separated by denaturing 6% PAGE (7 M urea, 0.5X TBE). Gels were dried and products were visualized using a Typhoon FLA 9000 PhosphorImaging system (GE Healthcare Life Sciences). Readthrough and termination RNA products were quantitated with ImageQuant™ and Microsoft Excel software.

3.6. References

1. Peters, J.M., et al., *Rho directs widespread termination of intragenic and stable RNA transcription*. Proc Natl Acad Sci U S A, 2009. **106**(36): p. 15406-11.
2. Bogden, C.E., et al., *The structural basis for terminator recognition by the Rho transcription termination factor*. Mol Cell, 1999. **3**(4): p. 487-93.
3. Dolan, J.W., N.F. Marshall, and J.P. Richardson, *Transcription termination factor rho has three distinct structural domains*. J Biol Chem, 1990. **265**(10): p. 5747-54.
4. Morgan, W.D., et al., *RNA sequence and secondary structure requirements for rho-dependent transcription termination*. Nucleic Acids Res, 1985. **13**(10): p. 3739-54.

5. Skordalakes, E. and J.M. Berger, *Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading*. Cell, 2003. **114**(1): p. 135-46.
6. Yu, X., et al., *Three-dimensional reconstruction of transcription termination factor rho: orientation of the N-terminal domain and visualization of an RNA-binding site*. J Mol Biol, 2000. **299**(5): p. 1279-87.
7. Miwa, Y., T. Horiguchi, and K. Shigesada, *Structural and functional dissections of transcription termination factor rho by random mutagenesis*. J Mol Biol, 1995. **254**(5): p. 815-37.
8. Thomsen, N.D. and J.M. Berger, *Running in reverse: the structural basis for translocation polarity in hexameric helicases*. Cell, 2009. **139**(3): p. 523-34.
9. Thomsen, N.D., et al., *Molecular mechanisms of substrate-controlled ring dynamics and substepping in a nucleic acid-dependent hexameric motor*. Proc Natl Acad Sci U S A, 2016. **113**(48): p. E7691-e7700.
10. Park, J.S. and J.W. Roberts, *Role of DNA bubble rewinding in enzymatic transcription termination*. Proc Natl Acad Sci U S A, 2006. **103**(13): p. 4870-5.
11. Schwartz, A., et al., *Transcription termination factor rho can displace streptavidin from biotinylated RNA*. J Biol Chem, 2007. **282**(43): p. 31469-76.
12. Brennan, C.A., A.J. Dombroski, and T. Platt, *Transcription termination factor rho is an RNA-DNA helicase*. Cell, 1987. **48**(6): p. 945-52.
13. Leela, J.K., et al., *Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in Escherichia coli*. Proc Natl Acad Sci U S A, 2013. **110**(1): p. 258-63.
14. Mitra, P., et al., *Rho Protein: Roles and Mechanisms*. Annu Rev Microbiol, 2017. **71**: p. 687-709.
15. Ruteshouser, E.C. and J.P. Richardson, *Identification and characterization of transcription termination sites in the Escherichia coli lacZ gene*. J Mol Biol, 1989. **208**(1): p. 23-43.
16. Peters, J.M., et al., *Rho and NusG suppress pervasive antisense transcription in Escherichia coli*. Genes Dev, 2012. **26**(23): p. 2621-33.
17. Richardson, J.P., *Preventing the synthesis of unused transcripts by Rho factor*. Cell, 1991. **64**(6): p. 1047-9.
18. Sullivan, S.L. and M.E. Gottesman, *Requirement for E. Coli NusG protein in factor-dependent transcription termination*. Cell, 1992. **68**(5): p. 989-994.

19. Tomar, S.K. and I. Artsimovitch, *NusG-Spt5 proteins-Universal tools for transcription modification and communication*. Chem Rev, 2013. **113**(11): p. 8604-19.
20. Mooney, R.A., et al., *Two structurally independent domains of E. coli NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators*. J Mol Biol, 2009. **391**(2): p. 341-58.
21. Burmann, B.M., et al., *A NusE:NusG complex links transcription and translation*. Science, 2010. **328**(5977): p. 501-4.
22. Nehrke, K.W., F. Zalatan, and T. Platt, *NusG alters Rho-dependent termination of transcription in vitro independent of kinetic coupling*. Gene Expr, 1993. **3**(2): p. 119-33.
23. Burns, C.M. and J.P. Richardson, *NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator*. Proc Natl Acad Sci U S A, 1995. **92**(11): p. 4738-42.
24. Valabhoju, V., S. Agrawal, and R. Sen, *Molecular Basis of NusG-mediated Regulation of Rho-dependent Transcription Termination in Bacteria*. J Biol Chem, 2016. **291**(43): p. 22386-22403.
25. Li, J., S.W. Mason, and J. Greenblatt, *Elongation factor NusG interacts with termination factor rho to regulate termination and antitermination of transcription*. Genes Dev, 1993. **7**(1): p. 161-72.
26. Richardson, J.P., *Activation of Rho protein ATPase requires simultaneous interaction at two kinds of nucleic acid-binding sites*. J Biol Chem, 1982. **257**(10): p. 5760-5766.
27. Lawson, M.R., K. Dyer, and J.M. Berger, *Ligand-induced and small-molecule control of substrate loading in a hexameric helicase*. Proc Natl Acad Sci U S A, 2016. **113**(48): p. 13714-13719.
28. Downing, W.L., et al., *Sequence and transcriptional pattern of the essential Escherichia coli secE-nusG operon*. J Bacteriol, 1990. **172**(3): p. 1621-1627.
29. Sevostyanova, A., et al., *The beta subunit gate loop is required for RNA polymerase modification by RfaH and NusG*. Mol Cell, 2011. **43**(2): p. 253-62.
30. Said, N., et al., *Structural basis for lambdaN-dependent processive transcription antitermination*. Nat Microbiol, 2017. **2**: p. 17062.
31. Pasmán, Z. and P.H. von Hippel, *Regulation of Rho-dependent transcription termination by NusG is specific to the Escherichia coli elongation complex*. Biochemistry, 2000. **39**(18): p. 5573-85.
32. Krebs, J.E., E.S. Goldstein, and S.T. Kilpatrick, *Lewin's GENES XII*. 2017: Jones & Bartlett Learning.

33. Craig, N., et al., *Molecular Biology: Principles of Genome Function*. 2014: OUP Oxford.
34. Watson, J.D., et al., *Molecular Biology of the Gene*. 2013: Pearson Education.
35. Chen, C.Y. and J.P. Richardson, *Sequence elements essential for rho-dependent transcription termination at lambda tR1*. J Biol Chem, 1987. **262**(23): p. 11292-9.
36. Chalissery, J., et al., *Interaction surface of the transcription terminator Rho required to form a complex with the C-terminal domain of the antiterminator NusG*. J Mol Biol, 2011. **405**(1): p. 49-64.
37. Hu, K. and I. Artsimovitch, *A screen for rfaH suppressors reveals a key role for a connector region of termination factor Rho*. MBio, 2017. **8**(3): p. e00753-17.
38. Knowlton, J.R., et al., *A spring-loaded state of NusG in its functional cycle is suggested by X-ray crystallography and supported by site-directed mutants*. Biochemistry, 2003. **42**(8): p. 2275-81.
39. Drogemuller, J., et al., *An autoinhibited state in the structure of Thermotoga maritima NusG*. Structure, 2013. **21**(3): p. 365-75.
40. Kang, J.Y., et al., *Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex*. Elife, 2017. **6**.
41. Martinez-Rucobo, F.W., et al., *Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity*. EMBO J, 2011. **30**(7): p. 1302-10.
42. Kohler, R., et al., *Architecture of a transcribing-translating expressome*. Science, 2017. **356**(6334): p. 194-197.
43. Saxena, S., et al., *Escherichia coli transcription factor NusG binds to 70S ribosomes*. Mol Microbiol, 2018.
44. Porrua, O. and D. Libri, *A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast*. Nat Struct Mol Biol, 2013. **20**(7): p. 884-91.
45. Leonaite, B., et al., *Sen1 has unique structural features grafted on the architecture of the Upf1-like helicase family*. Embo j, 2017. **36**(11): p. 1590-1604.
46. Han, Z. and O. Porrua, *Helicases as transcription termination factors: Different solutions for a common problem*. Transcription, 2017: p. 1-7.
47. Martin-Tumas, S. and D.A. Brow, *Saccharomyces cerevisiae Sen1 Helicase Domain Exhibits 5'- to 3'-Helicase Activity with a Preference for Translocation on DNA Rather than RNA*. J Biol Chem, 2015. **290**(38): p. 22880-9.
48. Kang, J.Y., et al., *Structural Basis for Transcript Elongation Control by NusG Family Universal Regulators*. Cell, 2018. **173**(7): p. 1650-1662.e14.

49. Ehara, H., et al., *Structure of the complete elongation complex of RNA polymerase II with basal factors*. Science, 2017. **357**(6354): p. 921-924.
50. Bernecky, C., J.M. Plitzko, and P. Cramer, *Structure of a transcribing RNA polymerase II-DSIF complex reveals a multidentate DNA-RNA clamp*. Nat Struct Mol Biol, 2017. **24**(10): p. 809-815.
51. Svetlov, V. and I. Artsimovitch, *Purification of bacterial RNA polymerase: tools and protocols*. Methods Mol Biol, 2015. **1276**: p. 13-29.
52. Kabsch, W., *XDS*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 125-32.
53. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.
54. Davis, I.W., et al., *MolProbity: all-atom contacts and structure validation for proteins and nucleic acids*. Nucleic Acids Res, 2007. **35**(Web Server issue): p. W375-83.

Chapter 4

Dissertation dissociation: conclusions and future directions

Chapter contributions:

M.J.B. wrote the text.

4.1. Conclusions

The overarching goal of my dissertation research was to explore how conformation changes in the elongation complex (EC) and associated elongation and termination factors mediate the two processes of programmed bacterial transcription termination. In the nearly fifty years since the discovery and reporting of Rho [1], great strides have been made in determining the mechanisms that drive both termination mechanisms. The thermodynamic contribution of specific sequences within an intrinsic terminator are well-understood [2], and within the last decade we have successfully probed how the terminator hairpin (T_{hp}) drives subtle movements in the nucleic-acid (NA) scaffold to promote destabilization of the EC [3, 4] and how the U-tract drives pausing due through elemental pausing [5-7]. Likewise, the contributions of NA sequence and structure during Rho-dependent termination have been broadly understood for several years. The role of transcriptional pausing in the kinetics of Rho-dependent termination have long been appreciated [8], and structural and biochemical data explains Rho's preference for pyrimidines in both its primary and secondary binding sites [9-11]. These strides in understanding how NA sequences and structures contribute to termination have not, however, been matched by characterizing the roles of protein conformation changes in promoting the process during termination, despite recent advances in understanding these movements throughout elongation [12-14].

During these studies we showed how movements in the clamp affect the intrinsic termination efficiencies at three classes of terminators. By utilizing a cysteine-pair ligation scaffold *in vitro* transcription assay, we demonstrated that inducing conformational biases in the clamp had dramatic effects on the elongation rates of the enzymes, and these changes were the primary drivers of the termination decision (Chapter 2). Incredibly, in most cases constraining the clamp had no significant effects on the termination commitment rates, casting doubt on the

longstanding hypothesis that disruption of RNAP–NA contacts through clamp displacement should facilitate EC inactivation at an intrinsic terminator. However, analyses of crystal and cryo-EM structures of the EC support the notion that inactivation through hybrid shortening should not require large-scale motions of the clamp [15, 16]. A notable exception to this trend occurred on the imperfect t_{500} terminator, where displacement of the clamp both suppressed the elongation rate but also enhanced the commitment rate. At t_{500} , the hybrid is typically shortened through hypertranslocation, or downstream movements of RNAP without concomitant nucleotide addition [4]. Displacement of the clamp may facilitate commitment at t_{500} either by promoting the movements necessary for hypertranslocation or by enabling hybrid shortening through the shearing mechanism exploited at near-perfect terminators.

While we determined that clamp movements were unnecessary for commitment, we remained interested in the role of the module during dissociation of RNA and DNA from RNAP. By performing dissociation assays we determined that clamp movements do indeed play a role in the intrinsic termination process. The width of the RNAP main cleft during elongation is too narrow to allow reannealing of the transcription bubble into double-stranded DNA [13, 15, 17, 18]. Displacement of the clamp during termination, probably through clamp opening, sufficiently widens the main cleft to enable these movements and subsequent dissociation. On the other hand, movements of RNA out of the exit channel were not predicted to require displacement of the clamp. Consistently, we found that inducing a topological constraint on the clamp did not affect the rate of RNA release from an EC.

The dissociation assay also yielded a curious role of the active-site proximal trigger loop (TL) during termination. A recent study had found that efficient termination commitment required free motion of the TL. These results suggested a multistate-multipath (MS-MP) model

of commitment, where the dynamic, unconstrained TL expanded the conformational diversity of the EC, therefore opening up numerous lower energy barriers to commitment [7]. Thermal flux in proteins have previously demonstrated that small conformation shifts can significantly affect the relative height of an activation barrier [19-21], and so the MS-MP model is simply the natural extension of the original thermodynamic models for termination [2]. We provided additional evidence that constraining the TL in an unfolded conformation and reducing its conformational flexibility using an antimicrobial lariat peptide slows both RNA and RNAP. The finding that constraining the TL significantly slowed the rate of RNA dissociation was most stunning, given the large distances between the TL and the 3'-end of the RNA in a proposed model of the inactivated EC. While it was not possible to discern whether these ECs had been trapped prior to or subsequent to commitment (and therefore the effects could have been on either the commitment or dissociation step), the striking result is that preventing TL movements at a terminator stabilizes the EC, in support of the MS-MP model of termination.

Finally, we solved a contradiction in the Rho–NusG literature (Chapter 3). Genome-scale studies of Rho suggested that a class of NusG-dependent terminators acted on RNAs with a low G/C ratio, suggesting a role in binding of Rho to subpar *rut* sequences [22] whereas biochemical data implied NusG acted on a step subsequent to binding [23, 24]. Through collaborative crystallographic and *in vitro* termination assays we identified an amino acid network in Rho that promotes the open-to-close ring movement necessary for activation of its translocase activities. The studies, taken together with those of Chapter 2, demonstrate the important roles that protein movements, from small Brownian motion to large concerted modular movements, play during both programmed bacterial transcription termination mechanisms.

4.2 Future Directions

4.2.1 Determining the structures of the committed EC with cryo-EM

The earliest direct evidence of EC inactivation during intrinsic termination was discovered through single-molecule assays nearly twenty years ago [25]. In the intervening time period, comparisons of the relative rates of elongation (which occurs on the timescale of seconds even at a stable pause [26]) and dissociation (which can take minutes [3]) have provided rigorous support for the necessity of termination commitment. In Chapter 2 I likewise used a dissociation assay to demonstrate that release of both RNA and DNA from RNAP occurs over the course of several minutes even at the strong *t_{hisL}* terminator. However, the actual mechanism of commitment remains unknown. The current models propose that that hybrid shortening through shearing or hypertranslocation and displacement of the RNA 3'-end from the active site serves as the mechanism of commitment by preventing nucleotide addition [3, 4, 27] but rigorous biochemical and structural data in support of these models are lacking.

The advent of the cryo-electron microscopy (cryo-EM) revolution will at last enable mechanistic study of the commitment process [28]. Within the past year alone, cryo-EM has enabled the study of elongation mechanisms as complex as elemental and hairpin-stabilized pausing [29], the structural role of the transcription elongation factor NusA in promoting hairpin formation [30], defined how the coliphage transcription factor Nun induces transcriptional arrest [16], and defined the binding modes of the elongation factors NusG and RfaH [31]. Unlike classical crystallographic techniques, cryo-EM can be performed on protein complexes engaged in catalysis in solution, bypassing the need to establish a homogeneous population of static, “trapped” complexes. By reconstituting an EC on a weak terminator (*e.g.*, the artificial *t_{his2}* terminator used in kinetic analyses [7]), performing *in vitro* transcription at temperatures ranging

from low to physiological, and freezing reactions over a time course, it will be possible to identify meta-stable intermediates during the commitment process.

These analyses will can be used to structurally explore the hybrid shortening model of commitment [3, 4] and, just as importantly, be useful in probing the role of RNAP conformation throughout the commitment process. Although the MS-MP model of termination commitment predicts that the committed EC will be represented by a family of distinct RNAP conformations [7], it is likely that some members of this family will have deeper energy minima along its landscape of energy barriers to termination and will thus be enriched in the cryo-EM reaction mixture. The flexible nature of the unconstrained TL during MS-MP termination may disallow structural analysis *via* cryo-EM. On the other hand, if commitment is most favored by a class of closely-related TL conformations, then cryo-EM will readily enable detection.

Study of commitment by cryo-EM will likewise offer insight into the roles of other RNAP modules and motifs. For example, during hairpin-stabilized pausing the active site-proximal bridge helix (BH) “kinks” by 8°. These movements, combined with shifts in the downstream DNA, may prevent loading of the incoming template base into the active site [29]. Pausing also causes rearrangements in Switch 2, an element of RNAP that interacts with the RNA-DNA hybrid and mediates movements of the clamp [30]. Cryo-EM will serve as a powerful tool to evaluate whether these RNAP rearrangements, and others, persist beyond pausing and into termination commitment, and many of their findings can be complimented using biochemical tools such as the Cys-pair *in vitro* transcription assay.

4.2.2 MS-MP energetics and termination commitment during Rho-dependent termination

The expanded MS-MP model of intrinsic termination was proposed after the development of an assay that disentangled the relative rates of elongation from termination commitment [2, 7].

Previously, specific TL conformations had been proposed that drove the termination process.

During intrinsic termination, it was proposed that a folded TL drove termination [32], whereas an unfolded TL was proposed to drive Rho-dependent termination [33]. These conclusions arose from poorly-controlled experiments, however. In the first study (that proposed the folded TL as promoting termination), an anticolon antibody that stabilizes the unfolded form was found to suppress intrinsic termination efficiency. In the second, the antibody tagetotoxin, which targets the folded form of the TL, was likewise found to suppress TE by Rho. Importantly, the reciprocal tests were not performed, and the effects on elongation rate by these molecules was not evaluated. Given our newfound knowledge that termination works through MS-MP energetics [7], a more parsimonious explanation is that both intrinsic and Rho-dependent termination was inhibited because of restricted movements of a dynamic TL, not because the TL was biased into one specific conformation. That Rho-dependent termination may function through MS-MP sampling of conformation states is supported by the finding that backtracked pauses, known to block TL folding due to physical obstruction of the active site by the backtracked RNA [34], significantly reduced Rho termination efficiencies [35].

Mechanistic study of Rho during *in vitro* transcription is innately more difficult than intrinsic termination because it requires the additional of Rho, a fine-tuning of buffer and NTP conditions that permit its function, and because termination is induced over a long stretch of sequence rather than at one defined locus as in intrinsic termination. Nonetheless, recent innovations in the study of termination will enable these tests. Experiments based upon the

microcin J25 dissociation assay described in Chapter 2 can be used to determine how Rho responds to ECs with constrained TLs over the model λ_{R1} Rho terminator. Determining buffer conditions and NTP concentrations to ensure that the small molecule of choice binds and constrains the TL as the EC is positioned over sequences amenable to Rho action will require careful optimization. It will also be crucial that these experiments be done in the absence of TL inhibitors, as complete arrest of an EC (as opposed to temporary offline pausing) may in and of itself suppress Rho activity through an undefined mechanism [35].

The timescale of dissociation and the necessity for a stable inactivated EC during Rho-dependent termination is not as well defined as during intrinsic termination. Biochemical studies suggest that Rho most effectively induces termination as the EC is poised at non-stabilized elemental pauses [8, 35]. Unlike at intrinsic terminators, where hybrid shortening is the natural thermodynamic consequence of the sequence of the GC-dyad and the U-tract [3, 36, 37], hybrid shortening during Rho-dependent termination is promoted by the translocase activity of the Rho motor protein so strong that it can dissociate biotin-tethered RNAs from streptavidin [38-40]. Once Rho encounters the EC, hybrid shortening is driven by continued ATP hydrolysis. Whether the nascent transcript is extracted from the EC in one continuous pull/push, or whether the hybrid is first shortened by ~2–3 base pairs followed by EC inactivation will require the disentanglement of EC elongation rates, Rho translocase rates, termination efficiency, and RNA dissociation rates. Both the dissociation assay described in Chapter 2 and the termination kinetic assays developed by Ray-Soni *et al.* [7] will enable these investigations.

4.2.3 Recruitment of Rho to the EC

Early ChIP-chip experiments demonstrated similar distributions of Rho and RNAP along template DNA [41], but the nature of these Rho–DNA crosslinks remain poorly defined. Three explanations have been proposed: (i) Rho is bound directly DNA at the site of toxic R-loops, where the RNA invades the DNA duplex through base pairing to the template-strand and nontemplate-strand displacement, (ii) crosslinking is mediated through Rho binding to the nascent transcript and RNAP, or most intriguingly, (iii) that an undefined Rho–RNAP binding interaction occurs. One study demonstrated this potential interaction through pull-down assays of His₆-tagged RNAP and Rho across a transcription unit [33]. Based on these pulldowns, a “Rho-on-RNAP” recruitment model was proposed, wherein Rho stably associated without the EC throughout elongation, enabling rapid detection and binding to *rut* RNA upon extrusion from the exit channel. These results, which lacked negative controls, could not be replicated by another group under similar conditions [42], causing the authors to prefer the classical “Rho-on-RNA” model of direct binding to a *rut*. The most recent experimental evidence appears to favor the Rho-on-RNA model, with the original ChIP-chip results being best interpreted as the capture of transient interactions between RNAP and Rho as the latter is translocating along RNA through formaldehyde crosslinking [41]. Nonetheless, the Rho-on-RNAP model still persists within the literature, and future ChIP-chip studies could clarify the scientific record by varying formaldehyde concentrations or using Rho mutants that are deficient for *rut* binding or RNA translocation.

4.2.4 The termination mechanism in other species

The wide breadth of knowledge regarding the mechanisms of transcription termination have almost exclusively relied upon its study in *E. coli*. Remarkably little is known about Rho function and essentiality in other species, a glaring deficiency in the literature given our desire to engineer microorganisms for human benefit (e.g., biofuel production). To date, genome-level analysis of Rho activity has only been performed on one other species, *Bacillus subtilis*, a Gram-positive bacterium in which Rho activity is dispensable in rich medium [43, 44], and this analysis identified only 174 putative Rho termination sites. *B. subtilis* Rho shares 60% sequence identity with that of *E. coli* [44], but the size and sequences of the remaining 1,257 known Rho variants show high divergence in size and sequence [45]. Many of these variants contain large sequence insertions within their NTD. Biochemical analyses of the *Mycobacterium tuberculosis* Rho-NTD insertion demonstrated an enhancement for Rho affinity to *rut* sites at the expense of slowed motor activity [46]. The function of these insertions in other organisms deserves serious study.

The mechanism of intrinsic termination is not perfectly conserved across all bacterial lineages. *Mycobacterial* RNAPs are known to terminate at noncanonical U-tracts with higher GC-content, and its NusG paralog actually facilitates T_{hp} formation, promoting termination rather than acting as a pause suppressing factor as in *E. coli* [47]. Recent kinetic studies of the *E. coli* EC support the notion that T_{hp} formation induces a hairpin-stabilized pause [7]; however, hairpin formation in many species (e.g., *B. subtilis*) has no effect on pause strength [48] in the absence of elongation factors such as NusA or NusG. The role of the upstream A-tract in promoting TE *in vivo* has only been thoroughly characterized in *E. coli* [49]. Genome-scale analyses of A-tracts across diverse lineages could help define the mechanism by which they

enhance TE. Yet while the exact mechanism of intrinsic termination is not perfectly conserved, the general process shows remarkable consistency across species. Even the eukaryotic RNAPII, typically associated only with factor-dependent termination mechanisms, has been shown to respond to intrinsic terminator sequences *in vitro* [50]. Though we increasingly begin to rely on genome-scale bioinformatics to further our knowledge, it will still be necessary to complement “big data” with careful biochemical interrogations of fundamental mechanisms across these diverse lineages.

4.3 References

1. Roberts, J.W., *Termination factor for RNA synthesis*. Nature, 1969. **224**(5225): p. 1168-74.
2. Yager, T.D. and P.H. von Hippel, *A thermodynamic analysis of RNA transcript elongation and termination in Escherichia coli*. Biochemistry, 1991. **30**: p. 1097-1118.
3. Komissarova, N., et al., *Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination*. Mol Cell, 2002. **10**(5): p. 1151-62.
4. Larson, M.H., et al., *Applied force reveals mechanistic and energetic details of transcription termination*. Cell, 2008. **132**(6): p. 971-82.
5. Larson, M.H., et al., *A pause sequence enriched at translation start sites drives transcription dynamics in vivo*. Science, 2014. **344**(6187): p. 1042-7.
6. Peters, J.M., A.D. Vangeloff, and R. Landick, *Bacterial transcription terminators: the RNA 3'-end chronicles*. J Mol Biol, 2011. **412**(5): p. 793-813.
7. Ray-Soni, A., R.A. Mooney, and R. Landick, *Trigger loop dynamics can explain stimulation of intrinsic termination by bacterial RNA polymerase without terminator hairpin contact*. Proc Natl Acad Sci U S A, 2017. **114**(44): p. E9233-E9242.
8. Jin, D.J., et al., *Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho*. Proc Natl Acad Sci U S A, 1992. **89**(4): p. 1453-7.
9. Thomsen, N.D. and J.M. Berger, *Running in reverse: the structural basis for translocation polarity in hexameric helicases*. Cell, 2009. **139**(3): p. 523-34.

10. Thomsen, N.D., et al., *Molecular mechanisms of substrate-controlled ring dynamics and substepping in a nucleic acid-dependent hexameric motor*. Proc Natl Acad Sci U S A, 2016. **113**(48): p. E7691-E7700.
11. Richardson, J.P., *Activation of Rho protein ATPase requires simultaneous interaction at two kinds of nucleic acid-binding sites*. J Biol Chem, 1982. **257**(10): p. 5760-5766.
12. Windgassen, T.A., et al., *Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase*. Nucleic Acids Res, 2014. **42**(20): p. 12707-21.
13. Chakraborty, A., et al., *Opening and closing of the bacterial RNA polymerase clamp*. Science, 2012. **337**(6094): p. 591-5.
14. Nayak, D., et al., *Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase*. Mol Cell, 2013. **50**(6): p. 882-93.
15. Vassylyev, D.G., et al., *Structural basis for transcription elongation by bacterial RNA polymerase*. Nature, 2007. **448**(7150): p. 157-62.
16. Kang, J.Y., et al., *Structural basis of transcription arrest by coliphage HK022 N_{un} in an Escherichia coli RNA polymerase elongation complex*. Elife, 2017. **6**.
17. Felklistov, A., et al., *RNA polymerase motions during promoter melting*. Science, 2017. **356**: p. 863-866.
18. Cramer, P., D. Bushnell, and R. Kornberg, *Structural basis of transcription: RNA polymerase II at 2.8 Å resolution*. Science, 2001. **292**(5523): p. 1863-76.
19. Frauenfelder, H., S.G. Sligar, and P.G. Wolynes, *The energy landscapes and motions of proteins*. Science, 1991. **254**(5038): p. 1598-1603.
20. Noé, F., et al., *Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations*. Proc Natl Acad Sci U S A, 2009. **106**(45): p. 19011-19016.
21. Wang, H., et al., *Protein dynamics control the kinetics of initial electron transfer in photosynthesis*. Science, 2007. **316**(5825): p. 747-750.
22. Peters, J.M., et al., *Rho directs widespread termination of intragenic and stable RNA transcription*. Proc Natl Acad Sci U S A, 2009. **106**(36): p. 15406-11.
23. Burns, C.M. and J.P. Richardson, *NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator*. Proc Natl Acad Sci U S A, 1995. **92**(11): p. 4738-42.
24. Valabhoju, V., S. Agrawal, and R. Sen, *Molecular Basis of NusG-mediated Regulation of Rho-dependent Transcription Termination in Bacteria*. J Biol Chem, 2016. **291**(43): p. 22386-22403.

25. Yin, H., et al., *Nonequilibrium mechanism of transcription termination from observations of single RNA polymerase molecules*. Proc Natl Acad Sci U S A, 1999. **96**(23): p. 13124-9.
26. Touloukhonov, I., et al., *A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing*. Mol Cell, 2007. **27**(3): p. 406-19.
27. Ray-Soni, A., M.J. Bellecourt, and R. Landick, *Mechanisms of Bacterial Transcription Termination: All Good Things Must End*. Annu Rev Biochem, 2016. **85**: p. 319-47.
28. Binshtein, E. and M.D. Ohi, *Cryo-electron microscopy and the amazing race to atomic resolution*. Biochemistry, 2015. **54**(20): p. 3133-41.
29. Kang, J.Y., et al., *RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing*. Mol Cell, 2018. **69**(5): p. 802-815 e1.
30. Guo, X., et al., *Structural Basis for NusA Stabilized Transcriptional Pausing*. Mol Cell, 2018. **69**(5): p. 816-827 e4.
31. Kang, J.Y., et al., *Structural Basis for Transcript Elongation Control by NusG Family Universal Regulators*. Cell, 2018. **173**(7): p. 1650-1662.e14.
32. Epshtein, V., et al., *An allosteric path to transcription termination*. Mol Cell, 2007. **28**(6): p. 991-1001.
33. Epshtein, V., et al., *An allosteric mechanism of Rho-dependent transcription termination*. Nature, 2010. **463**(7278): p. 245-9.
34. Sekine, S., et al., *The ratcheted and ratchetable structural states of RNA polymerase underlie multiple transcriptional functions*. Mol Cell, 2015. **57**(3): p. 408-21.
35. Dutta, D., J. Chalissery, and R. Sen, *Transcription termination factor rho prefers catalytically active elongation complexes for releasing RNA*. J Biol Chem, 2008. **283**(29): p. 20243-51.
36. Lubkowska, L., A.S. Maharjan, and N. Komissarova, *RNA folding in transcription elongation complex: implication for transcription termination*. J Biol Chem, 2011. **286**(36): p. 31576-85.
37. d'Aubenton Carafa, Y., E. Brody, and C. Thermes, *Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures*. J Mol Biol, 1990. **216**(4): p. 835-58.
38. Schwartz, A., et al., *Transcription termination factor rho can displace streptavidin from biotinylated RNA*. J Biol Chem, 2007. **282**(43): p. 31469-76.

39. Park, J.S. and J.W. Roberts, *Role of DNA bubble rewinding in enzymatic transcription termination*. Proc Natl Acad Sci U S A, 2006. **103**(13): p. 4870-5.
40. Richardson, J.P., *Rho-dependent termination and ATPases in transcript termination*. Biochim Biophys Acta, 2002. **1577**(2): p. 251-260.
41. Mooney, R.A., et al., *Regulator trafficking on bacterial transcription units in vivo*. Mol Cell, 2009. **33**(1): p. 97-108.
42. Kalyani, B.S., et al., *Interaction with the nascent RNA is a prerequisite for the recruitment of Rho to the transcription elongation complex in vitro*. J Mol Biol, 2011. **413**(3): p. 548-60.
43. Nicolas, P., et al., *Condition-dependent transcriptome reveals high-level regulatory architecture in Bacillus subtilis*. Science, 2012. **335**(6072): p. 1103-6.
44. Quirk, P.G., et al., *Identification of a putative Bacillus subtilis rho gene*. Journal of Bacteriology, 1993. **175**(3): p. 647-654.
45. D'Heygere, F., M. Rabhi, and M. Boudvillain, *Phyletic distribution and conservation of the bacterial transcription termination factor Rho*. Microbiology, 2013. **159**(Pt 7): p. 1423-36.
46. D'Heygere, F., et al., *ATP-dependent motor activity of the transcription termination factor Rho from Mycobacterium tuberculosis*. Nucleic Acids Res, 2015. **43**(12): p. 6099-111.
47. Czyz, A., et al., *Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators*. MBio, 2014. **5**(2): p. e00931.
48. Artsimovitch, I., et al., *RNA polymerases from Bacillus subtilis and Escherichia coli differ in recognition of regulatory signals in vitro*. J Bacteriol, 2000. **182**(21): p. 6027-35.
49. Chen, Y.J., et al., *Characterization of 582 natural and synthetic terminators and quantification of their design constraints*. Nat Methods, 2013. **10**(7): p. 659-64.
50. Reines, D., et al., *Identification of intrinsic termination sites in vitro for RNA polymerase II within eukaryotic gene sequences*. Journal of Molecular Biology, 1987. **196**(2): p. 299-312.

Appendix A

Hairpin-stabilized but not elemental pause durations are affected
by clamp topology

Appendix contributions:

M.J.B. wrote the text, performed the experiments, and constructed the figures.

A.1 Introduction

Transcriptional pausing is the process by which the elongation complex (EC) temporarily halts elongation due to rearrangements in the nucleic-acid (NA) scaffold and RNA polymerase (RNAP) that suppress the rate of nucleotide addition. Rather than being a “flaw” in the evolution of the EC, pausing facilitates the coordination of diverse cellular processes, including RNA folding, transcription factor recruitment, splicing, and is the necessary first component of the termination mechanism [1-3]. Structural studies suggest that the initial pause state, termed the elemental pause, arises in a sequence-dependent [4] manner during the translocation step of the nucleotide addition cycle. These sequences allow RNA to successfully move into the post-translocated register but “capture” the incoming template base on the active site-proximal bridge helix, trapping the template strand in the pre-translocated register [5]. These elemental pauses can be further stabilized either through reverse translocation onto more stable NA sequences (backtrack-stabilized pauses) or through formation of an RNA hairpin in the RNA exit channel (hairpin-stabilized pauses). Pause stabilization can increase durations of the pause more than ten-fold [6].

Structural and biochemical studies have suggested that the clamp is displaced *via* a swiveling motion from the closed conformation that it adapts throughout elongation at a well-studied hairpin-stabilized pause, the *his* pause (*hisP*) [5, 7, 8]. The role of clamp movements at the elemental pause are more poorly understood. Preliminary cryo-electron microscopy (cryo-EM) found no evidence of clamp displacement in a subpopulation of particles reconstituted on *hisP* sequences that lacked evidence of a hairpin [5]. On the other hand, a biochemical reporter of clamp conformation of ECs reconstituted on a consensus elemental pause sequence [4] found evidence that the elemental pause can drive a degree of clamp displacement intermediate to the

*his*PEC and a closed-clamp EC, suggesting that the elemental pause loosens the clamp but drives only transient movements that have not been captured by (R. Mooney, personal communications). Further, movements of the clamp (if any) during backtracking are completely unknown. To study the motions of the clamp during elemental pausing, we exploited the Cys-pair ligation scaffold assay and performed *in vitro* transcription through three separate pause sequences. Unfortunately, the complicated interplay between variable transcription elongation rates, pause strength, and crosslinking efficiencies precluded a fine kinetic analysis of pause strengths, but these studies enabled a general understanding of how topological constraints in the clamp perturbed pause duration at three pause signals.

A.2 Results

A.2.1 Topologically constraining the clamp perturbs the duration of the *his* hairpin-stabilized pause

We first sought to confirm that a ligation scaffold based approach (Fig. A.1A) could recapitulate the previous demonstration that using cysteine-pair (Cys-pair) RNAPs to introduce a topological constraint into the clamp would perturb the duration of the *his*P. The two variants used in the original study include C-LF, modelled to bias the clamp into a conformation observed in elongation complexes (closed-clamp, lid-flap; β P1044C β' 258iC), and D-LF, which should displace the clamp into a non-productive conformation by either biasing the swiveled conformation, the open conformation, or some other similar state (displaced-clamp, lid-flap; β T843C β' 258iC). In addition, we also utilized two alternative Cys-pair RNAPs generated to explore the role of lid shortening (C- Δ 3 and C-259A).

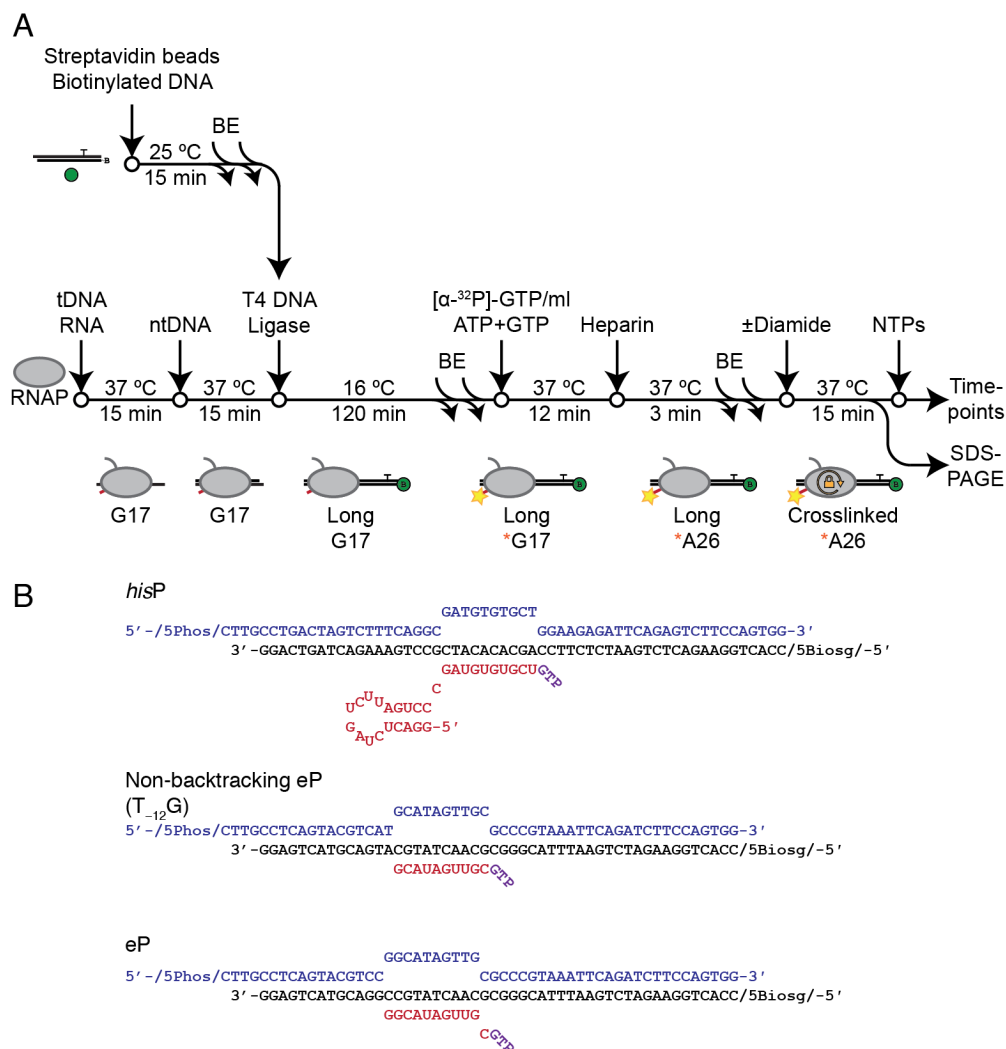


Figure A.1. Cys-pair ligation scaffold assay. **(A)** Assay schematic. Adapted from Chapter 2. BE, buffer exchange; the magnetic properties of the Streptavidin bead are exploited to eliminate unbound RNAPs and ECs that fail to ligate downstream DNAs of interest. **(B)** Pause sequence-containing oligonucleotides used in these experiments. RNA components represent bases within the RNA-DNA hybrid at the pause site and, in the case of the *hisP*, the sequence and structure of the pause hairpin.

During a Cys-pair ligation scaffold assay, RNAP is initially reconstituted on a short oligonucleotide scaffold with a template-strand 5' phosphate and 4-nt overhang for subsequent ligation (Fig. A.1A). Oligonucleotide pairs bearing pause sequences of interest (Fig A.1B) and a complementary overhang were tethered to paramagnetic streptavidin beads through a 5' biotin moiety on the template-strand and ligated to the G17-reconstituted ECs with T4 DNA Ligase, generating “long” ECs. These long ECs were extended to the A26 position with ATP and [α - ^{32} P] GTP and treated with diamide to generate Cys-pair crosslinks. Transcription elongation was then restarted by addition of all four NTPs at 150 μ M A/U/CTP and 30 μ M GTP, and its progress along the DNA template was monitored at various time points at 37 °C. These NTP concentrations were chosen to (i) promote elongation to the pause signal in as little time as reasonable while (ii) inducing as strong a pause as possible given the incoming G (Fig. A.1B).

As described above, the interconnected nature of the pause-free elongation rates, pause kinetics, and the incomplete nature of disulfide formation precluded analysis of pause strengths using the Cys-pair ligation scaffold. However, by plotting pause occupancy as a fraction of ECs present at the pause site as a function of time, we were able to determine that the ligation scaffold assay recapitulated previous findings (Fig. A.1). Across the set of five RNAPs tested under two redox conditions, the maximal pause occupancy was ~25%. Oxidation of the WT RNAP had no effect on the duration of the *hisP*, whereas formation of a crosslink by the C-LF RNAP decreased the duration of the pause. This effect was also seen with C-259A, an RNAP that was shown to behave similarly to C-LF in *in vitro* termination assays (see chapter 2). The D-LF RNAP was, as predicted, slow to escape the *hisP*; observing completion of pause escape under oxidizing conditions required taking time points more than 1 hr after adding chase NTPs. Interestingly, the C- Δ 3 RNAP, which behaved similarly to the D-LF RNAP in an *in vitro*

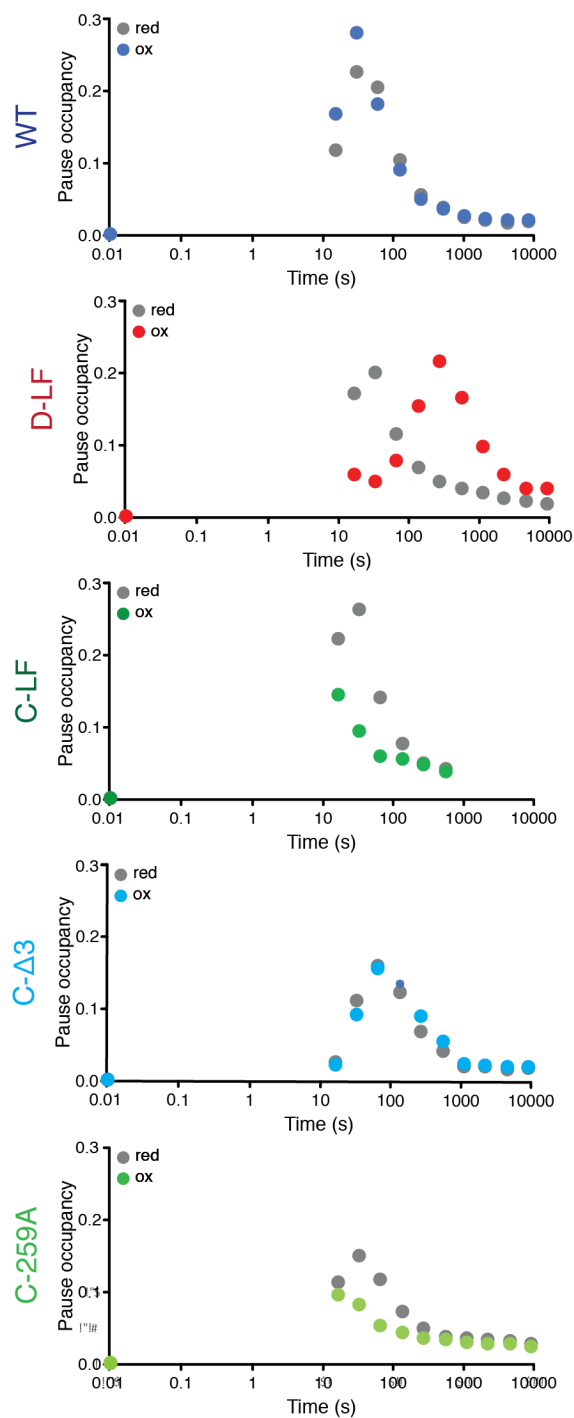


Figure A.2. Effects of topologically constraining the clamp on the *hisP*. Pause occupancy was calculated as the ratio of transcripts at the pause site over all transcripts (pre-pause, pause, and post-pause). Arrested ECs were subtracted out of the calculation. Grey points indicate ECs under reducing conditions, colored points indicate ECs under oxidizing conditions. Points are the average of three replicates.

termination assay due to its deficient elongation rate (see Chapter 2), did not display a differential response to the *hisP* between reducing and oxidizing conditions. Nonetheless, taken together these results show that the ligation scaffold assay can be a useful as an assay for obtaining a broader picture of elongation behavior, as it can recapitulate results obtained on minimal *hisP* scaffolds.

A.2.2 Biasing the clamp into a displaced but not closed conformation perturbs elemental pausing duration

Having confirmed that the ligation scaffold technique can recapitulate the effects of the *hisP* *in vitro* we next sought to evaluate the role of clamp conformation during the elemental pause, using oligonucleotides that encoded the consensus pause sequence (Fig. A.1B) but could not template hairpins in the nascent transcript. We created two derivatives of the elemental pause; one encoded a T₋₁₂ while the second encoded G₋₁₂. Both sequences are capable of inducing elemental pausing, but the second sequence (G₋₁₂) is more likely to promote backtrack-stabilized pausing because backtracking onto G₋₁₂ forms a more stable RNA-DNA hybrid.

In general, each RNAP responded to the the T₋₁₂ elemental pause and the G₋₁₂ pause under both reducing and oxidizing conditions in similar fashions (Fig. A.3). The WT RNAP, unable to form a crosslink, had no observable difference in pause occupancy between reducing and oxidizing conditions. While the C-LF and C-259A RNAPs were shown to have a strong resistance to this *hisP* under oxidizing conditions, they exhibited no change in elongation behavior over both classes of elemental pauses. The C-Δ3 RNAP likewise did not exhibit a changed response to either class of elemental pause upon formation of a crosslink.

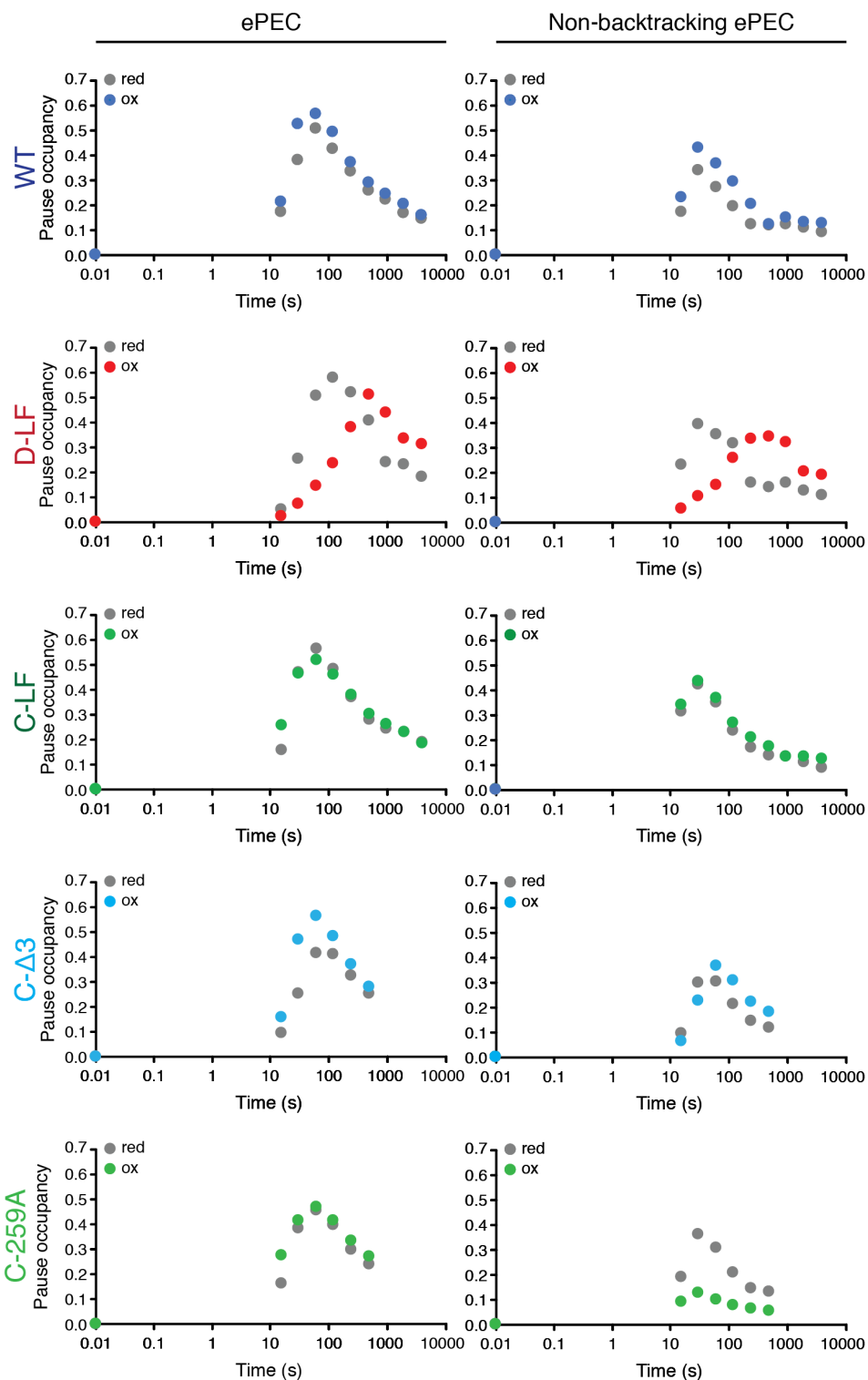


Figure A.3. Effects of topologically constraining the clamp on two classes of the elemental pause. Pause occupancy was calculated as the ratio of transcripts at the pause site over all transcripts (pre-pause, pause, and post-pause). Arrested ECs were subtracted out of the calculation. Grey points indicate ECs under reducing conditions, colored points indicate ECs under oxidizing conditions. Points are the average of three replicates.

The D-LF EC, on the other hand, did appear to exhibit a differential response to both tested classes of elemental pauses. The very slow elongation behavior of D-LF under conditions that favor formation of its crosslink complicated these analyses, and the semi-log plotting of pause occupancy obscures the fact that D-LF ECs approached the pause signal at ~ 1 min and escaped on the timeframe of minutes to hours.

A.4 Discussion

The complex interplay between basal elongation rates, pause kinetics, and crosslinking efficacy prevented us from deriving precise rates of pause escape from any of the three tested pause signals. Nonetheless, a mixed quantitative-qualitative approach to pause occupancies does begin to reveal the roles of the clamp during transcriptional pausing. We first established that our assay design could recapitulate the analyses of the *hisP* on a more minimal scaffold [8]. Having done so, we also tested how two other Cys-pair RNAPs behaved when permitted to elongate through the *hisP*. The C-259A RNAP behaved as was expected, based on its elongation and termination behavior; like C-LF, it was resistant to the *hisP* upon oxidation, likely because disulfide formation sufficiently closed the clamp to narrow the exit channel and slow the formation of the RNA hairpin.

The C- $\Delta 3$ EC has a slow elongation rate with behaviors that are minimally influenced after formation of a crosslink, even at the *hisP*. The reasons for C- $\Delta 3$'s slowed elongation rate remain unclear. The C-259A EC elongates at nearly the same rate as C-LF under both reducing and oxidizing conditions, suggesting that the deficiency in C- $\Delta 3$ is either due to (i) the loss of L255 or D256 or (ii) changes in the backbone structure of lid (or both). A recent cryo-EM study proposed that L255 (with V253) facilitated separation of RNA from the template-strand through hydrophobic interactions and that D256 forms a salt bridge with $\beta R841$ to prevent re-annealing

of the transcript and template [9]. An appealing hypothesis is that the C- Δ 3 deletions have remodeled the upstream fork junction, perhaps by changing how the lid moves and interacts with the -10 base pair during translocation. However, this is at odds with previous biochemical reports that demonstrated no effect on elongation by a full lid deletion [10]. It is possible that the C- Δ 3 deletions enable lid-NA interactions disfavorable for transcription, and that full deletion of the lid abolishes these interactions and restores the EC's ability to effectively translocate. Further study into the role of the complex lid-NA contact network in an EC will be necessary to fully understand the upstream fork junction and translocation.

Neither the C-LF nor C-259A ECs responded to the two tested elemental pause sequences, in strong contrast to their ability to suppress the duration of the *hisP*. These findings support the preliminary cryo-EM structures that suggest that elemental pausing does not require movements of the RNAP swivel module [5]. Curiously, however, formation of a crosslink by C-259A does appear to reduce the maximal occupancy of the pause for the T₋₁₂ ePEC variant that disfavors backtracking. The substitution eliminates β' R259, which forms a stacking interaction with the -9 base of the template strand [9], but it is unclear both why this effect would only occur on ePEC sequences that are disfavorable to backtracking and would not be recapitulated by the C- Δ 3 EC that likewise lacks β' R259.

Finally, displacement of the clamp from the closed conformation by D-LF dramatically increased the pause duration of both ePECs, such that some fraction of crosslinked ECs remained paused after two hours of transcription. While the elemental pause does not itself seem to drive movements of the swivel module [5], the usage of Cys-pairs to bias the swiveled conformation are evidently sufficient to induce stabilization of the ePEC. These results are in agreement with

previous Cys-pair biochemical data that demonstrated that displacement of the clamp could stabilize the *hisP* even in the absence of a pause hairpin or other exit channel duplex [8].

While the interplay between elongation, pausing, and crosslinking prevented a rigorous analysis of pause strength, these experiments nonetheless demonstrate that the ligation scaffold assay can be utilized to study any number of transcriptional processes. The assay bypasses the complex initiation mechanism and enables *in vitro* transcription of long stretches of DNA in the absence of σ factors, which are now known to interact with ECs even after promoter escape [11] and could confound data interpretation.

A.5 Materials and Methods

A.5.1 Materials

All DNA and RNA oligonucleotides were from IDT (Table A.1). Oligonucleotides used for reconstitution of elongation complexes were gel purified by denaturing PAGE. FPLC-purified NTPs were from Promega Corp. [α - ^{32}P] GTP was from Perkin-Elmer. T4 *Sty* I-HF and T4 DNA Ligase were from NEB, Ipswich, MA.

A.5.2 Protein purifications

Wild-type and variant RNAPs were overexpressed from plasmids (Table A.2). The parent plasmid pRM843 is a T7 RNAP-based overexpression plasmid encoding *E. coli rpoA*, *rpoZ*, *rpoB* and *rpoC*, yielding core RNAP. The N-terminus of β contains a His₁₀ tag followed by a PreScission protease cleavage recognition signal. The C-terminus of β' contains a *Strep*-tag[®] preceded by a Protein Kinase A (PKA) recognition signal. RNAPs were purified from *E. coli* BLR DE3 transformed as previously described [8] and dialyzed into a high [DTT] RNAP storage

Table A.1. Oligonucleotides used in this study.

Oligonucleotide name	Sequence	Purpose
RNA-8855	5'-UUUUUUUGAGCUAGAGG-3'	Reconstitution scaffold, RNA component
tDNA-10160	5'-/5Phos/CAAGGGACTGGTC TGAATCGGGAATACGGAACGGAA GATCTGAGTTCTCTTCCCCTCTA GCTCAGGACGTACTGACC-3'	Reconstitution scaffold, template DNA component
ntDNA-10161	5'-GTCAGTACGTCCTGAGCTAG AGGGGAAGAGAACTCAGATCTTC CGTTCCGTATTCCCGATTGAGAC CAGTCC-3'	Reconstitution scaffold, nontemplate DNA component
dsDNA-11029	5'-/5Phos/CTTGCCTCAGTAC GTCCGGCATAGTTGCGCCCGTAA ATTCAGATCTTCCAGTGG-3'	G ₋₁₂ ePEC, template strand
dsDNA-11030	5'-/5Biosg/CCACTGGAAGAT CTGAATTTACGGGCGCAACTATG CCGGACGTACTGAGG-3'	G ₋₁₂ ePEC, nontemplate strand
dsDNA-11031	5'-/5Phos/CTTGCCTCAGTAC GTCATGCATAGTTGCGCCCGTAA ATTCAGATCTTCCAGTGG-3'	T ₋₁₂ ePEC, template strand
dsDNA-11032	5'-/5Biosg/CCACTGGAAGAT CTGAATTTACGGGCGCAACTATG CATGACGTACTGAGG-3'	T ₋₁₂ ePEC, nontemplate strand
dsDNA-11033	5'-/5Phos/CTTGCCTGACTAG TCTTTCAGGCGATGTGTGCTGGA AGAGATTCAGAGTCTTCCAGTGG -3'	<i>his</i> PEC, template strand

dsDNA-11034	5'-/5Biosg/CCACTGGAAGAC TCTGAATCTCTTCCAGCACACAT CGCCTGAAAGACTAGTCAGG-3'	<i>his</i> PEC, nontemplate strand
-------------	---	------------------------------------

Table A.2. Plasmids used in this study.

Plasmid name	Source	Purpose
pRM843	[29]	Overexpression of wild-type RNAP (<i>rpoA</i> , <i>rpoZ</i> His ₁₀ :PPX: <i>rpoB</i> , <i>rpoC</i> :PKA: <i>strep</i>)
pRM890	[29]	pRM843 derivative. Overexpression of D-LF RNAP (β T843C β' 258iC)
pRM918	[29]	pRM843 derivative. Overexpression of C-LF RNAP (β P1044C β' 258iC)
pRM1100	This study	pRM843 derivative. Overexpression of C- Δ 3 RNAP (β P1044C β' Δ 255–256,258iC, Δ 259)
pRM1177	This study	pRM843 derivative. Overexpression of C-259A RNAP (β P1044C β' 258iC R259A)

buffer (10 mM Tris-Cl, pH 7.9, 20% [v/v] glycerol, 100 mM EDTA, 10 mM DTT, 100 mM NaCl, 1 mM MgCl₂, 20 μ M ZnCl) to ensure the integrity of the cysteine sulfhydryl groups. Aliquots were stored at -80°C .

A.5.3 Transcription on fully complementary oligonucleotide scaffolds

ECs were reconstituted step-wise. First, excess template strand (10 μ M) was annealed to 5 μ M RNA in reconstitution buffer [10 mM Tris-OAc, pH 7.9, 40 mM KOAc, 5 mM Mg(OAc)₂] by heating to 75°C for 2 min, rapidly cooling to 45°C , and then cooling to room temperature in $2^{\circ}\text{C}/2\text{-min}$ steps. Downstream oligonucleotides containing pause signals were annealed likewise. The RNA-tDNA scaffold was diluted to 500 nM by addition of elongation buffer (20 mM Tris-OAc, pH. 8.0, 40 mM KOAc, 5 mM Mg(OAc)₂, 0.1 mM EDTA, 1 mM DTT) followed by addition of 2.5 μ M RNAP and incubation at 37°C for 15 min. Nontemplate strand was then added to 1.5 μ M and the sample was again incubated at 37°C for 15 min, yielding an EC formed with the molar ratio 1 RNA:2 tDNA:3 ntDNA:5 RNAP. Downstream DNA oligonucleotides were pre-bound to paramagnetic beads through a 5' biotin moiety on the template strand. 150 nM bead-bound fragments were mixed with 50 nM ECs, a 3:1 ratio, and ligated together with 2,000 U T4 DNA ligase and 1 mM ATP in EB at 16°C for 2 hr. Non-ligated ECs and excess RNAP were removed from the mixtures through two EB washes on a magnetic rack. RNAs were labeled by addition of 0.2 mCi [$\alpha\text{-}^{32}\text{P}$] GTP and a heparin competitor was added to 50 $\mu\text{g}/\text{ml}$. Disulfide crosslinking of the Cys-pairs was initiated by adding CSSC to 1.5 mM or diamide to 15 mM and incubating ECs at 37°C for 15 min. To determine crosslinking efficiencies, aliquots were removed, quenched with 15 mM iodoacetamide, and separated through SDS-PAGE on 8% Bolt gels in 1X Bolt running buffer, and quantitated *via* PhosphorImaging. The remaining

sample was diluted in EB $\pm$ oxidant to a working concentration of 5 nM. Transcription was initiated by saturation with 150 μ M ATP, UTP, and CTP and 30 μ M GTP. Reactions were stopped at 15, 30, 60, 120, 240, and 480 s by addition of a urea stop buffer (final concentration 4 M urea, 44.5 mM Tris-borate, pH 8.3, 50 mM EDTA, 0.01% [v/v] xylene cyanol, and 0.01% [v/v] bromophenol blue). Time points were also taken at 960, 1920, 3840, and 7680 s for slowly-elongating ECs. Samples were separated by electrophoresis in a denaturing 8% (v/v) polyacrylamide gel and quantitated *via* PhosphorImaging.

A.6 References

1. Mayer, A., H.M. Landry, and L.S. Churchman, *Pause & go: from the discovery of RNA polymerase pausing to its functional implications*. Curr Opin Cell Biol, 2017. **46**: p. 72-80.
2. Zhang, J. and R. Landick, *A Two-Way Street: Regulatory Interplay between RNA Polymerase and Nascent RNA Structure*. Trends Biochem Sci, 2016. **41**(4): p. 293-310.
3. Ray-Soni, A., M.J. Bellecourt, and R. Landick, *Mechanisms of Bacterial Transcription Termination: All Good Things Must End*. Annu Rev Biochem, 2016. **85**: p. 319-47.
4. Larson, M.H., et al., *A pause sequence enriched at translation start sites drives transcription dynamics in vivo*. Science, 2014. **344**(6187): p. 1042-7.
5. Kang, J.Y., et al., *RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing*. Mol Cell, 2018. **69**(5): p. 802-815 e1.
6. Touloukhonov, I., I. Artsimovitch, and R. Landick, *Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins*. Science, 2001. **292**(5517): p. 730-3.
7. Guo, X., et al., *Structural Basis for NusA Stabilized Transcriptional Pausing*. Mol Cell, 2018. **69**(5): p. 816-827 e4.
8. Hein, P.P., et al., *RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement*. Nat Struct Mol Biol, 2014. **21**(9): p. 794-802.
9. Kang, J.Y., et al., *Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex*. Elife, 2017. **6**.

10. Touloukhonov, I. and R. Landick, *The role of the lid element in transcription by E. coli RNA polymerase*. J Mol Biol, 2006. **361**(4): p. 644-58.
11. Harden, T.T., et al., *Bacterial RNA polymerase can retain sigma70 throughout transcription*. Proc Natl Acad Sci U S A, 2016. **113**(3): p. 602-7.