Title
Mechanistic Studies of the Rho Helicase

Author
Lawson, Michael Robert

Publication Date
2016

Peer reviewed|Thesis/dissertation

Mechanistic Studies of the Rho Helicase

By

Michael Robert Lawson

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor James M. Berger, co-chair
Professor Donald Rio, co-chair
Professor Andreas Martin
Professor John Kuriyan
Professor David Wemmer

Fall 2016

Abstract

Mechanistic Studies of the Rho Helicase

by

Michael Robert Lawson

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor James M. Berger, Co-Chair

Professor Donald Rio, Co-Chair

All cellular organisms rely on transcriptional regulatory mechanisms to control the levels and timing of gene expression. Bacterial transcription is in part regulated by Rho, a RecA-family hexameric helicase that utilizes an ATP-dependent 5'→3' translocation mechanism to terminate synthesis of specific RNA transcripts. Rho is initially recruited to pyrimidine-rich transcripts in an open-ring, loading-competent configuration, and subsequently traps the RNA within Rho's central pore by transitioning to a closed-ring and catalytically active conformation. Using small-angle X-ray scattering and a novel fluorescence-based assay to monitor Rho's conformational state *in vitro*, I discovered that bicyclomycin, a known small molecule inhibitor of Rho, acts by sterically blocking isomerization of Rho into its closed-ring form. Conversely, I also demonstrated that two distinct Rho effectors – pyrimidine-rich nucleic acids (as are found in favored RNA elements) and the transcription factor NusG – directly promote Rho ring closure. To better understand how NusG modulates Rho activity, I determined a crystal structure of closed-ring Rho in complex with the NusG C-terminal domain. This structure reveals that NusG engages the C-terminal face of Rho's motor domains. Modeling of a Rho/NusG/RNAP complex based on the structure suggests that NusG may position Rho near the RNA exit channel of RNA polymerase to aid in the capture of non-ideal RNAs. Collectively, these findings delineate how a diverse set of ligands regulate bacterial transcription termination, and demonstrate how the conformational state of a hexameric helicase can be controlled by disparate classes of factors.

Table of Contents

Chapter 5 – Future Directions and Final Remarks

Acknowledgements

First and foremost, I am especially grateful for the love and support of my wife, Christi Preston, who has backed me through the many highs and lows of graduate school. I would also like to thank my thesis advisor, James Berger, for training me as a biochemist and teaching me how to focus on the task at hand. I am also grateful for guidance and feedback from my committee members Don Rio, Andy Martin, Dave Wemmer and John Kuriyan; members of the Berger lab for their support and input over the years; and both Andy and all of the members of his lab for welcoming me into their group after the move. I would also like to thank my collaborators Nathan Thomsen, Kevin Dyer, Tanja Kortemme, Roland Pache, Wen Ma and Klaus Schulten; James Holton (ALS 8.3.1), Greg Hura (ALS 12.3.1), Tzanko Doukov (SLAC), and Nat Echols (LBL) for assistance with experiments and data processing; and Michael Bellecourt (UW-Madison) for helpful feedback and suggestions. I am also grateful for support from my family (especially my mother, my father, and my uncle T.J.) and friends (including Jesse, Amanda, Brian, Tim and Romell). This work was funded by a National Science Foundation Graduate Research Fellowship and the Mathers Foundation. Finally, this dissertation is dedicated to my father, Donald Burke Lawson Jr., who is sorely missed.

Chapter 1: Introduction to bacterial transcription termination

(Portions of this chapter have been reproduced from Lawson M.R., Dyer K. and Berger J.M., PNAS 2016).

Regulation of transcription

All cellular organisms rely on transcriptional regulatory mechanisms to control the levels and timing of gene expression. Bacteria modulate transcription for important cellular roles such as altering metabolism in response to changes in available nutrients (e.g. the *lac* and *trp* operons (Cohn et al., 1953; Kennell and Riezman, 1977)), activating responses required for survival of stress conditions (Marles-Wright and Lewis, 2007), and synchronizing the initiation of virulence programs with conditions that are favorable for growth (DiRita and Mekalanos, 1989). Many changes in gene expression are governed by transcriptional activators or repressors, which regulate specific operons by promoting or blocking the formation of transcription initiation complexes (Browning and Busby, 2004; Jacob and Monod, 1961). Transcription elongation is subject to further regulation, often during polymerase pausing events (Belogurov and Artsimovitch, 2015). Finally, changes in gene expression can also be driven at the molecular level by adjustment of the point at which transcripts are terminated.

Intrinsic and Rho-dependent transcription termination

Transcription termination in bacteria is predominantly regulated through two mechanisms – intrinsic termination and Rho-dependent termination. Both regulatory approaches have been studied extensively on a genome-wide scale in *Escherichia coli* (Peters et al., 2009). In intrinsic termination (depicted in **Figure 1.1A**), which is responsible for ~80% of transcription termination in *E. coli*, RNA polymerase transcribes a hairpin that sits directly adjacent to a poly-uridine tract of roughly 7-8 Uridine bases (d'Aubenton Carafa et al., 1990). The presence of a run of uridines within the transcription bubble both weakens the RNA:DNA hybrid in the RNA polymerase catalytic center and elicits pausing of transcription by the enzyme (Gusarov and Nudler, 1999). Three different models have been proposed that explain how the subsequent folding of the hairpin directly adjacent to the transcription bubble drives termination of transcription. In the hybrid shearing model, the RNA:DNA hybrid is disrupted by the incorporation of bases from the poly-U tract into the hairpin (Yarnell and Roberts, 1999). In the hypertranslocation scheme, the application of a direct force to the polymerase by the newly-formed hairpin causes the polymerase to translocate before the next base can be incorporated, preventing subsequent rNMP addition (Santangelo and Roberts, 2004). In the hairpin invasion model, the RNA:DNA hybrid is displaced from the transcription bubble by the hairpin forming within the polymerase active site (Epshtein et al., 2007). Motifs required for intrinsic termination can be concealed or exposed through changes in local RNA structure, which is commonly controlled by regulatory elements such as riboswitches (Mironov et al., 2002).

Bacterial transcription termination events not controlled by intrinsic mechanisms typically require a protein-based transcription termination factor termed Rho (**Figure 1.1B**). Rho-dependent termination is initiated after RNA polymerase transcribes through

a generally C-rich and unstructured Rho utilization (or "*rut*") sequence, which serves as a recruitment site for Rho. It is unclear whether the binding of Rho to a *rut* sequence represents the first interaction Rho has with a transcription complex (Kalyani et al., 2011), or whether Rho first binds to RNA polymerase earlier in transcription, concurrent with the release of σ factor (Mooney et al., 2009a). Regardless of order, Rho is thought to engage a *rut*-containing transcript by forming a topologically open and loading-competent hexameric ring (**Figure 1.2A**) that can bind YC dinucleotide motifs (where Y is a pyrimidine, and C is a cytosine) within the *rut* sequence. An oligonucleotide/oligosaccharide binding ("OB") fold located in the N-terminal domain of each Rho subunit is responsible for *rut* engagement (Bogden et al., 1999; Dolan et al., 1990; Morgan et al., 1985). Because the OB fold regions of Rho encounter RNA first, they are often referred to as Rho's "primary" binding site in the literature (Galluppi and Richardson, 1980).

Following *rut* recognition, the central pore of the Rho ring next engages the nascent RNA transcript. During this process, Rho converts from an open ring state into a closed-ring form (**Figure 1.2B**), locking the RNA strand into a so-called "secondary" RNA binding site, which is formed by two conserved sequence elements known as the Q and R loops that emanate from a RecA-family ATPase fold (Miwa et al., 1995). This conformational change rearranges residues in the Rho ATP binding pockets into a hydrolysis-competent state (Thomsen and Berger, 2009); the molecular triggers for this rearrangement is explored in work described in Chapter 2 of this thesis. As with the primary site, Rho's secondary site also exhibits a preference for pyrimidine-rich nucleic acids (Richardson, 1982), a topic investigated further in Chapter 3. Upon RNA encirclement, Rho appears to maintain its primary site contacts with the *rut* sequence as it consumes ATP to translocate 5′ to 3′ along the RNA strand, a process known as tethered tracking (Galluppi and Richardson, 1980; Gocheva et al., 2015; Koslover et al., 2012a; Steinmetz and Platt, 1994).

The current model for how ATP hydrolysis by Rho drives directional translocation of RNA is based on structural studies and insights gleaned from studies of other ATP-dependent molecular machines. Rho's RecA ATPase domain comprises a highly conserved ASCE ("Additional Strand Conserved E") family nucleotide binding fold (Skordalakes and Berger, 2003), and is particularly similar to the catalytic regions of the F_1-ATPase (Abrahams et al., 1994; Dombroski and Platt, 1988; Walker et al., 1982). Key functional components of this fold (**Figure 1.2C**) include a Walker-A motif (also known as the "P-loop"), which both binds the phosphates of ATP and coordinates the Mg^{2+} ion, and the Walker-B motif, which further assists in coordination of Mg^{2+}. Also contained within this fold are a catalytic glutamate, which positions a water molecular for nucleophilic attack on the γ-phosphate, and a residue termed an "arginine finger" (contributed by an adjacent protomer), which assists with hydrolysis and conformationally communicates the hydrolysis status of ATP to the neighboring subunit. In a fashion similar to motor proteins such as the F_1-ATPase and the papillomavirus E1 DNA helicase, ATP hydrolysis by Rho is thought to proceed in a sequential, rotary manner, with changes in the status of ATP coupled to movement of the loops that engage RNA within Rho's secondary site (Enemark and Joshua-Tor, 2006; Noji et al.,

1997; Thomsen and Berger, 2009). A comparison of translocation intermediate structures of Rho (which moves 5'→3') and E1 (which runs 3'→5') has revealed that the firing order for ATP hydrolysis seems to be inverted in Rho relative to that of E1 (counterclockwise vs. clockwise), a feature that may explain why Rho and E1 traslocate in opposite directions (Enemark and Joshua-Tor, 2006; Thomsen and Berger, 2009). Rho's translocation mechanism is discussed extensively in Chapter 2.

Three models have emerged to account for how translocation by Rho eventually elicits termination of transcription by RNA polymerase, two of which resemble the models proposed for intrinsic termination. In the hybrid shearing model, the movement of RNA through Rho's central pore generates tension on the transcript, shearing the RNA/DNA hybrid in the transcription bubble formed within RNA polymerase. Support for this model includes the observation that Rho is able to generate a force sufficient to separate streptavidin from biotin (Schwartz et al., 2007a). In the hypertranslocation model, Rho moves towards, and eventually collides with, RNA polymerase in a fashion that causes the polymerase to move forward on the template strand before the next RNA base can be incorporated into the nascent RNA (Park and Roberts, 2006). The hybrid shearing and hypertranslocation models of termination are not mutually exclusive, and the mechanism by which Rho acts on the polymerase may actually be determined by whether the RNA:DNA hybrid in the transcription bubble is particularly stable at the moment at which Rho collides with the polymerase; e.g., a G/C-rich hybrid would favor hypertranslocation, while a less stable complex (due to an A/U-rich hybrid and/or a polymerase conformation that destabilizes the hybrid) would favor hybrid shearing (Ray-Soni et al., 2016). A third model, termed the allosteric model, posits that direct binding of Rho to RNA polymerase elicits a conformational change in the polymerase, which triggers inactivation and eventual dissociation of the nascent transcript (Epshtein et al., 2010).

Rho-dependent transcription termination is not limited to cytosine-rich sites in the 3' Untranslated Regions ("UTRs") of nascent transcripts. Indeed, Rho is also responsible for termination of transcription in the '5 UTRs of certain genes (e.g. *mgtA* and *rpoS*), where access to the corresponding *rut* elements is regulated by a Mg^{2+}-dependent riboswitch and the association of sRNAs, respectively (Hollands et al., 2012; Sedlyarova et al., 2016). Under stress conditions such as amino acid starvation, Rho can terminate transcription at many intragenic sites (termed "cryptic" *rut* sites) by a mechanism that may involve disruption of the physical tether that connects RNA polymerase to a trailing ribosome (Ruteshouser and Richardson, 1989). Rho is also required for suppression of antisense transcription (Peters et al., 2012), assists with the repression of transcription of horizontally transferred genes (Cardinale et al., 2008), and is responsible for the resolution of R-loops, which can trigger genomic instability if left unchecked (Leela et al., 2013). The diverse roles fulfilled by Rho suggest that, in addition to its role in normative termination, the enzyme is a global repressor of transcriptional events deemed to be aberrant by a bacterial cell, particularly those marked by the absence of a closely-associated ribosome on the nascent transcript.

Exogenous regulators of Rho-dependent transcription termination

Given the myriad cellular functions that are governed by Rho activity, it is perhaps unsurprising that Rho itself is subject to regulation by a wide variety of factors, including small molecule antagonists, specific nucleic acid sequences, and protein cofactors. Discussed below are the factors known to operate by directly interacting with Rho, and a summary of the insights gleaned from studies to date as to how these various factors affect Rho activity.

Inhibition of Rho by the antibiotic bicyclomycin

Although many nucleic acid helicases and protein translocases share a common ring-shaped structure, very few small molecules have been discovered that selectively inhibit this class of enzymes. The present-day dearth of selective inhibitors is likely due in part to the highly-conserved ATP binding pocket shared by such enzymes, which increases the likelihood of cross-reactivity with small molecules identified from high-throughput screens that focus on ATPase activity or motor function. One notable exception is the antibacterial agent bicyclomycin (**Figure 1.3A**), a highly specific inhibitor of Rho that was first discovered from fermentation studies of the bacterium *Streptomyces sapporonesis* (Miyoshi et al., 1972). Bicyclomycin (chemical formula $C_{12}H_{18}N_2O_7$) is a bridged diketopiperazine that is biosynthesized from leucine and isoleucine (Iseki et al., 1980); a formal synthesis of the drug was reported recently (Amatov et al., 2015). Although bicyclomycin is known to inhibit growth of gram-negative bacteria (Miyamura et al., 1972; Miyoshi et al., 1972), the drug is absorbed very poorly through the gut, which has limited its clinical utility to the prevention of traveler's diarrhea (Ericsson et al., 1985).

A crystal structure of an open-ring Rho hexamer bound to bicyclomycin has shown that the drug nestles between the ATPase domains of adjacent subunits (**Figure 1.3B**) (Skordalakes et al., 2005). Previous studies have shown that bicyclomycin is a noncompetitive inhibitor of Rho ATPase activity and a mixed-type inhibitor for RNA binding to Rho's secondary site (Magyar et al., 1996; Park et al., 1995). Although the structural work initially suggested that bicyclomycin antagonizes Rho by sterically preventing the binding of the nucleophilic water molecule that initiates ATP hydrolysis (Skordalakes et al., 2005), how bicyclomycin might interact with catalytically-competent Rho states, such as those thought to accompany ATPase activity and translocation (Thomsen and Berger, 2009), has not been defined. This question will be addressed as part of Chapter 3.

Activation and inhibition of Rho by nucleic acid cofactors

It has long been unclear whether the association of nucleic acids with Rho's primary sites simply acts as a mechanism to recruit Rho to a nascent transcript, or whether primary site occupancy has additional influences on Rho activity that are critical for proper termination of transcription. It is well established that the sequence of the RNA itself has a pronounced impact on whether a transcript will be acted upon by Rho (Hart and Roberts, 1991; McSwiggen et al., 1988; Morgan et al., 1985; Schwartz et al., 2007b). In addition, the binding of pyrimidine-rich sequences to the N-terminal RNA binding domains of Rho has been demonstrated to accelerate Rho's ATPase activity

(Richardson, 1982), with pre-steady-state ATPase assays showing that the formation of a catalytically competent Rho ring is governed by a rate-limiting RNA- and ATP-dependent conformational change (Jeong et al., 2004); work described in Chapter 2 suggests that this ligand-dependent isomerization likely corresponds to ring closure by Rho. Allosteric coupling between primary site occupancy and Rho translocation activity is further supported by time-resolved nucleotide analog interference mapping studies, which have revealed that the chemical modification of nucleotides engaged by Rho's primary sites slows the rate of unwinding of an RNA:DNA duplex (Soares et al., 2014). Collectively, these findings raise the intriguing possibility that the binding of pyrimidine-rich sequences to Rho's N-terminal RNA binding domains accelerates the formation of a closed-ring and catalytically-competent state. The dependence of Rho ring closure on primary site occupancy is explored in Chapter 3.

In addition to Rho activity being controlled by the presence or absence of a *rut* within a nascent transcript, Rho can be inactivated by motifs proximal to *rut* sequences that do not impact Rho's ability to associate with a *rut*. For example, in studies of the leader region of *mgtCBR* operon from *Salmonella enterica*, it was observed that inefficient translation of the operon exposes a Rho-antagonizing RNA element (termed "RARE") that inhibits ATPase activity and translocation by Rho (Sevostyanova and Groisman, 2015). RARE consists of a single-stranded, 13-nucleotide sequence located directly 5' to the *rut* site, and has no effect on Rho when sequestered within a stem-loop or replaced with a scrambled version of the sequence (Sevostyanova and Groisman, 2015). Although it appears likely that RARE stabilizes an inactive state of Rho, it is unclear whether RARE achieves this inhibitory effect through interaction with the primary sites, the secondary site, and/or some other region of the Rho hexamer.

Inhibition of Rho by the protein regulatory factors Psu, Yaeo and Hfq

Three protein regulatory factors have been found thus far that abrogate Rho-dependent transcription termination by a direct interaction with the protein. The first to be identified is the bacteriophage P4 capsid protein Psu (Linderoth and Calendar, 1991), which binds to Rho in a manner that weakens Rho's affinity for ATP (Pani et al., 2006). Another proteinaceous factor, named YaeO in *E. coli*, sterically blocks Rho:*rut* interactions by directly associating with Rho's primary sites (Gutierrez et al., 2007). The third Rho inhibitor, the bacterial RNA chaperone Hfq, can interfere with both ATPase and transcription termination activity, and when deleted, causes an observed increase in Rho-dependent termination efficiency *in vivo* (Rabhi et al., 2011a). It is possible that the increased efficiency of Rho-dependent termination in the absence of Hfq may in part be due to a decreased occupancy of Rho-inhibiting sRNAs on their cognate mRNAs (Sedlyarova et al., 2016), as sRNA delivery is known to be aided by Hfq (Chao and Vogel, 2010; Sledjeski et al., 2001).

Activation and repression of Rho-dependent termination by the transcription factor NusG

The NusG and Spt5-family of proteins (which are bacterial and archaeal/eukaryotic homologs, respectively) are the only transcription factors that are conserved across all living organisms (Tomar and Artsimovitch, 2013). The N-terminal

domains of NusG/Spt5-type factors bind to RNA polymerase in a region proximal to the polymerase active site (Martinez-Rucobo et al., 2011), and help oppose transcriptional pausing by locking the β'-clamp region of the enzyme in a productive configuration (Hirtreiter et al., 2010; Schulz et al., 2016; Sevostyanova et al., 2011). All NusG/Spt5-family proteins also contain at least one flexibly-attached C-terminal Kyrides-Onzonis-Woese ("KOW") domain, which serves in part as a recruitment site for accessory factors (Tomar and Artsimovitch, 2013). In bacteria, the NusG KOW domain (termed the NusG "CTD") predominantly associates with two proteins that bind in a mutually exclusive fashion – the transcription terminator Rho, and the dual-purpose transcription factor / ribosomal subunit NusE (also known as S10) (Burmann et al., 2010).

Although NusG is known to assist with the Rho-dependent termination of transcripts that bear non-optimal (i.e. cytosine-poor) *rut* sequences (Peters et al., 2012), the mechanism for this activity has remained unclear. The binding of Rho to NusG may serve to increase the effective local concentration of Rho around non-optimal RNA substrates, or it may favorably affect the orientation necessary for Rho to productively engage transcripts. These recruitment and positioning-based models for NusG's role in Rho-dependent termination, which are not mutually exclusive, are consistent with a report that NusG has no impact on Rho ATPase activity *in vitro* (Nehrke et al., 1993). Alternatively, the binding of NusG could modulate the rate of one or more critical steps in Rho's catalytic cycle, such as nucleotide binding/release, ring closure, translocation rate, or processivity. This kinetics-based model is supported by the finding that a requirement for NusG in helping Rho to terminate transcription at weak *rut* sites cannot be overcome by high concentrations of Rho *in vitro* (Burns and Richardson, 1995), and by the observation that NusG activity can be supplanted by hyper-active Rho mutants (Valabhoju et al., 2016). These various models have proven challenging to test, in part due to a lack of structural information on where the NusG CTD binds to Rho and to a dearth of *in vitro* assays that allow for monitoring of individual steps in Rho's functional cycle.

Binding of the NusG-CTD to NusE, which has been visualized using NMR, connects actively transcribing RNA polymerases to trailing ribosomes and blocks the association of Rho with the NusG-CTD (Burmann et al., 2010). Since conditions that slow translation rates (e.g., amino acid starvation) also promote Rho-dependent termination at intragenic sites (Ruteshouser and Richardson, 1989), it is possible that disruption of the NusE:NusG interaction (potentially in concert with the exposure of an unusually long stretch of RNA between the polymerase and the ribosome) is a prerequisite for Rho to act upon regions of certain transcripts. How Rho accesses conditional (e.g. cryptic) *rut* sites, and whether Rho requires interactions with the NusG-CTD to elicit termination at these loci, is an open question, however.

Another intriguing regulator of transcription termination is the NusG paralog RfaH. Although the NTD of RfaH greatly resembles NusG's NGN domains and binds to RNA polymerase in the same fashion, the RfaH CTD initially folds into an all α-helical conformation (compared to the β-barrel configuration of the NusG CTD). However, upon encountering a newly-transcribed *ops* sequence (which is an 8 nucleotide, single-

stranded RNA element found in a variety of loci in bacterial genomes (Artsimovitch and Landick, 2002; Belogurov et al., 2009)), the RfaH CTD unfolds and refolds as a β-barrel with close structural similarity to the NusG CTD (Burmann et al., 2012). After refolding, the RfaH CTD is capable of binding NusE, yet it exhibits no affinity for Rho (Burmann et al., 2012). How Rho discriminates between NusG paralogs that display a high degree of both structural and primary sequence similarity is an unanswered question.

To investigate the diverse roles of the NusG CTD in regulation of bacterial transcription, I determined the crystal structure of Rho in complex with the NusG CTD. This structure reveals that NusG engages Rho's C-terminus, and modeling of a Rho/NusG/RNAP complex suggests that NusG may position Rho's central pore near the RNA exit channel of RNA polymerase. By comparing the Rho/NusG structure to the structure of a complex between NusG and the ribosome-associated factor NusE, I discovered that NusG uses two loops that protrude off the KOW domain β-barrel to engage both client proteins, a finding that explains the mutual exclusivity of the Rho and NusE interaction with NusG. Using biochemical studies, I additionally found that NusG promotes RNA- and ATP-dependent Rho ring closure in the absence of any other factors, an observation that marks the first report of a direct effect of NusG the Rho catalytic cycle. These findings, and others, are described in Chapter 4.

Concluding remarks

The present dissertation is focused on mechanistic studies aimed at understanding how the binding of three distinct regulatory factors – the small molecule bicyclomycin, pyrimidine-rich nucleic acids, and the transcription factor NusG – control Rho-dependent transcription termination. Intriguingly, each of these factors acts in part by either promoting or repressing the interconversion of Rho between open-ring (loading competent) and closed-ring (translocation competent) states. The studies outlined here not only delineate the diverse mechanisms by which transcription termination is regulated in bacteria, but also exemplify how molecular machines can be controlled by the binding of dissociable regulatory factors that either accelerate or hinder the formation of a catalytically-active macromolecular assembly.

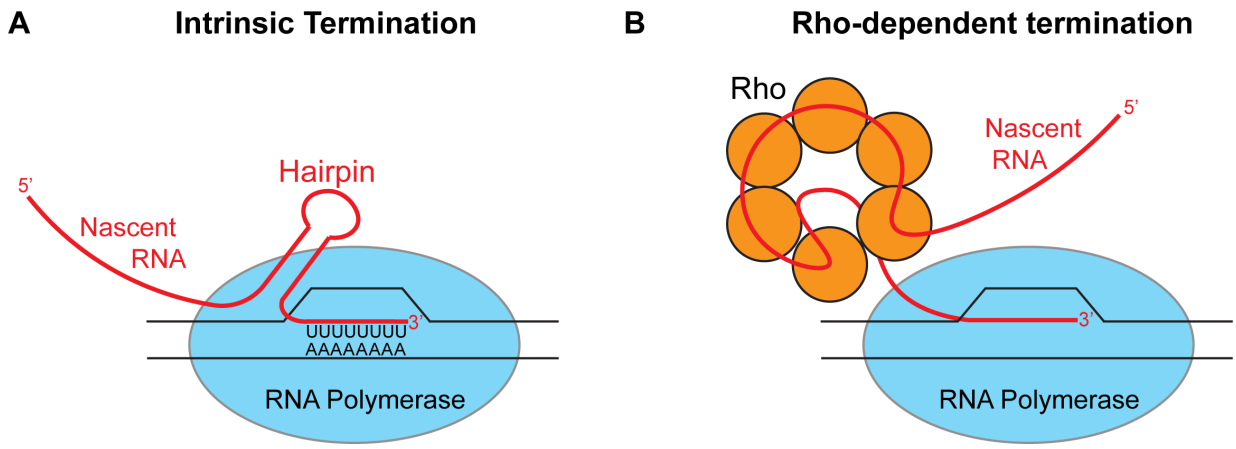

Figure 1.1: Transcription termination mechanisms in bacteria. **(A)** Intrinsic termination is driven by two adjacent motifs in the RNA strand – a poly-Uridine tract, which weakens the RNA:DNA hybrid and elicits pausing by RNA polymerase, and a hairpin, the folding of which is required for termination of transcription. (B) Rho-dependent termination is initiated by Rho's engagement of a *rut* sequence within a nascent transcript, and driven by subsequent ATP-dependent translocation of Rho towards the polymerase.

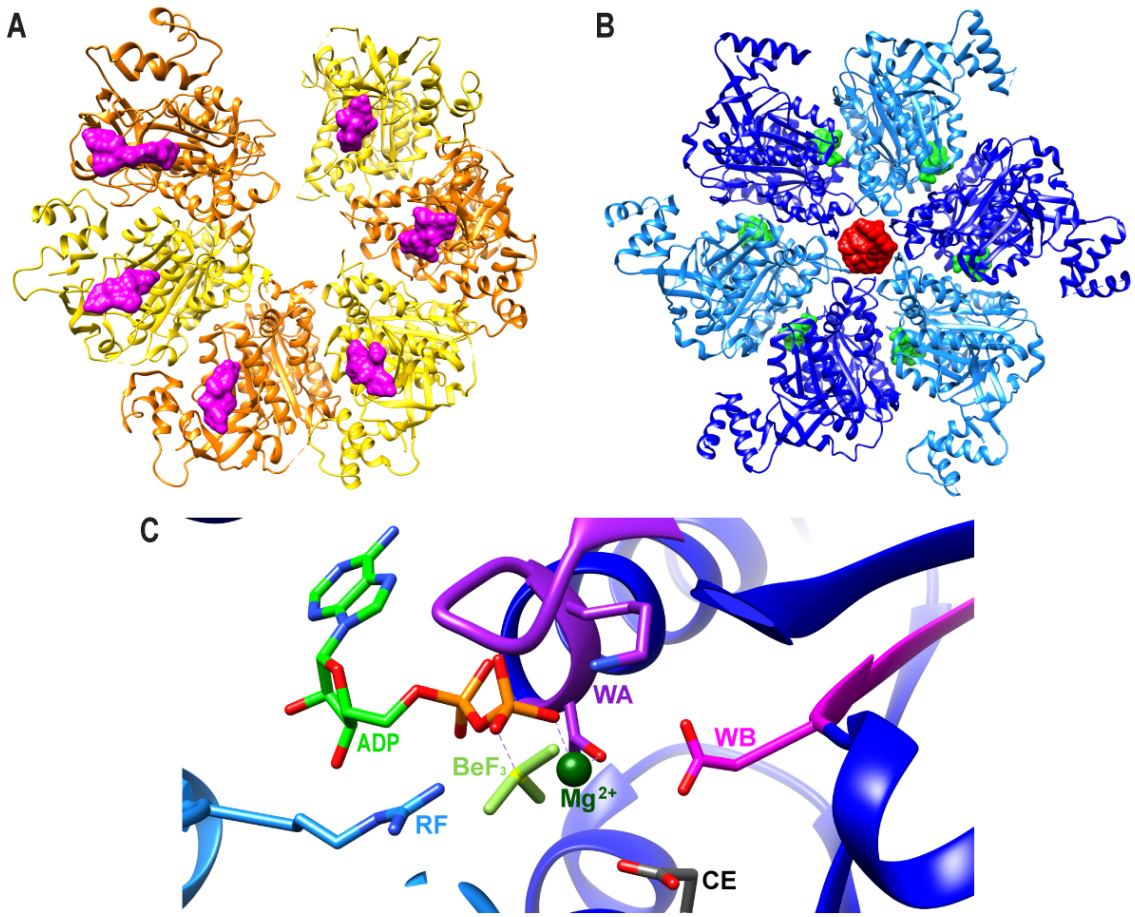

Figure 1.2: Rho is an ATP-dependent RNA helicase. **(A)** Crystal structure of an open-ring configuration of Rho (PDB ID 1XPO (Skordalakes et al., 2005)). Rho subunits are alternatingly colored orange and yellow, and primary site-bound RNA is represented in magenta. **(B)** Crystal structure of Rho in a closed-ring and translocation-competent state (PDB ID 3ICE (Thomsen and Berger, 2009)). Rho subunits are alternatingly colored light and dark blue, secondary-site bound RNA in red, and ATP in green. **(C)** A close-up view of an ATP binding site in a closed-ring Rho structure (PDB ID 3ICE (Thomsen and Berger, 2009)); WA = Walker-A motif, WB = Walker-B motif, CE = Catalytic Glutamate, and RF = Arginine Finger.

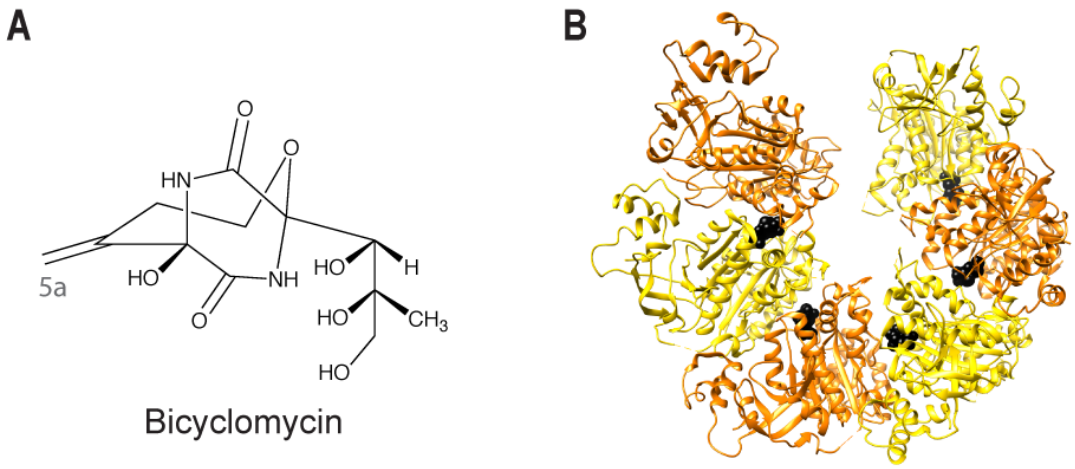

Figure 1.3: Bicyclomycin is an inhibitor of the Rho helicase. **(A)** Chemical structure of bicyclomycin. **(B)** Crystal structure of bicyclomycin-bound Rho (PDB ID: 1XPO (Skordalakes et al., 2005)). Rho subunits are alternatingly colored orange and yellow, and bicyclomycin is represented in black.

Chapter 2: Molecular mechanisms of substrate-controlled ring dynamics and substepping in a nucleic acid-dependent hexameric motor

(Portions of this chapter have been reproduced from Thomsen, N.D., Lawson M.R., Witkowsky L.B., Qu S., and Berger J.M., PNAS 2016. The work described in this chapter was a collaborative effort that I contributed to by assisting with the refinement of crystallographic models, the generation of figures and movies, and editing of the text).

Introduction

Hexameric helicases and translocases are motor proteins that play a central role in cellular transactions ranging from replication and repair to transcriptional regulation, chromosome packaging, and proteolytic homeostasis (Delagoutte and von Hippel, 2003; Patel and Picha, 2000; Sauer and Baker, 2011; Singleton et al., 2007). Used to drive the processive and at times highly-rapid movement of extended nucleic acid or protein chains through a central pore, ring-shaped motors face several challenges to their operation. One is that certain enzymes must transition through controlled ring opening and/or subunit assembly events to allow long, polymeric substrates that lack freely-accessible ends access to interior motor elements (Ahnert et al., 2000; Davey et al., 2002; O'Shea and Berger, 2014). Another is that, once loaded, the molecular plasticity inherent to these assemblies must be harnessed to precisely coordinate ATP binding and hydrolysis between multiple subunits for powering substrate translocation, while at the same time alternating between tight and loose grips on the substrate to allow for processive movement. The substrate-dependent molecular rearrangements that underpin ring dynamics during these events remain poorly understood not only for hexameric helicases and translocases, but for related ring-shaped switches as well.

The *E. coli* Rho transcription termination factor is a well-established model system for understanding hexameric translocase and helicase function (Boudvillain et al., 2010; Richardson, 2002). During termination, Rho uses a cytosine-specific RNA binding domain appended to the N-terminus of a RecA-type ATPase fold (Bear et al., 1985; Dombroski and Platt, 1988) to bind nascent RNA transcripts at cytosine-rich sequences (known as Rho utilization (*rut*) sites) (Chen and Richardson, 1987; Peters et al., 2012). Once loaded, Rho consumes ATP to translocate 5'→3' toward a paused RNA polymerase, eventually promoting transcription bubble collapse and RNA release (Brennan et al., 1987; Epshtein et al., 2010; Koslover et al., 2012b; Lau et al., 1983; Morgan et al., 1983; Park and Roberts, 2006).

Structural studies have provided insights into the mechanics of RNA loading by Rho, imaging the helicase in both closed-circular and notched/lockwasher-shaped states (Bear et al., 1988; Gogol et al., 1991; Skordalakes and Berger, 2003, 2006; Yu et al., 2000). These findings have suggested that Rho monomers assemble into a pliant hexamer that can spontaneously open to allow nucleic acid entry into the translocation pore (Richardson, 2003; Skordalakes and Berger, 2003; Yu et al., 2000), a property that has been observed in other hexameric helicases such as the eukaryotic MCM2-7 complex (Bochman and Schwacha, 2010; Costa et al., 2011; Lyubimov et al., 2012). By

comparison, the structure of a closed-ring Rho complex bound to single-stranded RNA and the ATP-mimetic ADP•BeF_3 has provided a picture of a prospective translocation intermediate of the motor (Thomsen and Berger, 2009). This work, combined with that of a prior landmark model for the human papillomavirus E1 helicase bound to single-stranded DNA and ADP (Enemark and Joshua-Tor, 2006) (as well as more recent work on the replicative helicase, DnaB (Itsathitphaisarn et al., 2012)), has led to two general principles for understanding hexameric helicase function. One is that, as proposed first for the T7 gp4 protein (Singleton et al., 2000), hexameric helicases likely employ a sequential, rotary ATPase mechanism similar to that used by the F1 ATPase to power translocation (Abrahams et al., 1994). Another is that both the relative order of hydrolysis between subunits and/or differences in subunit orientation within a motor can determine the polarity of nucleic acid translocation (Costa et al., 2014; Enemark and Joshua-Tor, 2006; Rothenberg et al., 2007; Thomsen and Berger, 2009).

Despite present-day insights, there remain many unanswered questions regarding the loading and translocation mechanism of Rho. For example, there has been no systematic assessment of the extent to which the various crystallographic states observed thus far correlate with the predominant structures adopted by Rho in solution. The relative roles that ATP and RNA play in promoting ring opening *vs*. ring closing are similarly unresolved. At the same time, the range and types of movements that accompany ATP turnover events to support RNA translocation have yet to be fully defined. Likewise, it is unclear how the physical nature of the nucleic acid substrate itself – its chemical identity and base composition – feeds back to ATPase centers to influence motor activity. Answers to these questions are needed to understand fundamental aspects of hexameric helicase loading and translocation not just by Rho, but by hexameric motor proteins in general.

To better define the physical basis by which substrate-coupled conformational changes underpin hexameric motor mechanism, we conducted both solution and crystallographic structural studies of *E. coli* Rho with the goal of characterizing novel conformational intermediates in conjunction with different RNA sequences and nucleotide-binding modes. Small-angle X-ray scattering (SAXS) confirms that the current open- and closed-ring models imaged crystallographically correspond to the dominant states adopted in solution in the absence and presence of substrates, respectively, and further establish that Rho ring closure occurs only when both RNA and ATP are present. Three new crystal structures of Rho (including an experimentally-phased structure at higher resolution than that determined previously) reveal new ring rearrangements accompanying ATP cycling, which together with comparisons to other oligomeric RecA-family motors, uncover a common allosteric linkage that physically connects nucleic-acid binding with ATPase activation in members of this widespread motor family. Together, our data highlight unexpected sub-steps in the Rho loading/translocation cycle, and help explain previous research implicating RNA sequence as a regulator of Rho function in transcription termination (Richardson, 1982; Schwartz et al., 2009; Soares et al., 2014; Wang and von Hippel, 1993).

RESULTS

Both RNA and ATP are required to drive Rho ring closure

To probe how Rho physically and differentially responds to the binding of substrates such as nucleotide and RNA, we first used small angle X-ray scattering (SAXS) to verify whether the available crystal structures of the helicase correspond to major or minor population states in solution (**Figure 2.1a-b and Figure S2.1a-b**) (Putnam et al., 2007). SAXS is an ideal technique for analyzing the ensemble of states accessed in solution by macromolecular assemblies with conformational or configurational flexibility (Rambo and Tainer, 2013). SAXS data were initially collected on Rho in the absence of ligand, and on Rho mixed with one molar equivalent per hexamer of a poly-pyrimidine RNA (rU_{12}) and 2 mM ADP•BeF_3, a non-hydrolyzable ATP analog. Consistent with size-exclusion chromatography (**Figure S2.1c-d**), SAXS curves for Rho exhibited no sign of aggregation under any of the conditions tested, and distance distribution ($P_{(r)}$) profiles, Guinier analysis, and Kratky plots also were consistent with a well-behaved, hexameric particle (**Figure S2.2 and Table S2.1**). Notably, data collected in the absence and presence of both RNA and ADP•BeF_3 proved to be extremely close matches to the theoretical curves calculated from crystal structures of models of open and closed-ring states of Rho, respectively (**Figure 2.1a-b**). Although our theoretical models do not fully account for inter-domain and inter-subunit flexibility (which would require polymer physics-based modeling of all flexible connections (Capp et al., 2014)), our data indicate that the open and closed-ring crystal structures seen previously for Rho are consistent with the conformations of native particles observed in solution.

To assess the individual contributions of RNA and nucleotide to ring closure, we next conducted SAXS experiments in either the presence or absence of RNA, and over several ADP•BeF_3 concentrations. Data collected in the absence of RNA but with 0, 0.5, 1, or 2 mM ADP•BeF_3 induced only modest changes in the scattering profile (**Figure S2.1a**). By contrast, scattering data collected in the presence of RNA exhibited distinct changes as ADP•BeF_3 concentrations were increased (**Figure S2.1b**). Data collected at intermediate concentrations of nucleotide could not be properly modeled by individual Rho crystal structures, so we used the programs FOXS and OLIGOMER to fit a two-component model consisting of both the open- and closed-ring Rho structures (**Tables S2.2-3**) (Konarev et al., 2003a; Schneidman-Duhovny et al., 2010b; Schneidman-Duhovny et al., 2013). The two-component analysis produced greatly improved fits for the intermediate states, indicating that in the absence of RNA and nucleotide, the vast majority of Rho hexamers are open in solution, and remain so even at high (2mM) ADP•BeF_3 concentrations (**Figure 2.1c**). By contrast, with the addition of RNA, nucleotide titration leads to ring closing in Rho, such that by 2mM ADP•BeF_3, the hexamer population is dominated by the closed ring state. Together, these data confirm that Rho forms an open lock-washer in solution, and further demonstrate that both RNA and nucleotide are required to promote ring closure (Ahnert et al., 2000). This RNA- and ATP-dependent conformational change in Rho may correspond to the rate-limiting step previously observed in stopped-flow kinetics measurements of ATP hydrolysis (Jeong et

al., 2004), a concept supported by the marked acceleration of both ATPase and ring closure rates in response to primary site occupancy (See Chapter 3 and (Richardson, 1982)).

A closed-ring Rho state manifests a sequential shift in ATPase status

Having established that existing crystallographic structures are suitable models for the open and closed Rho states, we next set out to determine whether different substrate-bound, closed-ring forms of Rho might exist. We postulated that different RNA sequences or nucleotide-binding conditions might influence the formation of such intermediates. Since our initial closed-ring Rho crystals were difficult to work with (hereafter referred to as $Rho^{PolyU\text{-}P1}$, in reference to the RNA sequence and space group in which the protein crystallized) (Thomsen and Berger, 2012), we set out to find a new crystal packing arrangement that could be obtained in the presence of different RNA and nucleotide conditions, but that still possessed at least one complete Rho hexamer per asymmetric unit (so as to image the particle in the absence of crystal symmetry influences). We therefore conducted a new, broader set of screens with purified Rho in the presence of RNA and nucleotide, discovering a crystal form belonging to the spacegroup $P2_1$ that satisfied these criteria (**Methods**).

From the new Rho form, we were able to experimentally phase and refine a higher resolution (2.6 Å) structure of an intact hexamer bound to seven nucleotides of a centrally bound poly-U RNA, and that displayed electron density for ADP•BeF_3•Mg^{2+} bound to all six ATP-binding sites (hereafter referred to as $Rho^{PolyU\text{-}P2_1}$) (**Figure 2.2a-d and Table 2.1**). The experimental phases and higher resolution allowed us to correct two small building errors in previous Rho structures (**Fig. S2.3**). A global alignment of the new $P2_1$ model with the original P1 state reveals that the two hexamers are highly similar structurally (**Figure 2.2a**), with a root mean squared deviation (RMSD) for all shared backbone Cα atoms of 1.55 Å. As seen previously, five subunits of the $Rho^{PolyU\text{-}P2_1}$ state form a smoothly-ascending spiral staircase that wraps around a single-stranded RNA helix; the sixth (subunit F) occupies a “transit” position midway between the uppermost (A) and bottommost (E) protomers to close the ring (**Figure 2.2b**). In addition, four classes of nucleotide-binding sites were again evident at the six intersubunit interfaces. Using criteria such as nucleotide B-factors (a measure of positional accuracy, or “orderedness” (Frauenfelder et al., 1979)) and coordination geometry, these sites have been classified as corresponding to either “exchange” (*E*), “ATP-bound” (*T*), “hydrolysis” (*T**), or product (*D*) states along the ATPase cycle (**Figure 2.2a**) (Thomsen and Berger, 2009).

Globally, the conformations of the ATPase sites and the sequential order of ATPase states in the new $P2_1$ Rho model initially appeared similar to those seen in the original P1 structure, with the modeled nucleotides refining to a relative B-factor distribution consistent with the formation of three tight and three weak ATP-binding environments as observed biochemically (Geiselmann and von Hippel, 1992; Kim and Patel, 1999; Kim et al., 1999; Seifried et al., 1992; Stitt, 1988; Xu et al., 2003) (**Figure S2.4**). However, inspection of the distance vectors calculated between Ca atoms for the

P1 and $P2_1$ models revealed that subunit F, as well as the RNA binding R-loops of subunits E and A, have moved with respect to one other around a vector roughly parallel to the ring axis (**Figure 2.2e**).

Because the A, F and E subunits are thought to represent the likely sites for the release of nucleotide hydrolysis products and the binding of fresh ATP within the hexamer (Thomsen and Berger, 2009), we elected to compare the nucleotide-binding regions of the two Rho models in more detail. This analysis uncovered pronounced changes in some of the nucleotide-binding sites, which in turn revealed a shift in the pattern of functional states between the P1 and $P2_1$ forms of Rho (**Figures 2.2a, 2.3**). The most striking change is that the presumptive catalytic water molecules seen in subunit A and B active sites of the initial closed ring structure now appear in the catalytic centers of subunits B and C in the new structure (**Figures 2.2a, 2.3a-c and Figure S2.5a-c**). Moreover, each of these waters also has shifted its local position from a non-ideal location in the $Rho^{PolyU-P1}$ model to now take up a direct, in-line attack configuration with the BeF_3 group of the bound nucleotide in the $Rho^{PolyU-P2_1}$ structure (**Figure 2.3b**). Repositioning of the waters is mediated at both interfaces by a conserved arginine (the "arginine-switch" (RS), R269), the catalytic glutamate (CE) residue (E211) that polarizes the water to assist hydrolysis, and a backbone carbonyl (G337) of the adjacent protomer. The glycine-mediated contact was not observed in the prior $Rho^{PolyU-P1}$ structure. Consistent with biochemical studies implicating their importance to overall Rho function (Balasubramanian and Stitt, 2010; Dombroski and Platt, 1988; Miwa et al., 1995; Rabhi et al., 2011b; Wei and Richardson, 2001; Xu et al., 2002), a conserved network of ion-pairs at the B/C and C/D interfaces – which runs from the RNA-binding pore to the ATPase center – appears to help orient amino acids responsible for in-line placement of the attacking water (**Figure S2.6**). Overall, the new placement for the catalytic waters in $Rho^{PolyU-P2_1}$, together with other changes in the relative orientations of catalytic important amino acids (**Figure 2.3a-f and Supplementary Discussion**), supports the proposed existence of hydrolysis-competent active sites in the closed-ring Rho states (Thomsen and Berger, 2009). This shift also advances the relative position of the two *T** states by one subunit toward the lone *E*-state, and adds a second *D*-state compared to the $Rho^{PolyU-P1}$ structure determined previously (**Figure 2.2a**).

Inter-subunit relationships in the nucleotide-exchange region mirror those of an open-ring state

A key supposition of the rotary cycling model based on the $Rho^{PolyU-P1}$ crystal form has been that the E/F interface, which is relatively open and solvent exposed, corresponds to a nucleotide exchangable (*E*) site where substrate can bind to or product can dissociate from the hexamer, respectively (Thomsen and Berger, 2009). To test this idea, we back-soaked nucleotide out of the new $Rho^{PolyU-P2_1}$ crystal form and determined the structure of the helicase by X-ray crystallography. Inspection of all active sites in the model revealed very weak and discontinuous electron density for the nucleotide-binding site of subunit E, indicating a very low occupancy for ADP•BeF_3 that precluded modeling of nucleotide into the site (nucleotide density in the other active sites was unchanged) (**Figure 2.4a-b**). Intersubunit distance vectors calculated between

the back-soaked structure (referred to hereafter as Rho^{Empty}) and the fully occupied $Rho^{PolyU\text{-}P2_1}$ structure show that the relative positions of subunits E and F shift by ~1.7 Å with respect to one another, but that the rest of the subunits in the hexamer undergo only minor positional changes (**Figure 2.4c**). A comparison of the E/F interface in the presence and absence of ADP•BeF_3 reveals an increased separation of the two subunits that moves the Arg-finger out of the active site and that buries ~120$Å^2$ less total surface area at the interface (**Figure 2.4d**). Together, these observations further confirm the proposition that the E/F active site corresponds to a nucleotide-exchange point, and show that product release coincides with an increased opening of the subunit interface at this catalytic center. Interestingly, upon the disappearance of nucleotide from the *E*-state, the relative B-factors for the nucleotide in the active site of subunit F also increased relative to other Rho structures (**Figure S2.4**), suggesting that product release may cooperatively weaken the association of nucleotide in the immediately adjacent *D*-state.

Inspection of the E/F active site shows that the *E*-state is unique within the closed-ring Rho hexamer in adopting a relative intersubunit orientation markedly different from that of the other protomers. This altered configuration arises due to a transition of subunit F (which is bounded by subunits A and E, respectively) from the top step of the helical RNA staircase to the bottom (**Figure 2.2b**). Subunit F is also the only protomer that does not contact RNA in any closed-ring structure determined so far, and its interface with subunit E in the Rho^{Empty} structure represents a completely substrate-free component of the catalytic cycle. Given these considerations, we compared the intersubunit positions of the E/F protomer conformation of the back-soaked model with those of a completely RNA-free, open-ring Rho structure of Rho determined previously. Interestingly, the intersubunit orientation of motor domains in the open-ring structure (PDB: 1PVO), which all assemble with a left-handed helical pitch, proved to be a very close match to the orientation of the E/F protomer pair in both the Rho^{Empty} and $Rho^{PolyU\text{-}P2_1}$ structures, but distinct from other subunit dimers in the RNA-bound models (which adopt a right-handed pitch) (**Figure 2.4e**). This congruency indicates that, upon ATPase product and RNA release, a single interface in the Rho hexamer relaxes to an energetically favorable conformation present in the open-ring Rho structure. Therefore, the nucleotide-dependent structural transition, which promotes ring opening and closing during RNA loading, also plays a direct role in catalytic cycling.

Sequence and ATPase state-specific differences in RNA binding by Rho

The ATPase activity of *E. coli* Rho has been well-established to depend not just on RNA binding (Lowery-Goldhammer and Richardson, 1974), but also on RNA sequence, with pyrimidine-rich RNAs providing a greater increase in ATP turnover compared to purine-rich substrates (Richardson, 1982; Wang and von Hippel, 1993). To better understand this dependency, we crystallized and determined the structure of Rho in complex with a poly-purine RNA (rA_{12}) and ADP•BeF_3 (hereafter referred to as Rho^{PolyA}), and compared it to the other Rho structures determined thus far (**Figure 2.5**). Notably, the Rho•poly-A complex co-crystallized in the same $P2_1$ space group and unit cell as the new closed-ring Rho model bound to poly-U. The poly-A substrate similarly forms an overwound, single-stranded RNA helix in the Rho pore that is highly similar to

that formed by poly-U (**Figure 2.5 and Figure S2.7**), despite the extra bulk afforded by the additional five-member ring on the purine bases. This finding indicates that the tight coiling of RNA seen in the original Rho model is not an effect of RNA sequence or crystal packing, but likely reflects the standard manner by which the helicase encircles substrate during translocation.

Close inspection of the poly-A Rho model with its poly-U counterparts reveals notable differences in the specifics of RNA binding (**Figure 2.5 and Supplementary Discussion**). Interestingly, these differences are evident not just between purine- and pyrimidine-bound Rho hexamers, but also between the different poly-U models imaged thus far (**Figure 2.5a-c**). In particular, the $Rho^{PolyU\text{-}P2_1}$ and Rho^{PolyA} structures lack electron density for the 5'-most base present in original $Rho^{PolyU\text{-}P1}$ structure, and instead manifest new density for two additional bases at the 3'-end of the RNA (**Figure 2.5 and Figure S2.7b-g**). However, the bound RNAs also show conformational differences between these three structures, with the additional 3' nucleotides visible in the poly-A bound Rho model extending in a straight line from the helicase pore, but in the new poly-U structures folding back into a tight kink that deviates from the interior RNA spiral configuration but maintains base-stacking interactions (**Figure 5 and Figure S2.7**). The U-U kink seen in the poly-U models allows for a network of contacts to be formed between the ribose 2'-OH and the uracil O2 atom of the penultimate nucleotide, a newly evident water molecule, and Lys326 of both subunits B and C (the latter of which alters the contact that had seen previously between this invariant amino acid and the backbone phosphate of RNA in the original $Rho^{PolyU\text{-}P1}$ hexamer) (**Figure 2.5a-c and Figure S2.8**). Interestingly, in analyzing the Rho^{Empty} structure, we observed electron density for the additional, kinked 3' bases of RNA similar to that present in $Rho^{PolyU\text{-}P2_1}$; however, this model also exhibited density for the 5'-most base seen in $Rho^{PolyU\text{-}P1}$ (**Figure 2.5b and Figure S2.7b-d**). Thus, upon loss of nucleotide and opening of the *E*-state active site, the poly-U RNA bound to Rho^{Empty} appears to adopt a hybrid conformation between the original $Rho^{PolyU\text{-}P1}$ crystal form and the new Rho-$P2_1$ crystal form.

A conserved allosteric helix coordinates crosstalk between subunits and their catalytic centers

In considering how changes in the ATPase centers of our closed-ring Rho models might be coupled to RNA binding, we used the higher-resolution structures determined here to look more closely at the linkages that mediate intersubunit contacts between the nucleic acid and nucleotide binding sites. By co-aligning all six pairwise interfaces of the $Rho^{PolyU\text{-}P2_1}$ model, using the ATP-binding P-loop as an reference point, we found that a single α-helix (Rα1), which contains residues responsible for forming the salt-bridge network linking the RNA binding pore to the ATPase center, undergoes a large range of previously-unnoticed hinge movements (**Figure 2.6a and Figure S2.6**). The N-terminus of this helix (which corresponds to the "R-loop" of Rho and the "L2"-loop of RecA helicases such as DnaB and T7 gp4) binds RNA via K326 (**Figure 2.5**), while the C-terminus of this helix contains backbone carbonyls that interact directly with the catalytic water molecule and the arginine-based γ-phosphate sensors in the two T*

states (**Figures 2.3b and 2.6b**). This movement appears to be caused by an intersubunit levering motion of ~10Å at the helix's N-terminus, which alters the interactions in an ordered manner that correlates with the ATPase state of each subunit (**Figure 2.6b and Movie S2.1**). This observation suggests that R-loop mediated RNA binding, aided by interactions within an intersubunit allosteric network, not only positions the catalytic glutamate via the arginine switch residue as described previously (Thomsen and Berger, 2009), but also directly controls the position of the arginine finger, arginine valve, and catalytic water by coupled movements with the associated Rα1 helix.

In conducting this analysis, we note that the α-helix corresponding to Rα1 is a conserved feature of the RecA fold, and is present not only in Rho's closest homolog (the F_1-ATPase, (Kagawa et al., 2004)), but also in more distantly related RecA-family hexameric helicases and translocases. In all cases, the C-terminal carbonyls of the helix equivalent to Rα1 contact amino acids that interact with the γ-phosphate moiety (and catalytic waters when present), and that appear poised to help control the active-site organization of their respective hexameric motors through a similar mechanism (**Figure 2.6c-e**) (Chen et al., 2008; Itsathitphaisarn et al., 2012; Kagawa et al., 2004). These comparisons indicate that this helix, which lies downstream of the R/L2 loop region, is likely involved in allosterically coupling substrate movements to ATPase status in RecA-like motors in general.

Discussion

A possible role for ring distortion and strain in the mechanism of Rho and other hexameric motors

In the present study, we have used solution-based structural methods and X-ray crystallography to better understand how a RecA-family hexameric motor protein, Rho, couples changes in nucleotide status to conformational transitions that underpin ring loading and translocation events. Our SAXS data confirm that distinct open- and closed-ring structures are accessed prior to and following RNA engagement, respectively (**Figure 2.1a-c**), suggesting that the transition between these two states likely occurs as a relatively simple rigid body rotation of subunits with respect to each other (**Movie S2.2**). Notably, our studies indicate that at nucleotide concentrations approaching those in the cell (values between 2 and 10 mM ATP have been measured for *E. coli* (Albe et al., 1990; Bennett et al., 2009; Lowry et al., 1971)), Rho spontaneously and preferentially occupies an open-ring state, a form that is well-suited for engaging nascent RNA transcripts. A need for contacts engendered by both substrate nucleic acid and ATP binding to cooperatively induce ring closure is consistent with the idea that formation of a translocation-competent state may capture and harness strain as a motive force to help promote nucleotide cycling. Consistent with this model, the formation of a relaxed, left-handed inter-subunit orientation characteristic of the open-ring structure is seen only in the product release state, which also releases its grip on RNA to reset the RNA binding cycle (**Figure 2.4e**).

Strain imparted by substrate-induced conformational changes may underlie ATP cycling in other, more distantly-related hexameric helicases. For example, recent structures of DnaB-family helicases have shown that the ATPase domains of these hexameric motors adopt an uncoupled, planar conformation in the absence of nucleic acid substrate (Bailey et al., 2007), but form a steep right-handed helical conformation with a single open interface when DNA and GDP•AlF_4 are present (Bailey et al., 2007; Itsathitphaisarn et al., 2012). Electron microscopy data on the eukaryotic minichromosome maintenance proteins (MCMs) have revealed that a left-handed open-ring conformation is the default state of metazoan MCM hexamers (Lyubimov et al., 2012), and that the Cdc45•GINS complex, along with ATP and DNA are needed to drive the formation of a stable closed ring with an internal, right-handed pitch (Costa et al., 2011; Costa et al., 2014). Together, these studies support the idea that nucleic acids cooperatively activate ATP hydrolysis in many ring-shaped RNA and DNA translocases by serving as a template for the asymmetric assembly of ATPase subunits. This asymmetry in turn promotes the formation of a defined sequence and order of active site states (Enemark and Joshua-Tor, 2006; Thomsen and Berger, 2009), which allows the energetics of cooperative substrate binding in one subunit to be allosterically coupled to cooperative product release in another. One exception to this trend may be the papilloma virus E1 helicase, which has been found to adopt an asymmetric, right-handed closed ring in the absence of ATP and DNA (Sanders et al., 2007).

RNA sequence dependencies of the Rho functional cycle

In addition to the insights afforded into ring dynamics and the establishment of a sequential ATPase order, the structures presented here also help to explain some of the known effects that RNA sequence has on Rho's ATPase activity. In general, the molecular basis for substrate-specific sequence dependencies in hexameric helicases and translocases has remained relatively ill-defined. Some ring-shaped helicases, such as T7 gp4, decelerate when confronted with GC-rich duplexes (Johnson et al., 2007), indicating that motor function can be impacted by substrate stability. For Rho, transcription termination activity is stimulated by cytosine-rich sequences (Peters et al., 2012), due in part to the recognition requirements of its primary RNA (*rut*)-binding sites (Dombroski and Platt, 1988; Richardson, 1982), and also to the ability of primary site occupancy to promote Rho ring closure (see Chapter 3). However, Rho's RNA-stimulated ATPase and translocation activities also appear to be directly affected by pyrimidine-specific sequence preferences independent of primary site interactions (Richardson, 1982; Schwartz et al., 2009; Wang and von Hippel, 1993).

The differences in secondary site RNA binding seen here between poly-U and poly-A bound forms of Rho (along with the preferential ring closure capability of poly-U over poly-A described in Chapter 3) indicate that sequence discrimination in the helicase may in part arise from interrogation of RNA sequence at the point where the nucleic acid threads into the ring. In particular, the 3' kink seen in the new $Rho^{PolyU\text{-}P2_1}$ structure allows a ribose 2'-OH and two pyrimidine bases to interact with an invariant lysine, Lys326, on two different R-loops. These contacts are absent in the 3' RNA region when Rho is bound to poly-A, and in the original $Rho^{PolyU\text{-}P1}$ model (where the 3' nucleotides are unstructured, see **Figure 2.5**). Intriguingly, the lysines involved in binding the U-U kink in $Rho^{PolyU\text{-}P2_1}$ project from subunits B and C, which also manifest the two *T** states observed in the hexamer (**Figure 2.2a**). These interactions suggest that Lys326, in addition to its role neutralizing the backbone charge of RNA bound in Rho's translocation pore (Thomsen and Berger, 2009), may also serve as a sensor for RNA sequence to help properly organize the B- and C-subunit active sites (**Figure 2.3a**). Such a function would explain why pyrimidine-rich *rut* sites are particularly well-suited as secondary site ligands to promote nucleotide hydrolysis by Rho, in that they provide an RNA structural motif that helps set up an idealized ATPase configuration.

The differences in the 3' RNA structure seen here for the purine and pyrimidine substrates likewise have implications for understanding how RNA sequence is coupled to step size. For instance, using nucleotide analog interference mapping (NAIM) to insert dNTPs into RNA, Rho has been found to be inhibited by the presence of deoxyribose at intervals of every seven nucleotides, with the level of inhibition dependent on the local sequence environment (Schwartz et al., 2009). Our structures reveal that pyrimidine, but not purine substrates can specifically interact with Lys236 of Rho by forming a specific kink that involves the 2'OH and pyrimidine base of the incoming 3'-RNA end (**Figure 2.5**). Given that movements of the R-loop are directly coupled to changes in the allosteric network that couples ATPase status with RNA-binding (**Figure 2.6 and Figure S2.6**) (Thomsen and Berger, 2009), the ability to adopt a kink by pyrimidines may help underpin the positional and sequence dependency of

the 2'-OH activation step seen to occur roughly every seven steps by NAIM (Patel, 2009; Schwartz et al., 2009).

Structural snapshots reveal translocation sub-steps in a hexameric motor protein

The present study provides the most comprehensive structural analysis to date for a hexameric helicase bound to different nucleic acid substrates and nucleotide intermediates. In comparing the original P1 and the new $P2_1$ poly-U bound models of Rho, both hexamers can be seen to possess a single *E*-state active site (subunit E) that sits adjacent to a *T* (ATP-bound) state ATPase center (subunit D). The two structures diverge, however, in that the $Rho^{PolyU\text{-}P1}$ form contains a second *T*-state active site (subunit C) followed by two hydrolysis-like *T** states (subunits B and A), whereas the $Rho^{PolyU\text{-}P2_1}$ model is shifted, possessing a lone *T*-state followed immediately by two *T** catalytic centers (**Figure 2.2**). As a consequence, the $Rho^{PolyU\text{-}P1}$ structure contains only a single product (*D*) state (subunit F), whereas the corresponding $Rho^{PolyU\text{-}P2_1}$ structure contains two (subunits F and A). This observation is consistent with prior studies showing that there exists a 2-4 subunit delay in ATP hydrolysis and product release in Rho and related RecA-family hexameric helicases such as T7 gp4 (see (Liao et al., 2005a) and **Supplementary Discussion**).

Although the *E*-state active site (subunit E) remains the most solvent-accessible ATPase center in both crystal forms of Rho, changes in the relative disposition of key functional amino acids are nonetheless evident for this region between the P1 and $P2_1$ structures. In particular, the subunit E active site of the new $Rho^{PolyU\text{-}P2_1}$ model exhibits a dramatically displaced and off-line catalytic glutamate, but forms a new contact between the arginine-finger and the BeF_3 moiety; by contrast, in the original $Rho^{PolyU\text{-}P1}$ structure, the correct positioning of the catalytic glutamate is maintained but the arginine-finger/nucleotide contact is broken (**Figure 2.3d-e**). Concomitant with these changes, two bases at the 3' end of the bound RNA become ordered and coordinated in the $Rho^{PolyU\text{-}P2_1}$ structure, whereas in the $Rho^{PolyU\text{-}P1}$ model, only a single extra base is apparent at the 5' RNA end (**Figure 2.5**). Interestingly, back-soaking of the new poly-U Rho crystal form, which results in nucleotide dissociation from the E-subunit interface, produces a hybrid state in which RNA density is evident at both 5' and 3' ends.

Taken together, the suite of Rho structures not only recapitulates the movements expected for 5'→3' translocation along RNA, but also highlights the existence of sub-steps within the context of a sequential ATPase mechanism (**Figure 2.7 and Movie S2.3**). In this scheme, formation of an empty ATPase center would promote downstream (3') contacts between Rho and its substrate, while maintaining 5' interactions leftover from the prior round of hydrolysis. ATP binding would next disrupt the 5' RNA contact, initiating a new translocation cycle to pull the 3' end of the bound RNA into the motor pore. Finally, product formation would complete the translocation step, helping to shepherd the 5' RNA end out of the helicase, after which ADP release would reset the system by forming new contacts to the incoming 3' base.

Although conformational substeps have been observed structurally for superfamily 1 and 2 helicases and for motors such as the F1 ATPase, myosin, and

kinesin (Kull et al., 1996; Rees et al., 2012; Sablin et al., 1998; Schmidt et al., 2014), similar insights have not been forthcoming for hexameric helicase/translocase systems. Functional substeps (as well as macrosteps) are frequently observed in single molecule studies of motor proteins (Chistol et al., 2012; Yasuda et al., 2001), but the architectural underpinnings of these events often remained unresolved. The present work demonstrates the existence of distinct physical substates in a RecA-family hexameric helicase, and moreover implicates both RNA binding and different stages of the ATPase cycle in the formation and dissolution of these states. Future efforts involving high-resolution single-molecule studies and structure-guided mutagenesis will help establish the timing by which different states interconvert with each other and the role that specific, protein-mediated contacts between RNA and ATP play in controlling these transitions.

Materials and Methods

Small Angle X-ray Scattering (SAXS) data collection and analysis.

Native Rho protein was prepared in a similar manner to selenomethionine-labeled protein as described (Skordalakes and Berger, 2003). Rho protein (40 mg/mL) in size exclusion buffer (50 mM Tris 7.5, 500 mM KCl, 10% Glycerol, 1mM TCEP) was dialyzed overnight into SAXS buffer (150 mM KCl, 50 mM HEPES pH 7.5, 5 mM $MgCl_2$, 5% Glycerol). Dialyzed protein was then further diluted to 5 mg/mL in SAXS buffer containing various concentrations of RNA (rU12) or ADP•BeF_3, aliquoted into a 96 well plate, and used for data collection at room temperature using an automated system at ALS Beamline 12.3.1. Raw SAXS data were integrated, scaled, and buffer subtracted with the program OGRE (Greg Hura, Lawrence Berkeley National Laboratory). PRIMUS was used to merge data from a one second and ten second exposure and to manually remove outliers near the beam-stop in order to obtain high quality scattering curves with very little noise (Konarev et al., 2003b). Guinier analysis was performed using AUTORG (Petoukhov et al., 2007). Distance distribution functions (P(r)), maximum diameters (D_{max}) and real- and reciprocal-space radius of gyration (R_g) values were calculated with GNOM (Svergun, 1992). Structural models were generated using an open-ring Rho model (PDB: 1PVO (Skordalakes and Berger, 2003)) and the closed ring crystal structure determined here ($Rho^{PolyU\text{-}P2_1}$). Missing loops or domains were added to the structural models by superposition with other structures or subunits containing the missing residues. OLIGOMER and FOXS were used to compare theoretical scattering curves to the data and compute the relative contributions of multi-component datasets (Schneidman-Duhovny et al., 2010b; Schneidman-Duhovny et al., 2013; Svergun et al., 1995). Background subtraction as implemented in FOXS or OLIGOMER was used to account for errors in buffer composition and improve the fit of the high-resolution data.

Protein crystallization and data collection.

Rho was crystallized by mixing 250 nL of Rho complex (20 mg/mL) with 250 nL of a well solution containing 200 mM KOAc, 40% MPD and 0.5% ethyl acetate using a Mosquito Nanoliter Pipetting system (TTP Biotech), in a hanging-drop vapor diffusion format. Crystals grew within three days and reached maximum size within one week. Crystals were harvested directly by looping and plunging in liquid nitrogen. Crystals were mounted on bendable cryo-loops (Hampton Research), oriented to minimize spot overlap, and used for collecting diffraction data on Beamline 8.3.1 at the Advanced Light Source (96). Diffraction data from the new $Rho^{PolyU\text{-}P21}$ crystal form were significantly improved compared to the previous $Rho^{PolyU\text{-}P1}$ model (29), displaying lower mosaicity (~1° versus 2°), and higher resolution (2.6 Å versus 2.8 Å). Co-crystals of Rho with a short poly-adenosine RNA (rA_{12}) and ADP•BeF_3 were obtained under nearly identical conditions by substituting rA_{12} RNA for rU_{12} RNA during rho complex formation. To obtain the Rho^{Empty} crystals, we first grew $Rho^{PolyU\text{-}P21}$ crystals, and then conducted stepwise exchange of the mother liquor with an identical solution containing no ADP•BeF_3. After a series of 2-fold dilutions and equilibrations conducted over ~15 minutes, the concentration of ADP•BeF_3 in the mother liquor was reduced approximately 30-fold (further dilution caused the crystals to dissolve).

Structure solution and refinement.

All data were processed in HKL-2000 (Otwinowski and Minor, 1997). Molecular replacement/ single wavelength anomalous dispersion (MR/SAD) methods as implemented in PHENIX were used to phase the $Rho^{PolyU\text{-}P2_1}$ structure (Adams et al., 2010). Molecular replacement was conducted using a subunit from the $Rho^{PolyU\text{-}P1}$ structure. Phases from the resulting solution were used by the program PHASER to locate 92 selenium sites in the asymmetric unit and phase the structure using SAD (McCoy et al., 2007), producing a figure of merit (FOM) of 0.39 for unmodified phases. Multi-domain, six-fold non-crystallographic symmetry (NCS) averaging and phase extension into an isomorphous, low-energy remote dataset using RESOLVE produced a FOM of 0.63 (Terwilliger, 2000). These high quality maps allowed us to clearly locate bound ligands (**Figure 2.2c-d and Figure S2.6a-i**) and correct small errors in previous structures (**Figure S2.3**). The final model, which was refined to an R_{work}/R_{free} of 21.8/24.6%, currently represents the highest resolution view of a hexameric helicase bound to both nucleic acid and an ATP mimic. The Ramachandran statistics for this model are 98.17% favored, 1.79% allowed, and 0.04% outliers as reported by MOLPROBITY (Davis et al., 2007).

The nearly isomorphous Rho^{Empty} and Rho^{PolyA} structures were determined by molecular replacement, using the entire $Rho^{PolyU\text{-}P2_1}$ structure as a search model, and R_{free} was calculated using the same test set for all structures. Rigid body refinement of the N (1-126) and C (127-417) terminal domain of each individual subunit was conducted following molecular replacement for all structures. The structures were refined in PHENIX using custom bond and angle restraints for the bound ADP•BeF_3 complex, individual ($Rho^{PolyU\text{-}P2_1}$) or group (Rho^{PolyA} and Rho^{Empty}) B-factor modeling, and TLS modeling using selections obtained from the TLSMD server (Adams et al., 2010; Painter and Merritt, 2006). The Rho^{PolyA} structure was refined to 3.15 Å with an R_{work}/R_{free} of 23.3/26.5%, contains six bound ADP•BeF_3•Mg^{2+} moieties and resolves seven bases of poly-adenosine RNA (**Table 2.1 and Figure S2.5f-g**). The Ramachandran statistics for this structure are 97.48% favored, 2.52% allowed, and 0% outliers. The Rho^{PolyA} model closely aligns with both the $Rho^{PolyU\text{-}P1}$ and $Rho^{PolyU\text{-}P2_1}$ structures (RMSDs of 1.57 Å and 0.82 Å respectively for backbone C-α atoms). The Rho^{Empty} structure was refined to 3.2 Å resolution with an R_{work}/R_{free} of 23.2/26.6%, contains five clearly resolved ADP•BeF_3•Mg^{2+} complexes and resolves eight bases of poly-U RNA (**Table 2.1 Figure S2.5h-i**). The Ramachandran statistics for this structure are 97.68% preferred, 2.32% allowed, and 0% outliers. The structure displays the largest changes in overall architecture compared to the other Rho states, aligning with $Rho^{PolyU\text{-}P1}$, $Rho^{PolyU\text{-}P2_1}$and Rho^{PolyA} structures with Cα RMSDs of 2.39 Å, 0.82 Å, and 0.88 Å respectively.

Structural analysis.

All figures were prepared using PYMOL (Schrödinger). Structural superpositions were conducted with the CEALIGN PYMOL plugin (Shindyalov and Bourne, 1998). Energy minimized linear interpolations were computed using a CNS script written by the Yale Morph Server (Flores et al., 2006). To place atomic B-factors from different

structures on the same relative scale (**Figure S2.4**), atomic B-factors for each structure were divided by the average B-factor for that structure and then multiplied by the average B-factor for all structures. Distance vectors calculated between various Rho structures were calculated as described (Zhang et al., 2009). Vector length corresponds to distances between Ca atoms among different structures multiplied by a scale factor of three to assist visualization, while vector color corresponds to relative distance (dark blue: 0 Å; red: 1.5 Å or greater).

Table S2.1. Data collection, phasing and refinement.

	$Rho^{PolyU-P2_1}$	$Rho^{PolyU-P2_1}$	Rho^{PolyA}	Rho^{Empty}
Data collection				
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions				
a, *b*, *c* (Å)	68.72, 198.8, 111.8	69.00, 198.6, 111.4	68.76, 199.2, 112.2	69.09, 199.1, 111.7
α, β, γ (°)	90.0, 104.8, 90.0	90.0, 104.4, 90.0	90.0, 104.8, 90.0	90.0, 105.0, 90.0
	peak	*remote*	*remote*	*remote*
Wavelength	0.9796	1.116	1.116	1.116
Resolution (Å)	50-2.95 (3.06-2.95)	50-2.60 (2.69-2.60)	50-3.15 (3.26-3.15)	50-3.2 (3.31-3.20)
R_{sym} or R_{merge}	9.1 (46)	8.1 (63)	12.1 (60)	12.8 (58)
$I/\sigma I$	10.8 (2.4)	14 (1.8)	9.4 (2.0)	8.1 (1.8)
Completeness (%)	98.2 (97.1)	99.7 (99.1)	99.0 (99.5)	96.2 (96.4)
Redundancy	5.6 (5.2)	3.1 (3.1)	2.9 (2.9)	2.8 (2.8)
Refinement				
Resolution (Å)		47.4-2.6	49.8-3.15	49.5-3.2
No. reflections		88,380	49,848	46,060
R_{work}/R_{free}		21.8 / 24.6	23.3 / 26.5	23.2 / 26.6
No. atoms				
Protein		19,164	18,696	18,888
Ligand/ion		341	347	325
Water		271	19	15[a]
B-factors				
Protein		67.9	80.7	73.7
Ligand/ion		53.6	69.9	66.6
Water		45.2	64.5	46.6
R.M.S deviations				
Bond lengths (Å)		0.006	0.006	0.006
Bond angles (º)		0.575	0.560	0.691

Values in parenthesis correspond to the highest resolution bin

[a]Only waters involved in octahedral coordination of the active site Mg^{2+} ions were modeled

Figures

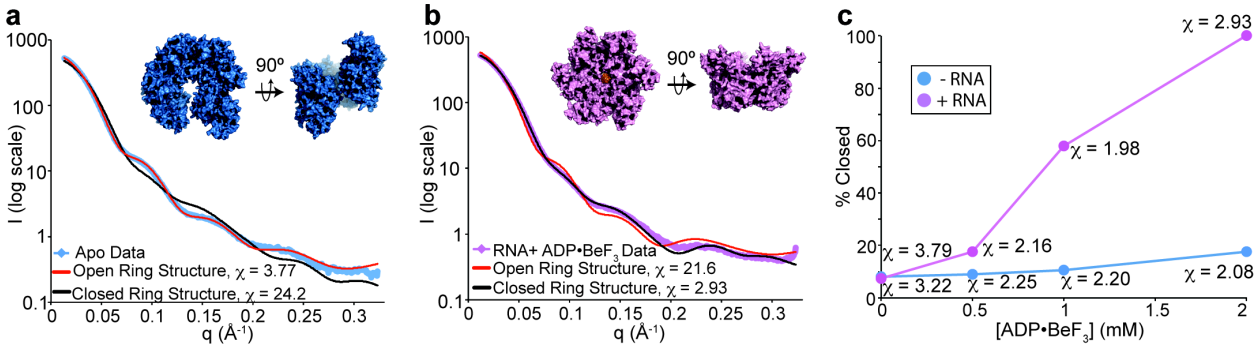

Figure 2.1. SAXS analysis of apo and substrate-bound Rho complexes. (a) Scattering curve for apo Rho in solution (blue) closely matches the theoretical curve (red) calculated from a model based on the open-ring structure (1PVO, inset (Skordalakes and Berger, 2003)), but not the theoretical curve (black) calculated from a closed ring-Rho model bound to RNA and ADP•BeF_3. (b) Scattering curve for Rho in the presence of RNA (1:1 RNA:hexamer) and 2 mM ADP•BeF_3 (pink) closely matches the theoretical curve (black) calculated from the closed ring-Rho structure bound to RNA and ADP•BeF_3 determined here but not the theoretical curve (red) calculated from a open ring Rho model (1PVO). (c) Plot of the percent-closed value obtained with FOXS *versus* ADP•BeF_3 concentration shows that both RNA and ATP are required for Rho ring closure.

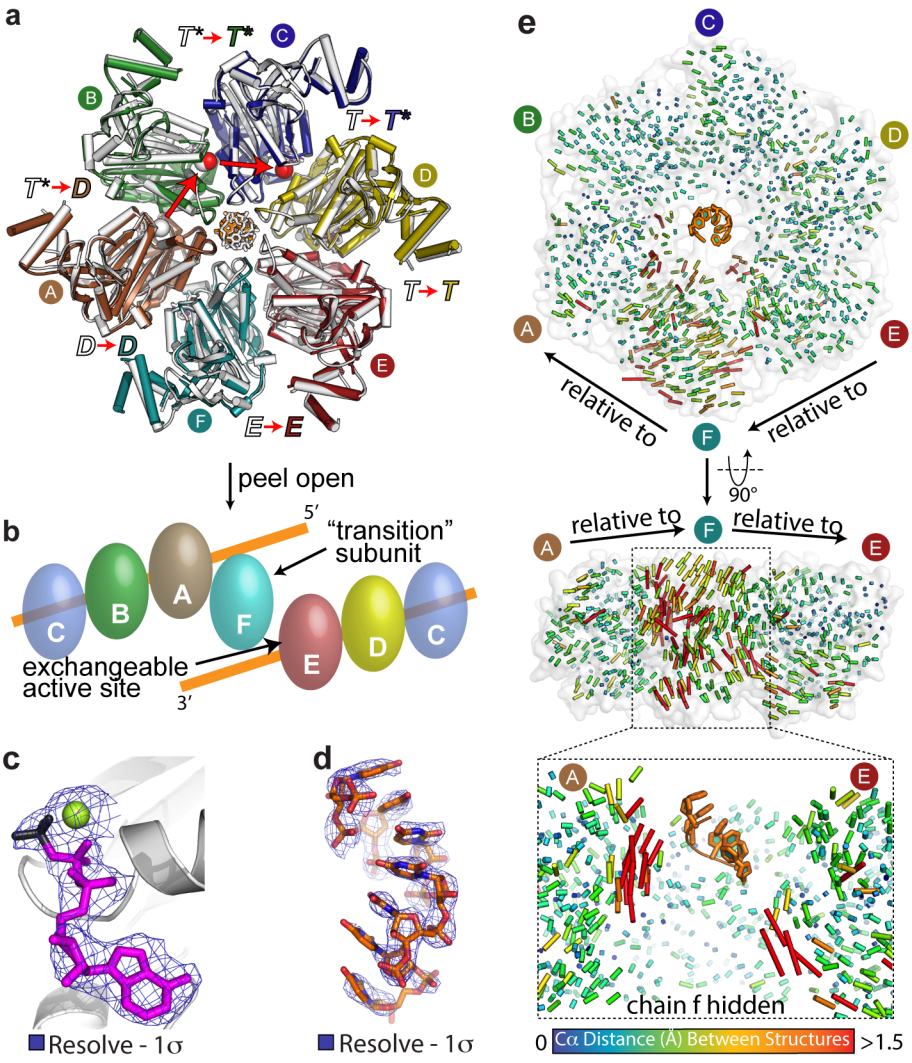

Figure 2.2. Overall structure and ligand binding in a new Rho translocation intermediate. (**a**) Superposition of a new Rho•RNA•ADP•BeF_3 structure ($Rho^{PolyU\text{-}P2_1}$, colored and labeled by subunit), with the previously-determined Rho•RNA•ADP•BeF_3 structure ($Rho^{PolyU\text{-}P1}$, white). Presumptive ATP hydrolysis states (*T* – ATP bound, *T** – ATP hydrolysis, *D* – product, *E* – exchange) are labeled in color ($Rho^{PolyU\text{-}P2_1}$) or white ($Rho^{PolyU\text{-}P1}$). (**b**) Schematic of closed-ring Rho hexamer (flattened on page) bound to RNA (one complete turn) highlighting the right-handed helical rise of the A/B/C/D/E interfaces that are bound to RNA, and the position of subunit F in transition between the top and bottom steps of the RNA binding staircase. Subunit C is shown twice to connect the two sides. (**c**) Experimental (density modified, SAD phased, $Rho^{PolyU\text{-}P2_1}$) electron density around one of the bound nucleotides (ADP – magenta, BeF_3 – black, Mg^{2+} – yellow-green). (**d**) Experimental (density modified, SAD phased, $Rho^{PolyU\text{-}P2_1}$) electron density around the bound RNA (orange). (**e**) Intersubunit distance vectors calculated independently for each intersubunit interface between the $Rho^{PolyU\text{-}P1}$ and $Rho^{PolyU\text{-}P2_1}$ structures. Vectors were calculated relative to either the clockwise-adjacent subunit (top) or to the counterclockwise-adjacent subunit (bottom) revealing that the new $P2_1$

structures differ from the original P1 model primarily in the position of subunit F, and in the Q and R-loops of subunits A and E.

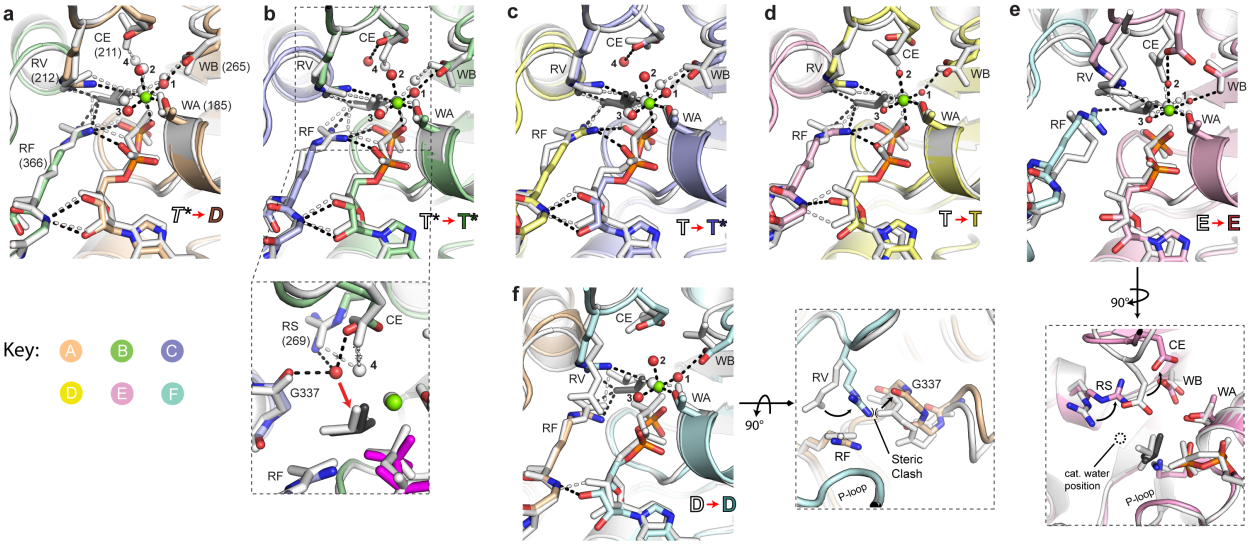

Figure 2.3. Changes in ADP•BeF_3 coordination as revealed by superpositions between six $Rho^{PolyU\text{-}P1}$ (white) and $Rho^{PolyU\text{-}P2_1}$ (colored) ATPase centers. In all subunits, changes are consistent with a partial shift in catalytic status of the six active sites around the ring. Abbreviations: WA – Walker A, WB – Walker B, RS – arginine switch, RF – arginine finger, RV – arginine valve. (**a**) Subunit A/B interface. While a water molecule is observed to associate with the catalytic glutamate (CE) at this site in $Rho^{PolyU\text{-}P1}$ (water 4), it is not observed in the new $Rho^{PolyU\text{-}P2_1}$ structure. (**b**) Subunit B/C interface. Both catalytic centers show bonding interactions (dashed lines) characteristic of the *T* and *T** states. Inset shows a zoomed-in view of the catalytic center with the RV removed (for clarity): whereas the prospective attacking water molecule in $Rho^{PolyU\text{-}P1}$ was not ideally positioned for in-line chemistry with the g-phosphate moiety (mimicked by the BeF_3 – white), the CE-associated water molecules seen in the *T** states of the new $Rho^{PolyU\text{-}P2_1}$ are positioned for ideal in-line attack (red arrow). (**c**) Subunit C/D interface. Concomitant with the disappearance of the catalytic water molecule in the A/B subunit interface of the $Rho^{PolyU\text{-}P2_1}$ structure, a catalytic water molecule appears in the C/D subunit interface compared to the $Rho^{PolyU\text{-}P1}$ structure. (**d**) Subunit D/E interface. This site retains contacts characteristic of the *T* state in both structures. The CE in $Rho^{PolyU\text{-}P2_1}$ has broken its atypical contact with the Mg^{2+} ion observed in $Rho^{PolyU\text{-}P1}$, suggesting that it is transitioning to a slightly more active arrangement. (**e**) Subunit E/F interface. The arginine finger in the $Rho^{PolyU\text{-}P1}$ *E*-state appears to be pointing out of the active site, while the arginine finger in $Rho^{PolyU\text{-}P2_1}$ clearly points into the active site and coordinates the bound BeF_3. However, the arginine valve in the $Rho^{PolyU\text{-}P2_1}$ structure is pulled slightly out of the active site, resulting in the same number of contacts between the arginine groups and the BeF_3 molecule. The catalytic glutamate in this active site is in an atypical conformation in both structures, interacting with the Mg^{2+} ion directly (P1) or indirectly *via* coordinated waters ($P2_1$). Rotated view – the catalytic glutamate and arginine switch in $Rho^{PolyU\text{-}P2_1}$ are moved out of the catalytic site in the *E*-state when compared to the $Rho^{PolyU\text{-}P1}$ structure. (**f**) Subunit F/A interface. The arginine valve in the $Rho^{PolyU\text{-}P1}$ *D*-state points out of the active site, while the arginine valve in the $Rho^{PolyU\text{-}P2_1}$ structure points into the active site and coordinates the bound BeF_3. Rotated view – the arginine valve now points into the active site due to an opening of the intersubunit

interface in $Rho^{PolyU\text{-}P2_1}$ relative to $Rho^{PolyU\text{-}P1}$ that relieves a steric clash between the guanidinium group of the arginine and G337 of the adjacent subunit.

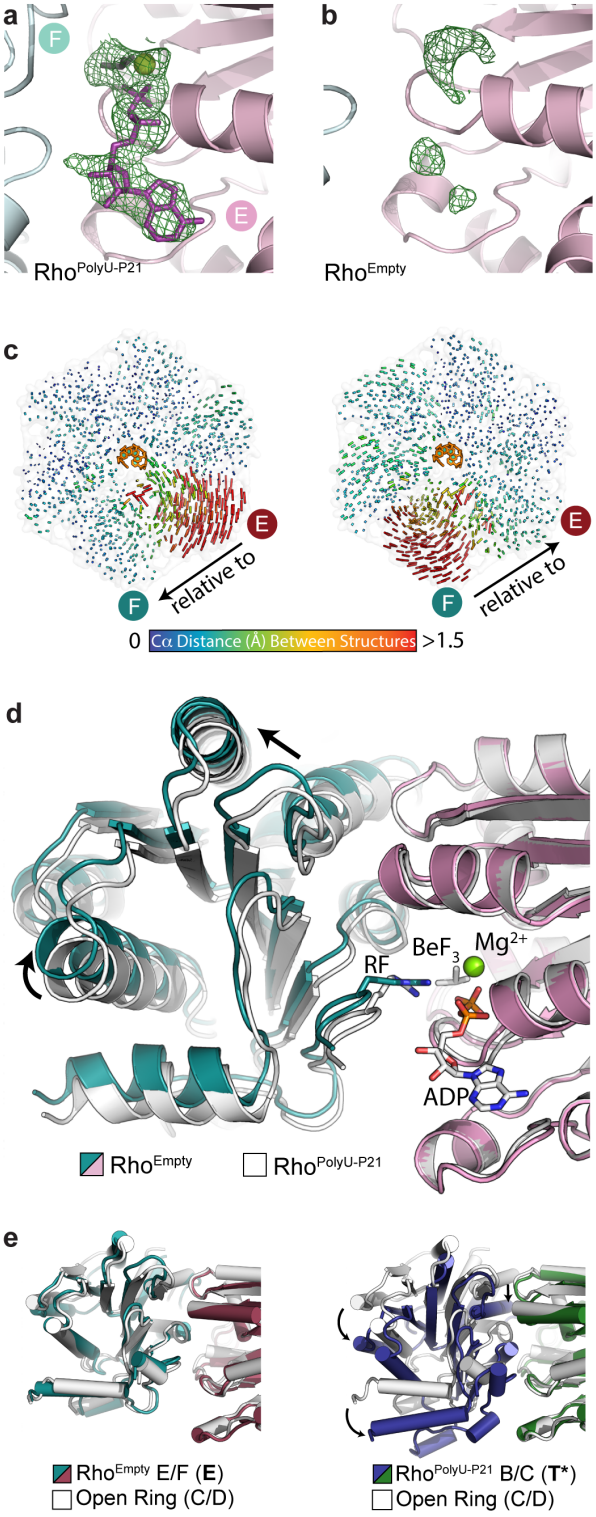

Figure 2.4. Opening of the active site in the exchange (*E*) state (subunit E) accompanies nucleotide release. (a) Simulated annealing omit map of the *E*-state in $Rho^{PolyU\text{-}P2_1}$ revealing strong density for the bound nucleotide. (b) Simulated annealing omit map of the *E*-state in the Rho^{Empty} structure reveals poor density suggesting low

nucleotide occupancy and/or order. (c) Intersubunit distance vectors calculated independently for each intersubunit interface and comparing Rho^{Empty} and $Rho^{PolyU\text{-}P2_1}$ structures. The vectors reveal that the only significant intersubunit changes between these structures are at the E/F interface. (d) Superposition of chain E from the $Rho^{PolyU\text{-}P2_1}$ and Rho^{Empty} structures reveals that nucleotide release induces a rigid body shift in chain F (arrows) that causes the active site to open. The arginine finger (RF), which also changes position, is labeled. (e) Structural alignments showing that the E/F suubnit (*E*-state) interface of Rho^{Empty} aligns almost perfectly with an open-ring intersubunit interface (left), as opposed contrast to the B/C subunit (*T**-state) subunit interface of $Rho^{PolyU\text{-}P2_1}$, which has adopted a right-handed orientation (right, red arrows).

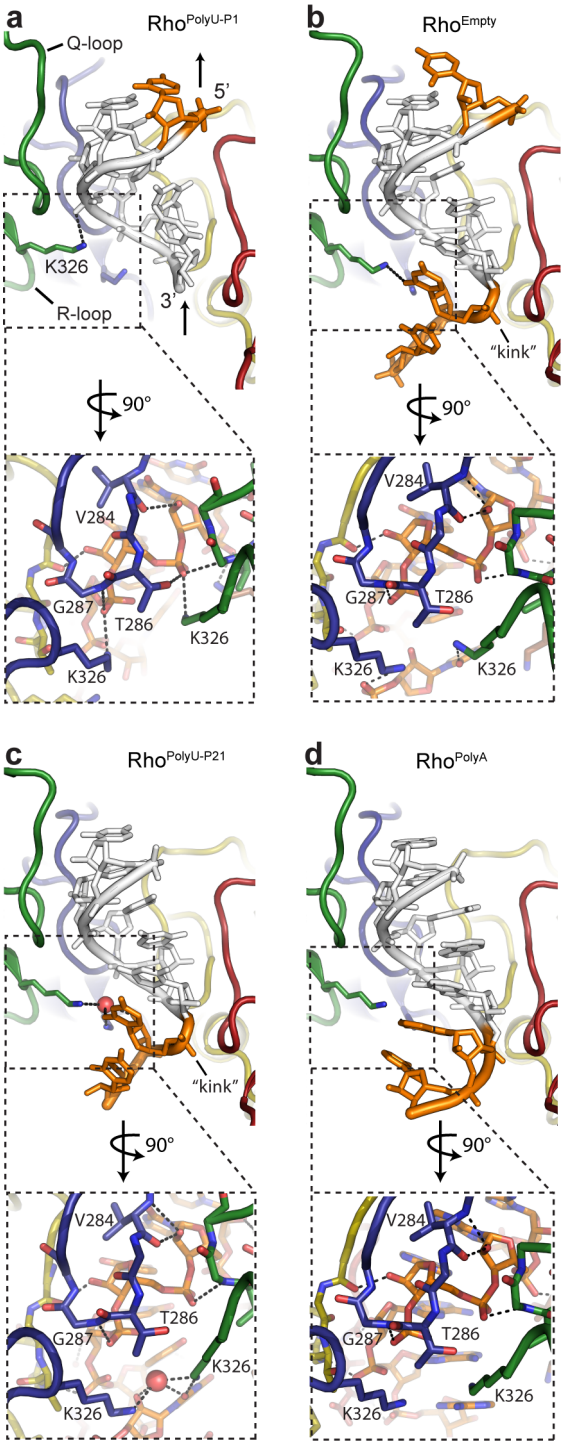

Figure 2.5. Comparison of RNA binding among closed-ring Rho structures. The five core RNA bases common to all structures are colored white, while additional 5'- and 3'-bases are colored orange. Arrows at the 5' and 3' end of the bound RNA indicate the direction of RNA translocation relative to the protein. Protein subunits are colored as in **Figure 2.2a**. Insets reveal protein-RNA hydrogen bonding details (dashed lines) at the B/C subunit interface. (**a**) $Rho^{PolyU-P1}$ contains six bases of RNA (orange/white) bound

by the Rho Q- and R-loops in a spiral staircase conformation. (**b**) The Rho^{Empty} structure allows for the modeling of eight bases of RNA (orange/white). Compared to the $Rho^{PolyU\text{-}P}{}_1$ structure, there is additional density for two 3' bases in a kinked conformation. (**c**) The $Rho^{PolyU\text{-}P2_1}$ structure allows for modeling of seven bases of RNA (orange/white). Compared to the Rho^{Empty} structure, density is lacking for the 5' base. This structure reveals a hydrogen bonding network stabilizing interactions between R-loop residue K326 of subunits B and C, a water molecule, and the 2'-OH and uracil base of the bound RNA kink. (**d**) The structure of Rho^{PolyA} solved under similar conditions to $Rho^{PolyU\text{-}P2_1}$, also reveals two additional 3' bases, but in a pseudo A-form helical arrangement, rather than the kinked conformation; in contrast to the H-bonding network observed with the kinked PolyU substrate, no additional protein-RNA contacts are formed.

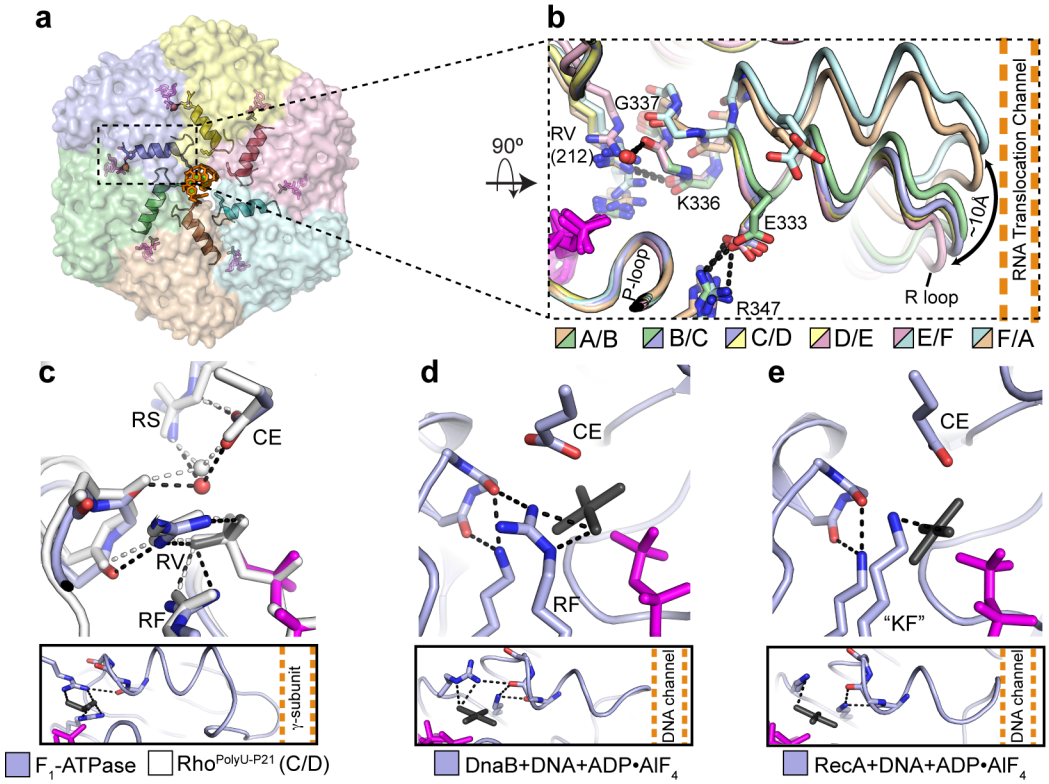

Figure 2.6. A conserved allosteric helix couples pore loop movement and ATPase status in RecA-like hexameric motors. (**a**) An α-helix (Rα1, cartoon) in each subunit of the Rho hexameric motor domain (transparent surface) directly connects the RNA and ATP binding sites. RNA – orange, nucleotide – magenta, catalytic water molecules – red spheres (subunits B and C only). Lys326 (colored sticks) directly connects the allosteric helix to bound RNA. (**b**) Alignment of all active sites in $Rho^{PolyU\text{-}P2_1}$ reveals the consequences of inter-subunit conformational transitions in the allosteric helix. The helix itself directly links the RNA binding R-loops in the center of the ring to the position of the γ-phosphate sensing residues in the ATP-binding sites at the periphery of the ring, as well as to the catalytic water in the *T** active sites. Dashed lines (colored by interface) show bonding interactions present in the B/C, C/D and D/E subunit interfaces. Abbreviations: RV – arginine valve, RF – arginine finger. (**c**) Alignment of the catalytic site in the C/D subunit interface of $Rho^{PolyU\text{-}P2_1}$ (white) with the catalytic site in the structure of the F_1-ATPase bound to ADP•BeF_3 (colored, PDB ID: 1W0J) reveals nearly identical alignments of the catalytic residues and the backbone carbonyl groups of the allosteric helix (Kagawa et al., 2004). Lower panel –position of the F_1 allosteric helix with respect to the rotating γ-subunit shaft. (**d**) The structure of DnaB bound to nucleic acid and GDP•AlF_4 reveals interactions between its analog of helix Rα1 and γ-phosphate sensing residues (Itsathitphaisarn et al., 2012). (**e**) The structure of RecA bound to double-stranded DNA and ADP•AlF_4 reveals a similar helix configuration as in panels **b-d** (Chen et al., 2008). KF – lysine finger.

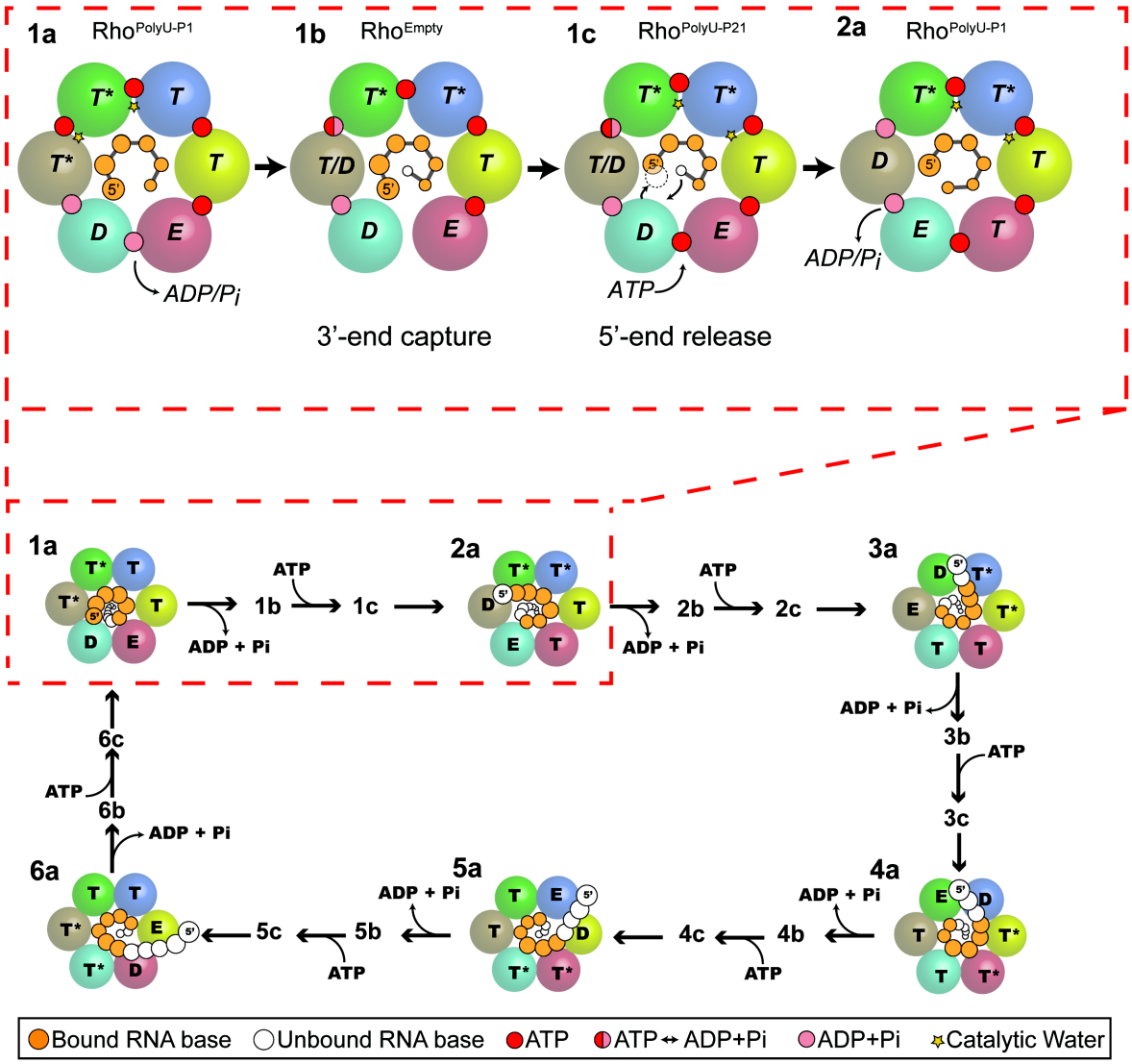

Figure 2.7. Sub-stepping in a hexameric helicase. (*Top, red dashed box*) Diagram of RNA bases and catalytic states in the Rho$^{PolyU\text{-}P1}$, Rho$^{PolyU\text{-}P2_1}$, and Rho^{Empty} structures. Protein subunits are shown as large spheres colored by chain as in **Figure 2.2a**. Arrows indicate the sequence of proposed sub-steps in which an incoming 3' RNA base is bound and a 5' RNA base is released consistent release of nucleotide and a shift in position of the *T** catalytic states. Yellow stars designate the sites of the nucleophilic water observed in the various Rho structures; only in the Rho$^{PolyU\text{-}P2_1}$ structure does this water adopt an in-line attack position. (*Bottom*) Relationship of the substeps to a complete series of six sequential hydrolysis and translocation steps by Rho.

Supplementary Material

Supplementary Discussion

Detailed analysis of ATPase site configuration

Concomitant with the observed changes in catalytic water position (**Figure 2.2e**), several functionally critical amino acids in the E/F and F/A interfaces alter their positions between the P1 and $P2_1$ states. At the F/A interface, which is proposed to correspond to an apparent product-release (*D*) state, an arginine that was previously displaced from the active site and that has been proposed to regulate P_i release in the related F_1-ATPase (Abrahams et al., 1994) – the Arginine Valve (RV) – can now be seen to coordinate the bound BeF_3 moiety (**Figure 2.3b**). A comparison of changes at the F/A inter-subunit interface reveals that this re-positioning is mediated by a small movement within the catalytic center, which relieves a steric clash that otherwise prevents the RV from contacting the BeF_3 moiety in the prior P1 Rho structure (**Figure 2.3c**). Intersubunit conformational changes mediate a similar phenomenon at the E/F protomer interface, where the arginine finger (RF) (proposed to stabilize the ATPase transition state in systems such as G-proteins (Wittinghofer et al., 1997)), now contacts the BeF_3 moiety, which itself is also bound in a more idealized conformation with ADP, similar to that observed in the T states (**Figure 2.3d**). The catalytic glutamate (CE) is likewise displaced from its position in the ATPase center of subunit E, a move mediated by shifts in the arginine switch (RS) residue from the allosteric network (**Figure 2.3e**); the positioning of the RV and RF residues in the *D*- and *E*-states of the F/A and E/F subunit interfaces likewise differ in their placement from the ATP-bound (*T*) and hydrolysis-competent (*T**) states found in other active sites (**Figure S2.6b**). Because of the sub-optimal liganding opportunities afforded by the amino acid shifts in the subunit A and F sites, the refined B-factors for the nucleotide modeled into the subunits adopting these states are concordantly high compared to other ADP•BeF_3 moieties present in the model (**Figure S2.4**).

Detailed analysis of sequence-specific RNA binding

Concomitant with the distinct 5'- and 3'-end structures of different bound RNAs, specific protein-RNA contacts also undergo a change. Overall, the new poly-U and poly-A models form a similar number of total polar contacts between Rho and RNA when compared to the prior $Rho^{PolyU\text{-}P1}$ structure, but these frequently involve different residues. In addition to the changes in Lys326 noted above, Thr286 of Rho's Q-loop (the L1-loop equivalent of DnaB and T7 gp4) loses its interaction with the phosphate of RNA that was seen in the original closed-ring structure of Rho, leading both the backbone amide (not observed in $Rho^{PolyU\text{-}P1}$) and carbonyl (observed in $Rho^{PolyU\text{-}P1}$) groups of Val284 to shift and make a new contact with the 2'-OH of its associated ribose (**Figure 2.5a-d**). These three models also manifest a slightly steeper RNA coil that is less compressed compared to the RNA seen previously in the $Rho^{PolyU\text{-}P1}$ model, likely due to changes in the intersubunit orientations between protomers E/F and F/A (**Figure 2.2e and Figure S2.7**).

Delay of product release in Rho and RecA helicases

Of the two active site states corresponding to the B/A and A/F subunit pairs, the one closest to the preceding *T** state (subunit A) at first glance appears superficially similar to a T-state in the geometric configuration of its active site. However, studies of Rho and related RecA-family hexameric helicases have shown that *ATP-γ-S* binds more tightly than ATP (5 vs. 3 tight binding sites in Rho (Xu et al., 2003)), suggesting that the non-hydrolyzable inhibitor can stably occupy product-binding sites. Moreover, nucleotide release has been shown to lag behind ATP hydrolysis in both Rho and the T7 gp4 helicase by ~2-4 subunits, and post-hydrolysis active sites have been reported to be capable of both hydrolyzing and synthesizing ATP (Adelman et al., 2006; Liao et al., 2005b; Xu et al., 2003), indicating that the active site environments that promote substrate binding *vs*. product formation are highly similar. Given these findings, and that the A/B and F/A catalytic centers are flanked by clear exchange (*E*) and hydrolysis-competent (*T**) states, the two (*D*) sites likely represent an ATP/ADP•P_i (*T/DP*) post-hydrolysis/pre-product release intermediate. Consistent with this assignment, both interfaces have lost the R347•E333 and K336 carbonyl•R212 intersubunit contacts that are found in the T and T* states and that appear crucial for orienting the arginine finger and catalytic glutamate residues (**Figure 2.3**).

Supplementary Tables

Table S2.1. Rho SAXS data statistics.

RNA	[ADP•BeF_3] (mM)	Guinier R_g (Å) **AutoRg**	Real Space R_g (Å) **Gnom**	Reciprocal Space R_g (Å) **Gnom**	D_{max} (Å) **Gnom**
-	0.0	49	47.6	47.7	134.5
-	0.5	47.6	47.28	47.38	143
-	1.0	47.1	47.23	47.36	135.5
-	2.0	47.4	46.84	46.96	131
+	0.0	49.7	49.3	49.3	164.5
+	0.5	48.3	47.24	47.34	139
+	1.0	47.3	46.31	46.4	141
+	2.0	45.8	45.3	45.4	139.5

Table S2.2. Rho SAXS data fitting to crystal structures (FOXS).

RNA	[ADP•BeF_3] (mM)	Open Ring Fit (χ)	Closed Ring Fit (χ)	Ensemble Fit (χ)	% Closed / Open
-	0.0	3.77	24.15	3.22	8.0 / 92.0
-	0.5	3.42	24.79	2.25	9.0 / 91.0
-	1.0	3.41	23.8	2.20	10.5 / 89.5
-	2.0	4.87	21.82	2.08	17.6 / 82.4
+	0.0	4.23	25.13	3.79	7.3 / 92.7
+	0.5	5.10	23.03	2.16	17.6 / 82.4
+	1.0	13.05	11.37	1.98	57.9 / 42.1
+	2.0	21.62	2.93	2.93	100 / 0

Table S2.3. Rho SAXS data fitting to crystal structures (OLIGOMER).

RNA	[ADP•BeF_3] (mM)	Open Ring Fit (χ)	Closed Ring Fit (χ)	Ensemble Fit (χ)	% Closed / Open
-	0.0	5.20	27.75	4.09	10.2 / 89.8
-	0.5	6.36	28.62	5.19	11.4 / 88.6
-	1.0	6.31	27.44	4.96	12.4 / 87.6
-	2.0	8.31	24.86	4.90	21.3 / 78.7
+	0.0	4.98	29.07	4.45	7.0 / 93.0
+	0.5	8.08	26.36	4.71	20.1 / 79.9
+	1.0	21.24	13.15	5.25	63.3 / 36.7
+	2.0	33.70	6.46	6.46	100 / 0

Supplementary Figures

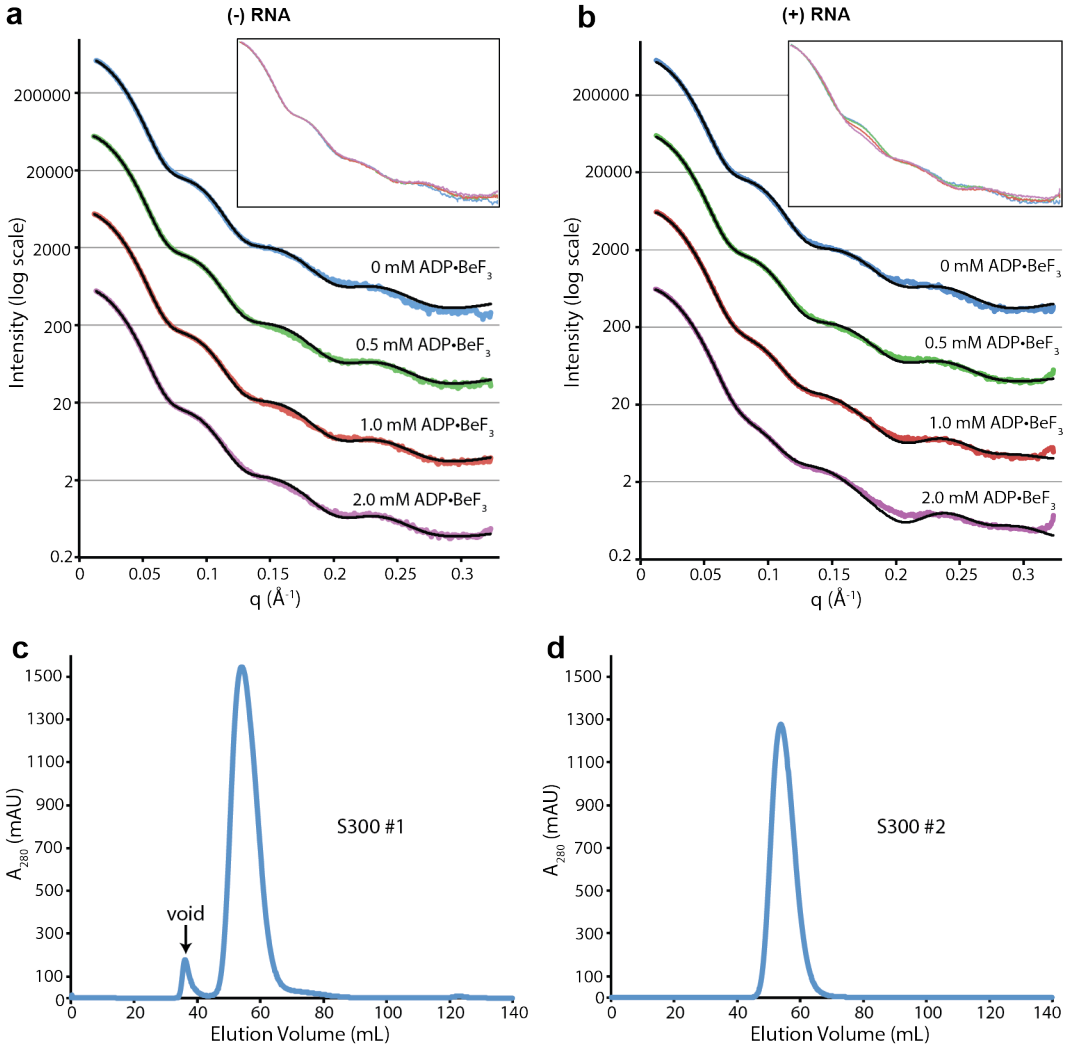

Figure S2.1. SAXS analysis and size-exclusion chromatography profiles for Rho. (**a**) SAXS curves for solutions of 5 mg/mL Rho in the presence of increasing concentrations of ADP•BeF_3. The intensity values on the log-scaled Y-axis are offset by powers of ten for clarity. The best fit obtained from combinations of two x-ray crystal structures (open-ring Rho, 1PVO, and closed-ring Rho, PolyU-$P2_1$) using the program FOXS is shown as a solid black line (**Methods**). Inset shows an overlay of all four SAXS data curves, showing the absence of obvious changes in the scattering profiles upon increasing nucleotide titration. (**b**) SAXS curves for solutions of 5 mg/mL Rho in the presence of a one molar equivalent of RNA per Rho hexamer and increasing concentrations of ADP•BeF_3 displayed as in panel (a). The inset shows an overlay of all four SAXS curves, showing the distinct changes in low-resolution portions of the scattering curve as the concentration of nucleotide is increased. (**c**) Initial size exclusion chromatography run on full length Rho protein using a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) reveals a small void peak and a larger peak consistent with a hexameric Rho particle (282 kDa). (**d**) After pooling the hexameric peak and re-concentrating the protein, a second size exclusion run reveals only a single peak corresponding to hexameric Rho, confirming the absence of aggregated protein.

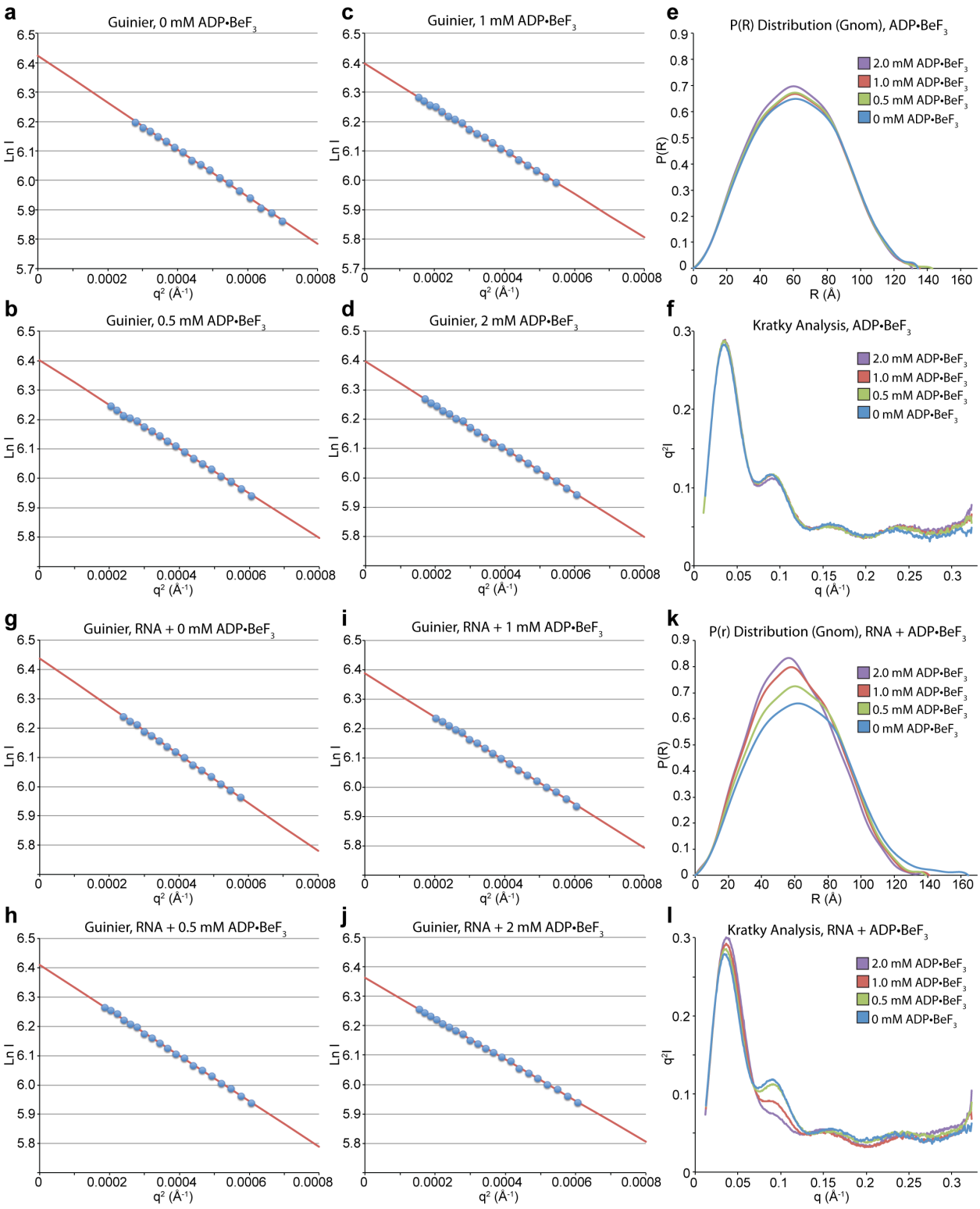

Figure S2.2. Analysis of SAXS data indicates that Rho complexes are well behaved in solution. **(a-d, g-j)** Guinier analysis. **(e, k)** Real-space distance distribution profiles. **(f, l)** Kratky analysis.

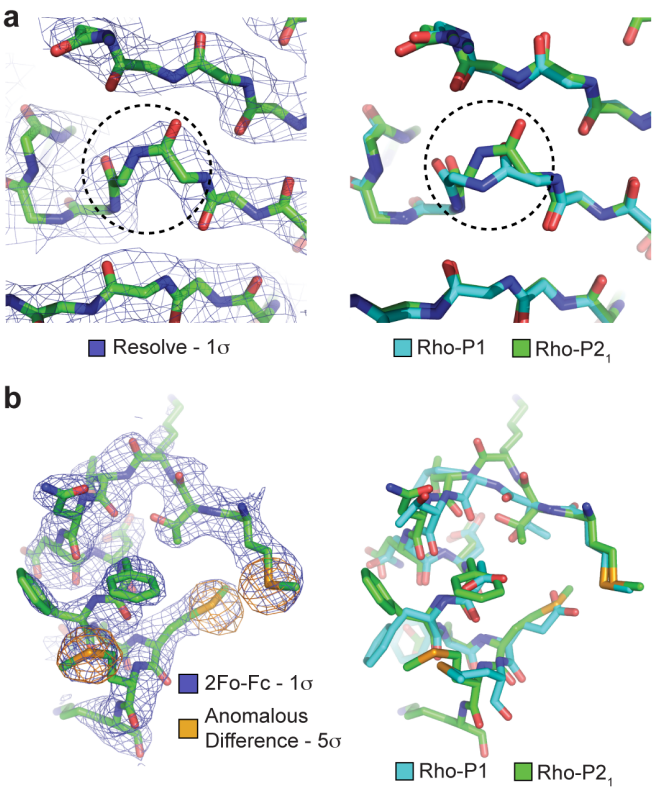

Figure S2.3. Phase information corrects minor building errors in previous Rho structures. **(a**, **left**) Density modified SAD phases calculated by the program RESOLVE and contoured at 1σ show clear electron density for a cis-peptide bond (Terwilliger, 2000) near the Walker-B motif. (**a, right**) Alignment of Rho$^{PolyU\text{-}P2_1}$ (green) with Rho$^{PolyU\text{-}P1}$ (cyan) how the backbone atoms (dashed circle) were misbuilt in the earlier Rho model. A *cis*-peptide bond in this position is a common feature of RecA-like proteins, as first observed in the structure of the RecA recombination protein (Story et al., 1992). Although this unique bond could not be modeled properly in earlier Rho structures due to insufficient resolution, the correction does not alter the conformation of the ATP binding site, and thus does not alter the functional interpretation of previous structures. (**b**) A second correction was facilitated by phased anomalous difference maps, which allowed us to precisely locate most of the 92 selenium sites in the Rho hexamer. Final $2F_o\text{-}F_c$ electron density maps contoured at 1σ and phased anomalous difference maps contoured at 5σ clearly reveal the location of adjacent selenomethionine residues in Rho$^{PolyU\text{-}P2_1}$ crystals (left). These maps show that the α-helix located at the C-terminus of Rho (containing two adjacent selenomethionine residues) was register-shifted by one residue in previous Rho structures (right). This error likely explains the high B-factors observed in this region previously (Skordalakes and Berger, 2003; Skordalakes et al., 2005; Thomsen and Berger, 2009), but has no bearing on functional interpretations of prior Rho models.

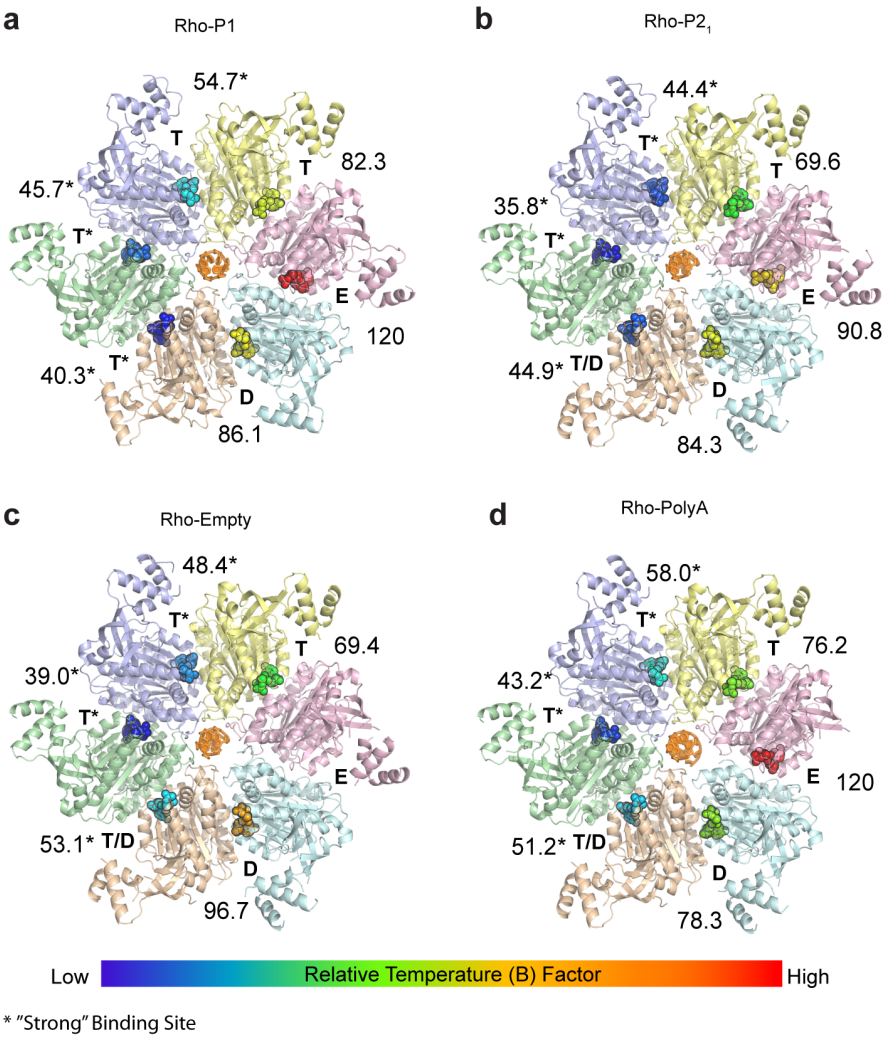

Figure S2.4. Temperature-factor analysis of bound nucleotides in different Rho structures. **(a-d)** Structures of various Rho hexamers with nucleotides colored by average temperature factors and scaled to correct for differences in average B-factors among structures determined at different resolutions (see **Methods**). The Rho hexamer is shown as a transparent cartoon, while the six ADP•BeF_3•Mg^{2+} complexes are shown as spheres. Consistent with the prior $Rho^{PolyU\text{-}P1}$ model, all structures have three active sites with low B-factors ranging from 35 – 58 Å^2, while the remaining three active sites have higher average B-factors ranging from 70-120 Å^2. Proposed nucleotide states for Rho^{PolyA} and Rho^{Empty} are based on the near-isomorphous but higher resolution $Rho^{PolyU\text{-}P2_1}$ structure.

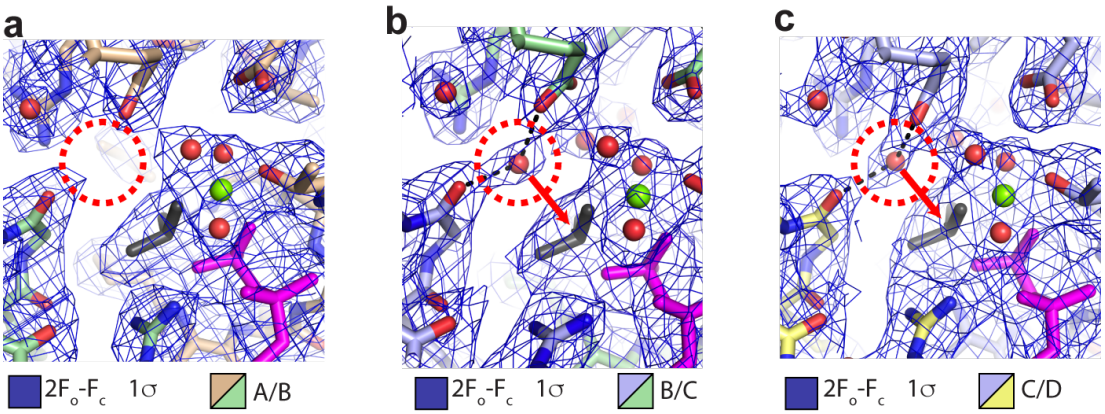

Figure S2.5. Final 2Fo-Fc electron density shown at 1σ. (**a-c**) Final $2F_o$-F_c electron density for the A/B, B/C and C/D subunit interfaces showing density for an in-line attacking water molecule in the subunit B and C active sites, but not in subunit A. Subunits colored as in **Figure 2.2**, with oxygen atoms colored red and nitrogen atoms colored blue.

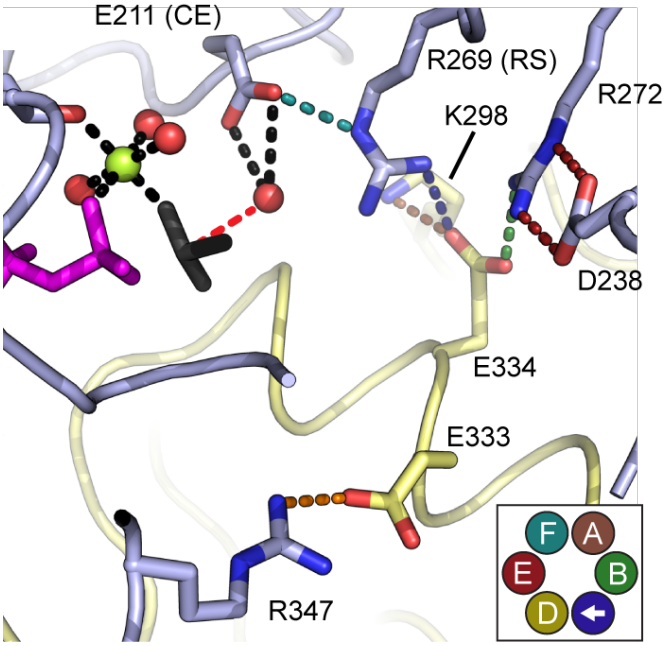

Figure S2.6. Allosteric communication network in $Rho^{PolyU\text{-}P2_1}$. A fully-linked allosteric network of conserved salt-bridges forms at the subunit C/D interface in the new $Rho^{PolyU\text{-}P2_1}$ structure(Thomsen and Berger, 2009). Interactions between members of the allosteric network appear to correlate with catalytic state, with maximal interactions serving to properly position the catalytic glutamate (CE) by the arginine switch (RS). Dashed lines representing salt bridging interactions are colored to highlight the six inter-residue interactions that form the allosteric network. Bonds involved in ATP binding and catalytic water coordination are colored black. The inset shows approximate view from within the context of the Rho hexamer (i.e., in the active site of subunit C (arrow), looking at subunit D).

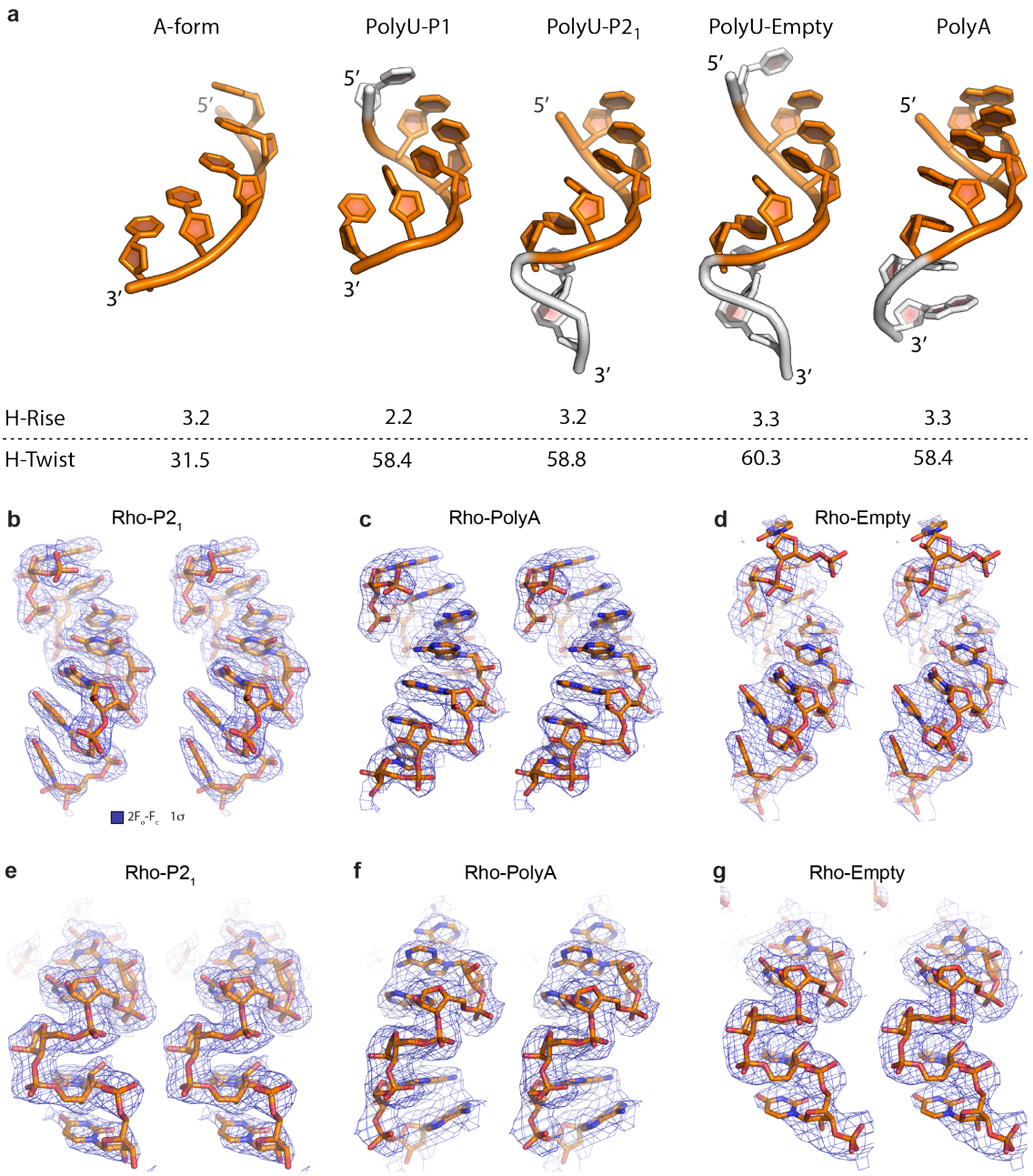

Figure S2.7. **(a)** Cartoon representations of idealized, A-form RNA alongside the RNA bound to each of the four Rho-RNA-ADP•BeF_3 structures. The five core RNA bases present in all structures are colored orange. Helical rise (H-Rise) and helical twist (H-Twist), as calculated by 3DNA(Lu and Olson, 2003) for the core RNA bases, are listed below each structure. The new $P2_1$ crystal structures presented in this study have a helical rise similar to A-form RNA and are thus less compressed than the RNA bound in the prior $Rho^{PolyU\text{-}P1}$ structure; however, to make a complete turn with one RNA base bound per subunit of Rho, all bound RNAs are similarly overwound with a helical twist of ~60°, almost double that observed in A-form RNA. (**b-g**) Stereo views of the final refined electron density for the $Rho^{PolyU\text{-}P2_1}$, Rho^{Empty} and Rho^{PolyA} structures. Close up views in

panels e, f, and g show the presence or absence of a kink in the bound RNA in the new poly-U and poly-A bound structures, respectively.

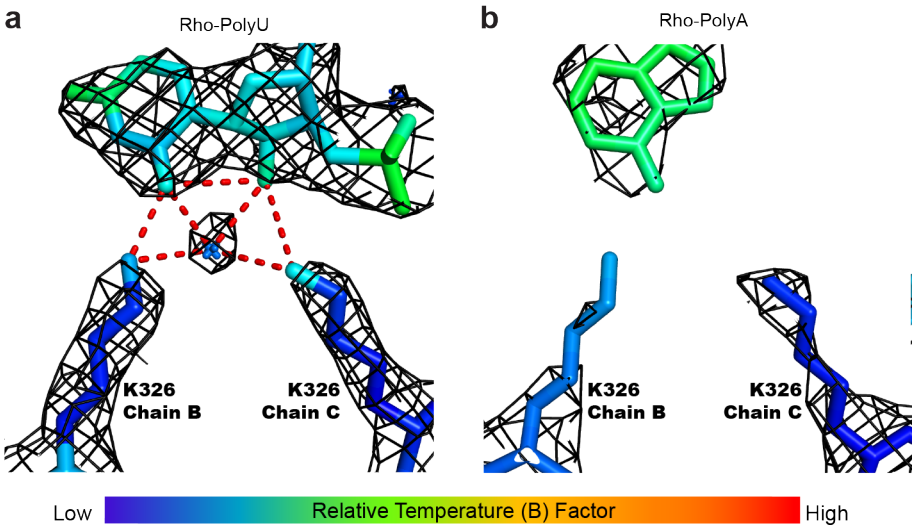

Figure S2.8. Comparison of temperature factors and electron density around the 3' RNA end of both the $Rho^{PolyU\text{-}P2_1}$ and Rho^{PolyA} structures. **(a)** Contacts between Lys326 residues of Rho in the $Rho^{PolyU\text{-}P2_1}$ structure and the uridine base and ribose 2'OH are mediated through a single tetrahedrally coordinated water molecule and direct interactions. Lys326 of Chain B is as well resolved as Lys326 of Chain C, and displays similar B factors. **(b)** Contacts between Lys326 residues of Rho in the Rho^{PolyA} structure and the adenine base are not formed compared the corresponding region in the $Rho^{PolyU\text{-}P2_1}$ structure. Lys326 of Chain B is less well-ordered than Lys326 of Chain C and displays higher B factors.

Chapter 3: – Ligand-Induced and Small-Molecule Control of Substrate Loading in a Hexameric Helicase

(Portions of this chapter have been reproduced from Lawson M.R., Dyer K. and Berger J.M., PNAS 2016).

Introduction

Ring-shaped hexameric helicases and translocases are motor proteins that control myriad essential viral and cellular processes. Many hexameric motors undergo substrate-dependent conformational changes that couple activity to the productive binding of client substrates (Lyubimov et al., 2011; Matyskiela and Martin, 2013; O'Shea and Berger, 2014; Singleton et al., 2007). Internal regulatory domains and exogenous proteins or small molecules frequently impact client substrate recruitment and engagement by these enzymes (Bashore et al., 2015; Froelich et al., 2014; Jeong et al., 2004; Strycharska et al., 2013); however, it is generally unclear how such factors control helicase or translocase dynamics.

Rho is a hexameric helicase responsible for controlling approximately 20% of all transcription termination events in *E. coli* (Peters et al., 2009). Rho is initially recruited to nascent transcripts in an open, lockwasher-shaped configuration (**Figure 3.1A**) (Skordalakes and Berger, 2003; Yu et al., 2000), where it binds preferentially to pyrimidine-rich sequences (termed Rho utilization of termination sequences, or '*rut*' sites) using a 'primary' RNA-binding site located in the N-terminal OB folds of the hexamer (Bogden et al., 1999; Dolan et al., 1990; Morgan et al., 1985). Following *rut* recognition, Rho converts into a closed-ring form (**Figure 3.1B**), locking the RNA strand into a 'secondary' RNA binding site (formed by two conserved sequence elements known as the Q and R loops (Miwa et al., 1995)) located within the central pore of the hexamer. This conformational change, which we show in Chapter 2 to be both RNA- and ATP-dependent , rearranges residues in the Rho ATP binding pockets into a hydrolysis-competent state (Thomsen and Berger, 2009). Once engaged, Rho maintains primary site contacts with the *rut* sequence as it translocates 5′ to 3′ along the RNA strand in an ATP-dependent manner, a process known as tethered tracking (Galluppi and Richardson, 1980; Gocheva et al., 2015; Koslover et al., 2012a; Steinmetz and Platt, 1994). Rho elicits termination by applying direct or indirect forces to RNA polymerase (Park and Roberts, 2006; Schwartz et al., 2007a), dislodging it from DNA and the newly made RNA in the transcription bubble.

Although the basic steps of Rho-dependent transcription termination have been largely elucidated, and it has become increasingly clear that Rho's activity, like that of many helicases and translocases, can be controlled by a variety of intrinsic and extrinsic factors. The small molecule bicyclomycin (**Figure 3.1C**), a highly specific chemical inhibitor of Rho that has been shown to wedge itself in between subunits of an open Rho ring (**Figure 3.1D** and **Figure S3.1**) (Skordalakes et al., 2005), is one such example. Previous studies have shown that bicyclomycin is a noncompetitive inhibitor of Rho ATPase activity and a mixed inhibitor of RNA binding to Rho's secondary site (Magyar et al., 1996; Park et al., 1995). Although subsequent structural studies

suggested that bicyclomycin antagonizes Rho by sterically preventing the binding of the nucleophilic water molecule that initiates ATP hydrolysis (Skordalakes et al., 2005), how bicyclomycin might interact with catalytically-competent Rho states, such as those thought to accompany ATPase activity and translocation (Thomsen and Berger, 2009), has not been defined.

It is similarly unclear how other dissociable factors – e.g., regulatory proteins and/or nucleic acids – influence Rho activity. It is well established that the sequence of the RNA itself has a pronounced impact on whether a transcript will be acted upon by Rho (Hart and Roberts, 1991; McSwiggen et al., 1988; Morgan et al., 1985; Schwartz et al., 2007b). Binding of pyrimidine-rich sequences to the N-terminal RNA binding domains of Rho is a particularly well-known accelerant of Rho's ATPase activity (Richardson, 1982), with pre-steady-state ATPase assays showing that the formation of a catalytically competent Rho ring is governed by a rate-limiting RNA- and ATP-dependent conformational change (Jeong et al., 2004). The ligand-dependence of this isomerization event correlates with the requirements for ring closure identified in the work described in Chapter 2. These findings raise the intriguing possibility that the sequence specificity of Rho-dependent termination is in part due to an increased efficiency of ring closure when the N-terminal RNA binding domains are occupied by pyrimidine-rich sequences.

Here we show that bicyclomycin and primary-site occupancy affect the structural state of Rho in an opposing manner. Using small-angle X-ray scattering and a fluorescence-based assay to track ring status *in vitro*, we demonstrate that bicyclomicin inhibits RNA binding to the central pore of Rho by sterically impeding ring closure. In contrast, the binding of pyrimidine-rich nucleic acids to the Rho N-terminal RNA binding domains promotes Rho ring closure, helping the capture of non-ideal, purine-rich RNA sequences within Rho's translocation pore. Collectively, these findings help highlight diverse mechanisms by which ligand binding to discrete sites on ring-type ATPases can activate or repress conformational changes and govern the function of the motor.

RESULTS

The bicyclomycin-binding pocket collapses upon Rho ring closure

During the course of examining the bicyclomycin-binding pocket across different Rho structural states ((Thomsen and Berger, 2009) and Chapter 2), we noticed that the site appeared substantially smaller in closed-ring states than in the open-ring states. Computational analysis confirmed that the pocket volume is substantially larger in the open-ring conformation (272 $Å^3$, **Figure S3.2A**) than in the closed-ring state (56 $Å^3$, **Figure S3.2B**), a trend that held true with all open- and closed-ring Rho structures that have been observed to date (**Table S3.1** and **Figure S3.2**). Correspondingly, alignment of a bicyclomycin-bound open ring structure (PDB ID 1XPO (Skordalakes et al., 2005)) with a closed-ring Rho model (PDB ID 3ICE (Thomsen and Berger, 2009)) revealed extensive steric clashes with bicyclomycin and residues P180, K184, and E211 in the closed-ring state that would appear to preclude binding (compare **Figure 3.1D** and **Figure 3.1E**). From these observations, we postulated that bicyclomycin might not sterically block the attacking water needed for ATP hydrolysis, as previously proposed, but instead antagonize the RNA- and ATP-dependent conformational switching of Rho from an open- to closed-ring state.

Bicyclomycin inhibits Rho ring closure over a range of nucleotide concentrations

To determine the impact of bicyclomycin on Rho conformational state, we first employed Small-Angle X-Ray Scattering (SAXS) to directly monitor structural dynamics in solution. The work described in Chapter 2 demonstrates that open-ring Rho molecules close in the presence of RNA and the nonhydrolyzable ATP analog ADP•BeF_3, but not in the presence of either ligand individually. Since bicyclomycin is a relatively modest inhibitor of Rho ATPase activity (I_{50} = 20 μM (Park et al., 1995)), we reasoned that its effects on ring closure might be most evident when examined over a range of nucleotide concentrations. As a control, we carried out SAXS studies using a series of ADP•BeF_3 concentrations in the presence of RNA and the absence of bicyclomycin. By plotting the average intensity of the observed scattering (I) as a function of scattering vector (q) for a variety of ADP•BeF_3 concentrations (**Figure 3.2A**), nucleotide-dependent differences in curve shapes were clearly evident, similar to those described in Chapter 2 (see **Figure S3.3**). Visual inspection of the curves over a q range from 0.07 to 0.13 $Å^{-1}$ (the region of the curve in which nucleotide-dependent differences are most pronounced, **Figure 3.2A**, *Inset*) revealed that the profiles converged between 1.5 and 4 mM ADP•BeF_3, suggesting that under these conditions Rho exists predominantly as a closed ring. Indeed, quantification of the percentage of open *vs*. closed states using a Minimal Ensemble Search ('MES', (Schneidman-Duhovny et al., 2010a)) suggested that 1.5 mM ADP•BeF_3 was sufficient to close the majority of Rho rings in solution, with an increase in the percentage of closed rings observed at 2, 3 and 4 mM ADP•BeF_3 (**Figure 3.2C** and **Table S3.2**). Similar trends were also observed in the reduction of the radius of gyration ('R_g') at increasing ADP•BeF_3 concentrations, consistent with the expected nucleotide-dependent compaction of the Rho ring upon transition to a predominantly closed-ring state (**Figure S3.4**, **Figure S3.5** and **Table S3.3**).

After establishing a SAXS regime for looking at changes in ring state, we next set out to determine whether bicyclomycin would antagonize the ability of the Rho ring to close in a nucleotide-dependent fashion. We first pre-incubated Rho with a high concentration of bicyclomycin (400 μM final concentration, an 8-fold molar excess over Rho monomer), and then added RNA and a variety of ADP•BeF_3 concentrations. The resulting SAXS curves (**Figure 3.2B**) also displayed differences that were clearly dependent upon nucleotide concentration; however, at intermediate concentrations of ADP•BeF_3 (0.5, 1 and 1.5 mM), curves were skewed towards the open-ring state relative to the drug-free curves (compare **Figure 3.2B,** *Inset* and **Figure 3.2A**, *Inset*). Quantification of ring state by MES showed a substantial shift toward the open form when bicyclomycin was present, which was particularly notable at 1.5 and 2 mM ADP•BeF_3 (**Figure 3.2C** and **Table S3.2**). This bicyclomycin-dependent inhibition of Rho ring closure was also consistent with observed increases in R_g upon addition of the drug (**Figure S3.4A**, **Figure S3.6** and **Table S3.3**).

Rho ring state can be controlled by varying bicyclomycin concentration

After observing that bicyclomycin inhibits ring closure, we wondered whether it would be possible to control the Rho ring state by varying the concentration of bicyclomycin with Rho prior to adding RNA and ADP•BeF_3. To address this question, we elected to conduct a SAXS experiment in the presence of 1.5 mM ADP•BeF_3, a nucleotide concentration that based on our previous titration is sufficient to close a substantial fraction (but not all) of the Rho rings present in solution. Pre-incubating Rho with varying concentrations of bicyclomycin yielded significant differences in the observed SAXS curves (**Figure 3.3A**), and by focusing on the mid-q region (**Figure 3.3A,** *Inset*), it is immediately evident that the ring state can be affected by bicyclomycin. Quantification of the percentage of the open *vs*. closed ring populations by MES (**Figure 3.3B** and **Table S3.4**) shows that ring state changed from predominantly closed in the absence of bicyclomycin to mostly open at high bicyclomycin concentrations. Observed R_g values ranged from 46.3 Å in the absence of bicyclomycin to 48.2 Å in the presence of 3200 μM bicyclomycin, consistent with bicyclomycin concentration-dependent inhibition of Rho ring closure (**Figure S3.4B**, **Figure S3.7** and **Table S3.5**).

A fluorescence-based assay tracks ligand-dependent effects on Rho ring status in vitro

As SAXS studies constitute only one assessment of Rho state, we next set out to develop an *in vitro* assay to further characterize ligand-dependent effects on the conformation of the hexamer. We reasoned that a short, fluorescein-derivatized RNA would have a significantly faster tumbling rate when free in solution than when stably bound to the secondary (translocation pore) site inside a closed Rho ring (see diagram in **Figure 3.4A**), thus enabling us to track ring closure by monitoring changes in Fluorescence Anisotropy ('FA'). To rule out the possibility that labeled RNA binding to Rho's primary (RNA-recruitment) site might convolute the signal, we first pre-incubated the protein with a short pyrimidine-rich DNA ('dC_5') at a concentration equimolar to that of Rho monomer (polypyrimidine DNAs bind both tightly and selectively to the primary sites of Rho (Morgan et al., 1985; Richardson, 1982)). No anisotropy changes were observed with a fluorescein-labeled poly-U RNA ('rU_{12}*') in the absence of nucleotide

and in the presence of a primary site competitor (**Figure 3.4B**, black), indicating that rU_{12}* alone does not have high affinity for the secondary site in open-ring Rho. By contrast, robust rU_{12}* binding was observed as nucleotide was added (**Figure 3.4B**), showing that the secondary site becomes capable of stably binding RNA as the Rho ring closes. These findings are consistent with the SAXS studies described in Chapter 2, showing that both RNA and nucleotide are needed to cooperatively promote ring closure in Rho.

To further characterize the mechanism by which bicyclomycin inhibits Rho, we next conducted an order-of-addition experiment examining the impact of bicyclomycin on rU_{12}* binding either before or after Rho ring closure. As expected, pre-incubation of bicyclomycin with Rho before the addition of RNA and nucleotide inhibited rU_{12}* binding in a bicyclomycin concentration-dependent manner (**Figure 3.4C**). Conversely, the impact of bicyclomycin on rU_{12}* binding was far less evident when closed rings were pre-assembled before addition of the drug (**Figure 3.4D**). These data indicate that although bicyclomycin is able to bind to an open Rho ring and prevent ring closure, the molecule is unable to force open a pre-closed Rho ring.

Rho ring closure is promoted by the binding of pyrimidine-containing nucleic acids to the N-terminal, primary site domains

Although RNA binding to Rho's secondary site (the helicase translocation pore) is a major effector of ring closure, RNA binding to Rho's primary site (its OB folds) also has been implicated in controlling Rho function (McSwiggen et al., 1988). To directly probe the role of primary site ligands in modulating ring closure, we tested whether the presence or absence of dC_5 could impact the binding of rU_{12}* to the secondary site. We also conducted the same experiments with fluorescein-labeled poly-A RNA ('rA_{12}*'), which is a weaker secondary site ligand than poly-U and is incompatible with primary site binding (Bogden et al., 1999; Faus and Richardson, 1990). A clear increase for rA_{12}* affinity was evident in the presence of dC_5 (**Figure S3.8**), especially at low ADP•BeF_3 concentrations. Under all nucleotide concentrations tested, rU_{12}* bound more tightly to Rho than rA_{12}* (**Figure S3.8**), with dC_5 subtly promoting ring closure at moderate (1 mM ADP•BeF_3) nucleotide concentrations. Primary site occupancy appeared to have either no effect (1.5 mM ADP•BeF_3) or a slight inhibitory effect (2-3 mM ADP•BeF_3) on rU_{12}* binding at higher nucleotide concentrations. Collectively, these results signify that nucleic acids bound to the primary site promote ring closure under conditions that are sub-optimal for RNA binding to the secondary site.

Based on the spacing of the OB folds in the Rho structure and the observation that the isolated N-terminal domain only engages a pyrimidine dinucleotide (**Figure 3.1a-b**) (Bogden et al., 1999; Skordalakes and Berger, 2003), each dC_5 substrate would be expected to bind one primary site at a time. Having observed that this short oligonucleotide can nonetheless promote Rho ring closure, we wondered whether longer DNAs capable of bridging multiple primary sites in a single Rho hexamer would have a more pronounced impact on ring state (see diagram in **Figure 3.5A**). We repeated the rU_{12}* and rA_{12}* binding assays with either a 15-mer DNA capable of

bridging two primary sites (dC_{15}) or a 79-mer DNA sufficiently long to potentially bridge all six primary sites (dC_{75}), using an ADP•BeF_3 concentration (1 mM) at which primary-site binding of DNA has the most significant effect on RNA binding to the secondary site. The number of putative sites that could be occupied by the DNAs was held constant relative to the dC_5 experiments by adjusting the relative molar ratios between Rho and the DNA (**Methods**). For both rU_{12}* and rA_{12}*, inclusion of dC_{15} or dC_{75} more robustly promoted RNA binding than the shorter dC_5 oligo (**Figure 3.5B-C**). Using fluorescein-labeled derivatives of the various primary-site ligands (denoted as dC_5*, dC_{15}*, and dC_{75}*), titration experiments revealed that DNAs long enough to bridge multiple primary sites have a much higher affinity for Rho than dC_5 (**Figure 3.5D**). By fixing Rho at a concentration insufficient to drive ring closure with rU_{12}* or rA_{12}* and varying the concentration of primary site ligands, we found that ring closure can be driven by sufficiently high concentrations of dC_5, dC_{15} or dC_{75} (**Figure 3.5E-F**). These data indicate that primary site occupancy favors binding of both optimal (poly-pyrimidine) and non-optimal (poly-purine) RNAs to Rho's secondary site.

DISCUSSION

Ring-type helicases and translocases control numerous essential biological processes (Etlinger and Goldberg, 1977; Roberts, 1969). Although generally capable of standalone motor activity, multiple lines of evidence are increasingly highlighting complex mechanisms by which these enzymes are regulated. Protein-protein and protein-ligand interactions, small molecule effectors, and post-translational modifications have all been implicated in controlling disparate types of ring-ATPases; how a majority of these factors exert their various stimulatory or antagonistic effects is generally not well understood.

In the present study, we discovered that bicyclomycin, one of the few small molecule agents known to inhibit a ring-type motor, acts on its target – the Rho transcription termination factor – by sterically blocking a conformational change from an open- to a closed-ring state. Rho ring closure, which the work described in Chapter 2 shows is dependent upon both ATP and secondary-site bound RNA, triggers RNA strand engagement and rearranges residues in the ATP binding pocket into a hydrolysis-competent state. By comparing a bicyclomycin-bound, open-ring conformation of Rho to other structures of Rho sub-states, we found that the drug-binding pocket collapses upon ring closure (**Figure 3.1**). Follow-up SAXS studies show that bicyclomycin does not simply occlude binding of a catalytic water as previously proposed (Skordalakes et al., 2005), but instead directly impedes nucleotide-dependent closure of the Rho ring (**Figures 3.2 and 3.3**). Using a FA-based assay to track ring closure *in vitro*, we further observed that bicyclomycin acts on pre-opened Rho rings, and has little effect on ring topology after ring closure (**Figure 3.4**). Since inhibition of secondary-site RNA binding is indicative of bicyclomycin binding to Rho (Magyar et al., 1996), these results strongly indicate that bicyclomycin is unable to bind to pre-closed Rho rings. Thus, bicyclomycin appears to inhibit Rho by stabilizing a conformation that is incapable of ATP hydrolysis and incompatible with stable RNA binding to its motor regions.

In the course of investigating bicyclomycin's effects on Rho, we also found that binding of nucleic acids to Rho's primary RNA recruitment site promotes the switching of Rho from an open- to closed-ring state. This finding is in accord with prior studies showing that Rho's ATPase activity can be stimulated by primary site occupancy (Richardson, 1982), and that Rho's ATPase rate is in part regulated by an RNA- and ATP-dependent conformational change (Jeong et al., 2004), presumably into a closed-ring state (Chapter 2 and (Thomsen and Berger, 2009)). The inclusion of polypyrimidine DNAs, which bind selectively to the primary site and also promote ATPase activity (Richardson, 1982), increased Rho's affinity for both rU_{12} and rA_{12} substrates at nucleotide concentrations that are otherwise too low to stably drive ring closure (**Figure S3.8**). This result suggests that the binding of a *rut* sequence to Rho's primary site, which is formed by an N-terminal OB-fold in the protein (Dolan et al., 1990; Morgan et al., 1985), allosterically promotes ring closure.

We further found that DNAs capable of bridging two (dC_{15}) or six (dC_{75}) primary sites within the Rho hexamer promoted ring closure more robustly than short (dC_5) DNAs capable of only occupying one primary site (**Figure 3.5A-C**). The more robust impact of these longer DNAs is likely due to higher affinity for Rho (**Figure 3.5D**) rather than by a particular signal that might be propagated by a nucleic acid segment bridging multiple primary sites, as all three ligands drive complete ring closure at sufficiently high concentrations (**Figure 3.5E-F**). We also found that ring closure was more sensitive to primary site occupancy with rA_{12} than with rU_{12}, which is of interest because polypurines bind to the secondary site more weakly than polypyrimidines (Galluppi and Richardson, 1980). Collectively, these results suggest that the RNA recruitment sites on Rho work together to count the number and spacing of pyrimidine recruitment motifs in an RNA transcript, and promote ring closure around non-optimal secondary site binding sequences such as purine-rich regions (**Figure 3.6**). Future studies looking at different types of di-pyrimidine patterns on both natural and synthetic substrates will be needed to probe how this interrogation occurs at a molecular level.

In using an auxiliary ligand-binding element to aid activity, Rho joins a growing list of processive ring helicases and translocases that are subject to both intrinsic and extrinsic regulation. These factors include the bacterial DnaB helicase (Strycharska et al., 2013), the MCM2-7 helicase (Barry et al., 2007; Froelich et al., 2014), and the proteasomal Rpt1-6 ATPases (Bashore et al., 2015). In each of these motors, a built-in accessory domain or a dissociable set of adaptor proteins is used to choreograph proper substrate recruitment and engagement with motor activity. How control linkages between the accessory elements and the motor regions are established through long-range binding and allosteric transitions remains a frontier question.

In closing, it is worth noting that there is substantial interest in identifying small molecule approaches to inhibiting ring-ATPases (Deshaies, 2014). Unfortunately, because ATP binding pockets are often well conserved across different motor families, developing such inhibitors has been challenging. Interestingly, Rho is one of the few ring-type motor proteins for which a small molecule with therapeutic properties has been identified. Here we have shown that this agent, bicyclomycin, does not antagonize the chemistry of the ATPase reaction, but instead acts as a conformational inhibitor of ring dynamics, entering a binding pocket that is only accessible during a portion of the Rho RNA loading and translocation cycle. These findings suggest that by using appropriate screens for ring dynamics, coupled with counter screens against general ATPase activity, it may be possible to identify new generations of selective inhibitors with clinical potential.

Materials and Methods

Calculation of bicyclomycin-binding pocket volumes:

Bicyclomycin-binding pocket volumes in various Rho structures were calculated using the program POVME, using a spherical search with an 8 Å radius and default parameters (Durrant et al., 2011). The center of the bicyclomycin binding pocket was defined as the location of the C1 atom of bicyclomycin in Chain B of the structure 1XPO (Skordalakes et al., 2005), and the corresponding site in all other structures as determined through superposition to this site. Waters were removed from all structures before calculation of binding pocket size.

Small-angle X-ray scattering (SAXS) sample preparation:

Rho was purified as described previously (Thomsen and Berger, 2009), concentrated to 60 mg ml^{-1}, and then dialyzed overnight into SAXS collection buffer (150 mM KCl, 50 mM HEPES-NaOH (pH 7.5), 5 mM $MgCl_2$, 5% Glycerol). Aliquots of bicyclomycin (obtained as a generous gift from Dr. Y. Itoh and the Fujisawa Pharmaceutical Co., Ltd., Japan) were generated through the addition of dialysis buffer to the lyophilized drug. Rho samples were then diluted with bicyclomycin aliquots (for drug containing samples) or dialysis buffer (for drug-free samples) and incubated together overnight. RNA (obtained as 250 nmol syntheses of rU_{12} from IDT) was initially suspended in DMPC-treated RNase-free water (MP Biomedicals), supplemented with an equal volume of 2× SAXS buffer (300 mM KCl, 100 mM HEPES-NaOH (pH 7.5), 10 mM $MgCl_2$, 10% Glycerol) immediately before use, and added at a concentration equimolar to that of Rho hexamer. ADP•BeF_3 aliquots were generated using a 1:3:15 ratio of ADP:$BeCl_2$:NaF and diluted with an equal volume of 2× SAXS buffer. Final SAXS samples were distributed into an Axygen 96-well PCR plates, and obtained through the sequential addition of 8 µl of Rho/bicyclomycin, 8 µl of RNA, and finally 8 µl of ADP•BeF_3 after a two-hour wait. Plates were then flash frozen in liquid nitrogen and stored at −80°C.

Small-angle X-ray scattering (SAXS) data collection and analysis:

Frozen plates were thawed in water, spun for 1 min at 1000 × g, and data collected at 12°C using an automated, high-throughput system at BL 12.3.1 at the Advanced Light Source (Dyer et al., 2014). Buffer-subtracted curves were generated using an automated Ogre script provided by ALS BL12.3.1 (Greg Hura). We focused our analysis on data from 1 s exposures of 2.5 mg ml^{-1} Rho samples because the buffer-subtracted curves showed the least evidence of radiation damage or aggregation under these conditions. Data scaling, Guinier analyses and P(r) calculations were conducted using the *ScÅtter* software suite (Rambo, 2015). MES calculations were conducted using the FOXS web server (Schneidman-Duhovny et al., 2010a), using default parameters with the open (PDB ID 1PV4 (Skordalakes and Berger, 2003)) and closed (PDB ID 5JJI, Chapter 2) ring Rho structures provided as inputs. Unbuilt regions of the open and closed ring structures were modeled in using alignments with Rho structures in which these regions were visualized. All plots were generated using the Prism 6 software suite.

Flourescence Anisotropy (FA)-based RNA binding data collection and analysis:

Rho was concentrated to 60 mg ml^{-1} and dialyzed overnight into SAXS buffer. DNA oligonucleotides (dC_5, dC_{15} and "dC_{75}" ($dC_{15}TC_{15}TC_{15}TC_{15}TC_{15}$)) utilized as primary site ligands were purchased from IDT, were resuspended in the SAXS buffer used for dialysis, and concentrations were calculated based on A_{260} readings and extinction coefficients provided by IDT. For assays, Rho was diluted with SAXS buffer in the presence or absence of DNA, and left to incubate for >1 h on ice. 5′ 6-Fluorescein amidite-labeled rU_{12} and rA_{12} oligonucleotides were purchased from IDT, resuspended in MilliQ-purified water to 100 µM and stored as small aliquots at −80° C until used. Rho was pre-incubated with fluorescein-labeled oligonucleotides, BSA and DTT (20 nM, 0.5 mg ml^{-1} and 5 mM final concentrations respectively) in SAXS buffer on ice for at >30 min, and then combined with ADP•BeF_3 in SAXS buffer for >30 min before reading (referred to herein as assembly of closed rings). In the case of pre/post bicyclomycin incubations, Rho was either incubated with bicyclomycin for 1 h before switching to conditions that favor the ring closure ("pre-incubation"), or Rho was incubated with bicyclomycin for >2h after ring closure was triggered ("post-incubation"). All FA measurements were made at 30° C using a Biotek Neo2 plate reader in Corning 384-well low volume plates (#3544). Baseline corrections (where appropriate) and averaging of duplicates into single points were calculated using a custom Python script; graphs were created in the Prism 6 software suite, and fits generated by nonlinear regression to one-site binding models (with the exception of dC_{75}*, which fit notably better to a two-site model – here, both one- and two-site fits are shown on the corresponding plot).

Figures

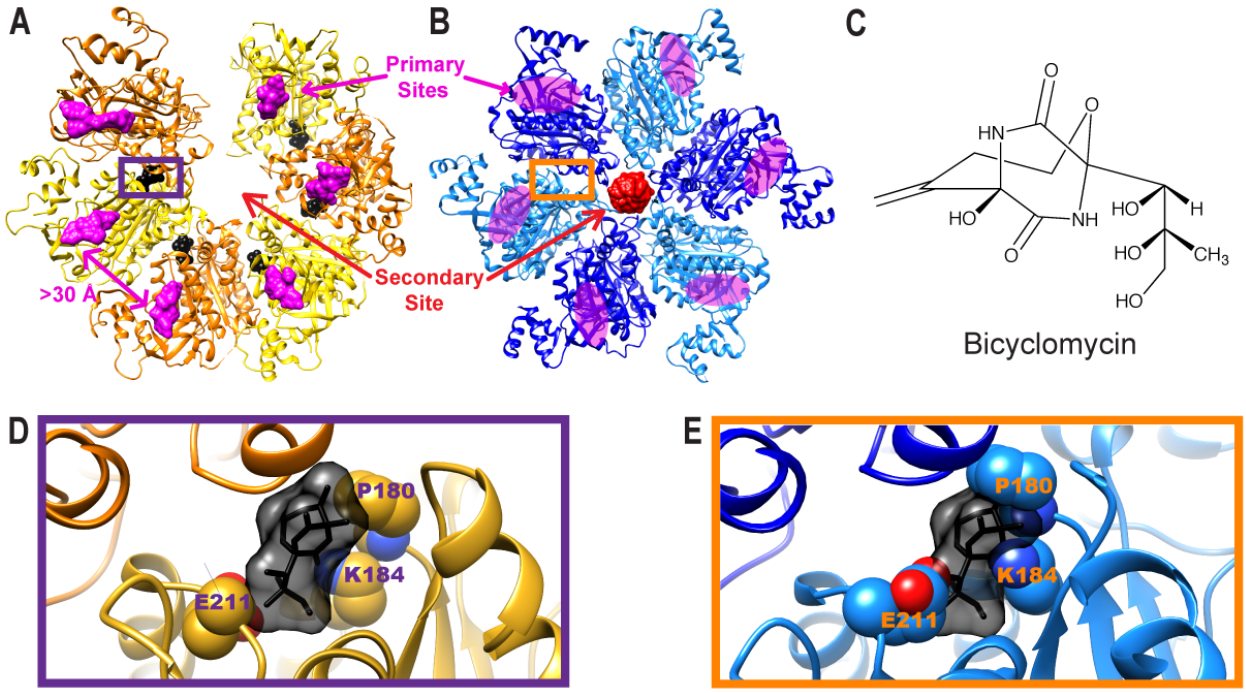

Figure 3.1. Bicyclomycin is an inhibitor of the Rho helicase. **(A)** Crystal structure of open-ring and bicyclomycin-bound Rho (PDB ID: 1XPO (Skordalakes et al., 2005)). Rho subunits are alternatingly colored yellow and orange for distinction, bicyclomycin is black, and primary site RNAs are magenta. **(B)** Crystal structure of closed-ring and translocation-competent Rho (PDB ID: 3ICE (Thomsen and Berger, 2009)). Rho subunits are alternatingly colored cyan and dark blue, and RNA bound in the secondary site is red. Magenta-colored ovals denote the location of the primary sites that are not occupied in this crystal form. **(C)** Chemical structure of bicyclomycin. **(D)** Close-up view of the bicyclomycin-binding pocket shows that bicyclomycin nestles into a small pocket between Rho subunit interfaces. **(E)** Modeling of bicyclomycin into the closed-ring Rho structure shows clear steric clashes between the drug and Rho.

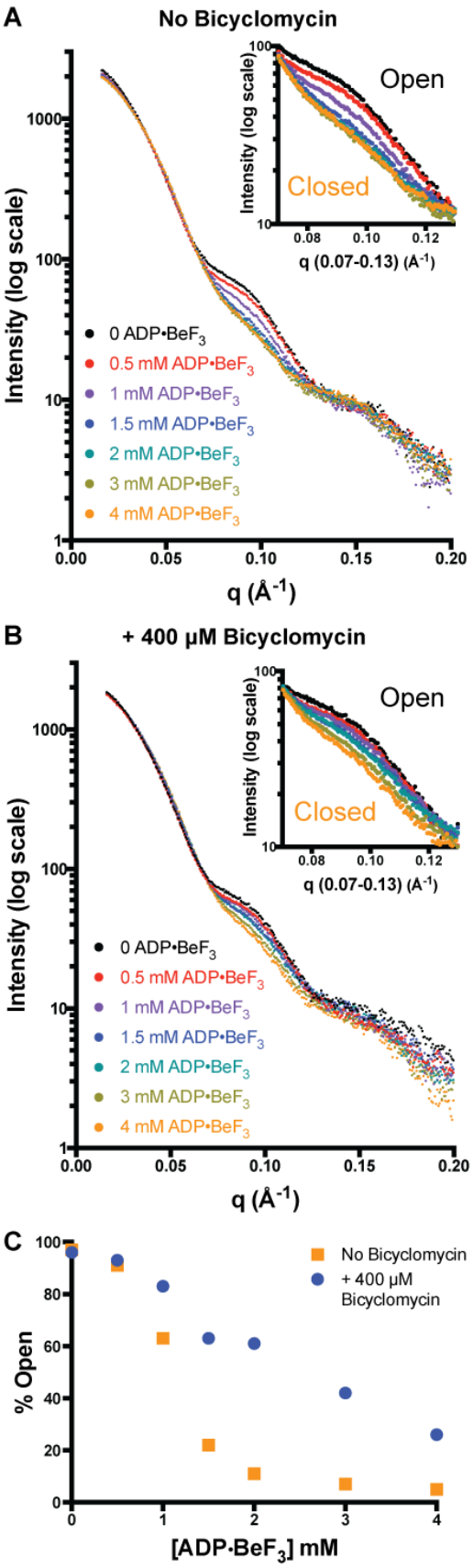
A
No Bicyclomycin
Open
Closed
Intensity (log scale)
q (0.07-0.13) (Å-1)
0 ADP•BeF3
0.5 mM ADP•BeF3
1 mM ADP•BeF3
1.5 mM ADP•BeF3
2 mM ADP•BeF3
3 mM ADP•BeF3
4 mM ADP•BeF3
q (Å-1)
B
+ 400 µM Bicyclomycin
C
No Bicyclomycin
+ 400 µM Bicyclomycin
% Open
[ADP•BeF3] mM

Figure 3.2. Bicyclomycin inhibits Rho ring closure over a range of nucleotide concentrations. **(A)** Scaled SAXS intensities of Rho in the presence of one rU_{12} RNA per hexamer and a variety of concentrations of the ATP mimetic, ADP•BeF_3. **(Inset)** Zoom-in of the mid-q region, in which differences between the various SAXS curves are most pronounced. **(B)** Scaled SAXS intensities of Rho pre-incubated with 400 μM bicyclomycin, and then combined with one rU_{12} RNA per hexamer and a variety of concentrations of ADP•BeF_3. **(Inset)** Zoom-in of the mid-q region reveals a skew of the SAXS curves towards the open state in the presence of bicyclomycin. **(C)** Quantification of the curves represented in **(A)** and **(B)** by MES shows a clear skew towards the open ring state in the presence of bicyclomycin.

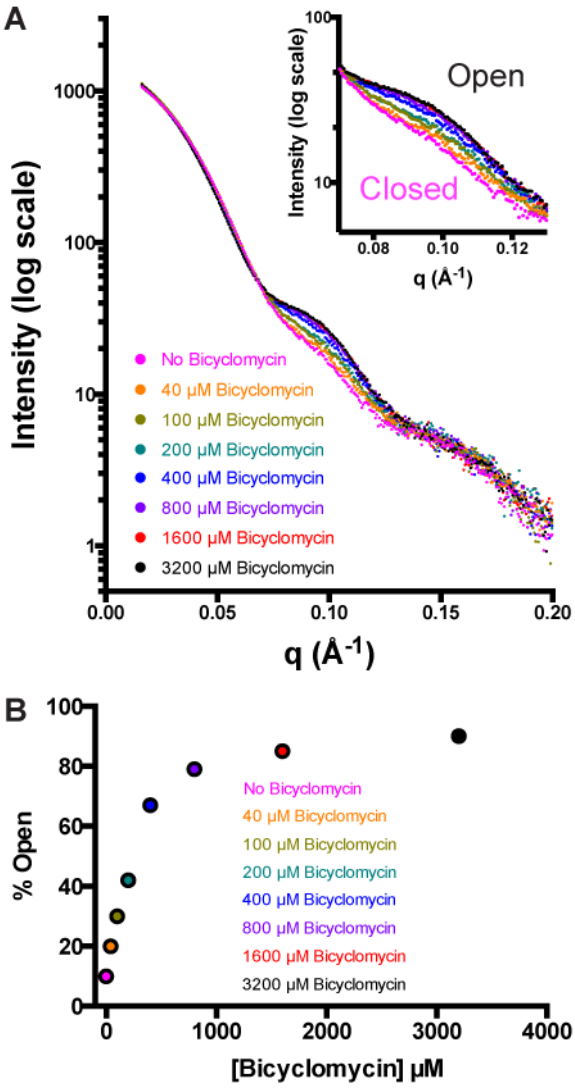

Figure 3.3. Rho ring state can be controlled by varying bicyclomycin concentration. **(A)** Scaled SAXS intensities of a variety of bicyclomycin concentrations, prior to adding one rU_{12} RNA per Rho hexamer and 1.5 mM ADP•BeF_3. **(Inset)** Zoom-in of the mid-q region shows that ring state can be titrated from closed to open by increasing bicyclomycin concentration. **(B)** Quantification of the curves in **(A)** shows a clear titration of Rho ring state that is dependent upon bicyclomycin concentration.

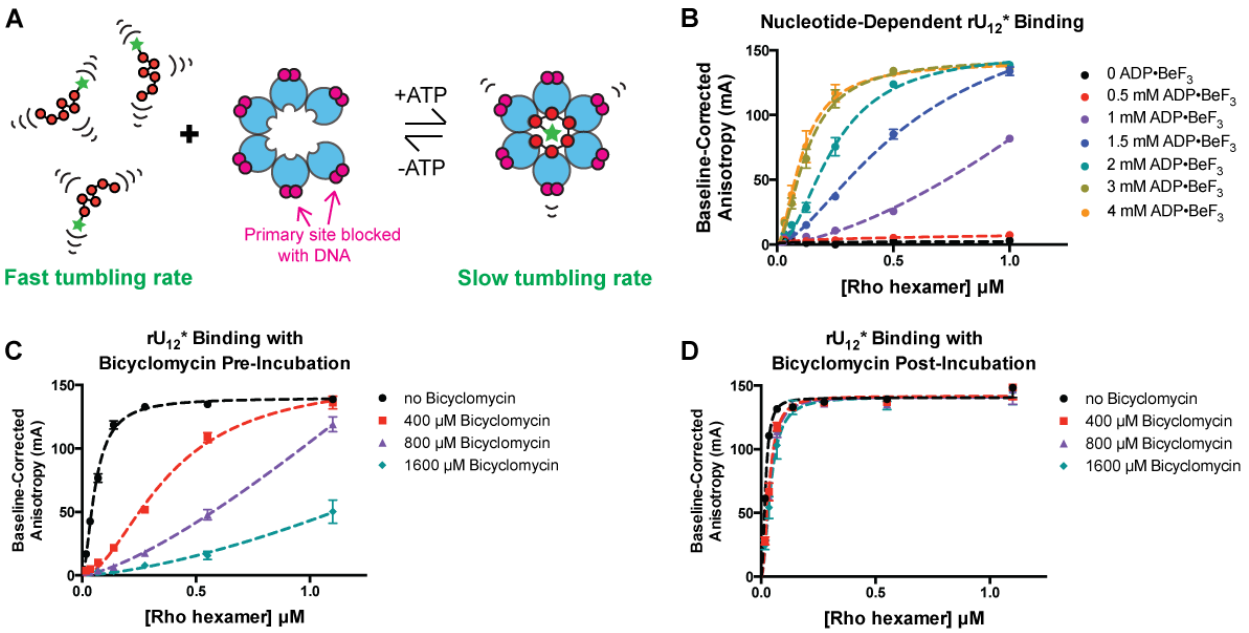

Figure 3.4. A Fluorescence Anisotropy-based RNA binding assay to track Rho ring closure *in vitro*. **(A)** Schematic illustrating how the tumbling rate of a short fluorescein-tagged RNA is slowed by capture within the Rho ring. The Rho primary sites were blocked with short DNAs, represented in magenta. **(B)** Binding of rU_{12}* to Rho is dependent upon ADP•BeF_3 concentration. **(C)** Pre-incubation of Rho with bicyclomycin before addition of RNA and ADP•BeF_3 demonstrates bicyclomycin concentration-dependent inhibition of rU_{12}* binding. **(D)** Incubation of bicyclomycin with a pre-closed Rho ring has little effect on rU_{12}* binding.

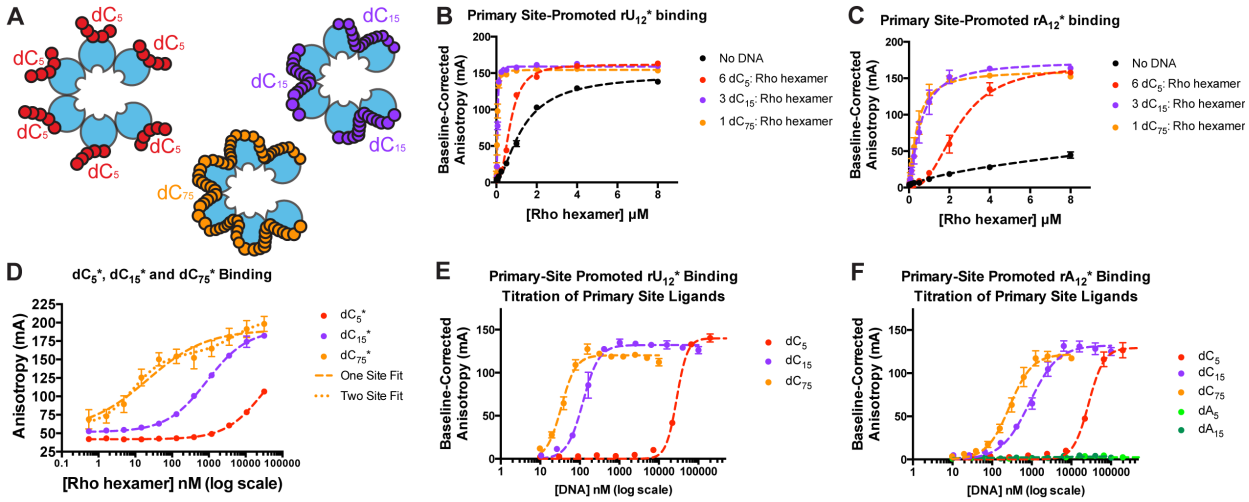

Figure 3.5. Primary site occupancy promotes Rho ring closure. **(A)** Cartoon representation of how dC_5, dC_{15} and dC_{75} could occupy one, two, or six primary sites, respectively. **(B)** Binding of rU_{12}* to Rho is promoted by polycytidylic DNAs of various lengths. **(C)** Binding of rA_{12}* to Rho is promoted by polycytidylic DNAs of various lengths. **(D)** Longer DNAs (dC_{15}* and dC_{75}*) capable of bridging multiple primary sites progressively promote secondary site binding more effectively than does a short DNA (dC_5*) that can only occupy one primary site. **(E)** Binding of rU_{12}* can be fully driven at saturating concentrations of dC_5, dC_{15} or dC_{75}. **(F)** Binding of rA_{12}* can be fully driven at saturating concentrations of dC_5, dC_{15} or dC_{75}. Polypurine DNAs (dA_5 and dA_{15}, light and dark green, respectively), which do not bind Rho's primary site, do not facilitate Rho ring closure.

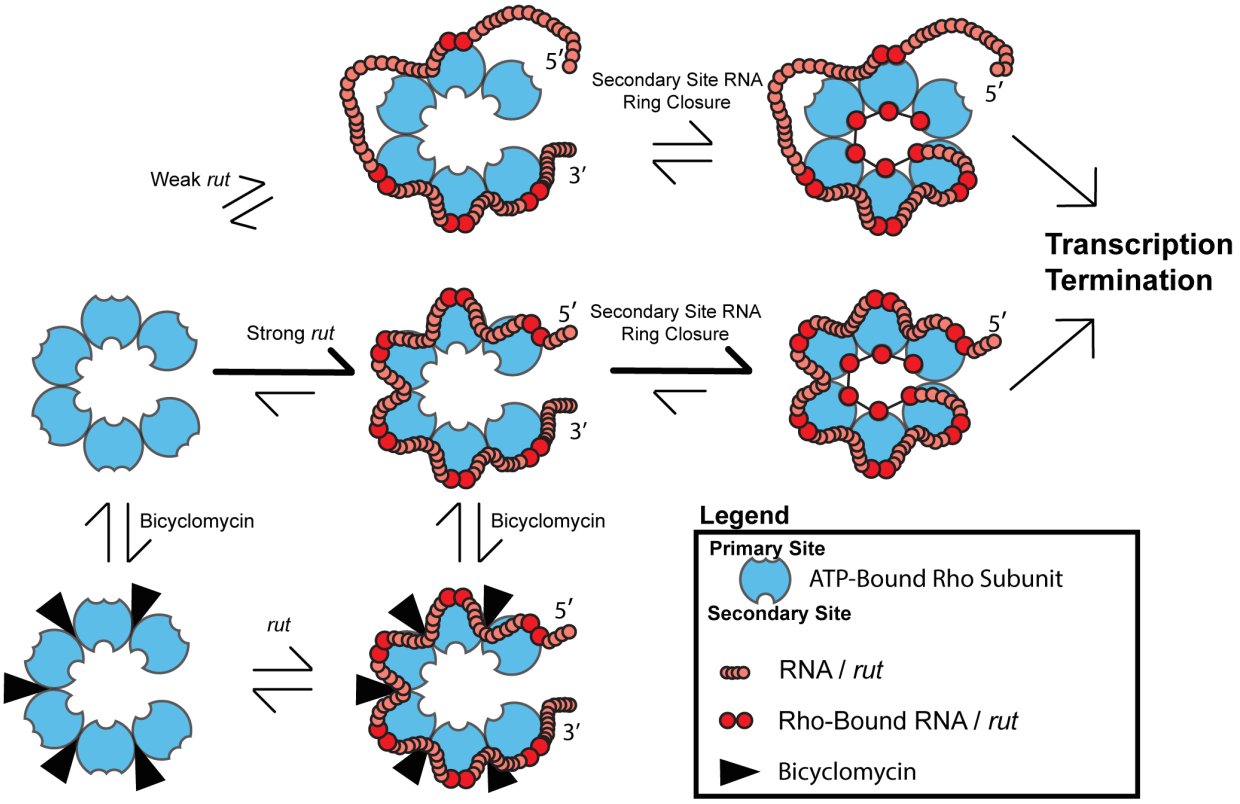

Figure 3.6. Model for how bicyclomycin and primary site ligands regulate Rho ring closure, and thereby the capacity of the helicase to promote transcription termination. After a *rut*-containing RNA first binds to the Rho primary sites and subsequently the secondary site, the Rho ring snaps shut to trap the RNA within the central pore. This conformational change to a translocation-competent closed ring state is less efficient in the presence of a weakly binding *rut* sequence with non-optimal di-pyrimidine number/spacing (top), and is completely blocked by bicyclomycin (bottom, black wedges). Open-ring Rho must bind to the *rut* sequence via primary site interactions and close before its ATPase activity can promote 5′ to 3′ RNA translocation and transcription termination.

Supplementary Material

Tables

Table S3.1: Volume of BCM binding pocket in a variety of Rho structures calculated using the program POVME.

Structure	Interface	Volume (Å^3)	Ring State	BCM bound?	Oligo bound?
1XPO	A / B	272	Open	Yes	1° Site RNA
1PVO	A / B	345	Open	No	1° Site RNA
3ICE	A / B	56	Closed	No	2° Site rU_{12}
3JJI	A / B	55	Closed	No	2° Site rU_{12}
3JJL	A / B	53	Closed	No	2° Site rU_{12}
3JJK	A / B	60	Closed	No	2° Site rA_{12}

Table S3.2: FOXS fitting statistics of SAXS data represented in **Figure 3.2**.

[ADP•BeF_3] mM	[BCM] µM	Open Fit χ	Closed Fit χ	Ensemble χ	% Open (MES)	% Closed (MES)
0	0	4.08	14.16	4.05	97	3
0.5	0	3.10	13.26	2.85	91	9
1	0	5.87	8.41	4.00	63	37
1.5	0	8.60	3.11	2.05	22	78
2	0	9.65	2.11	1.73	11	89
3	0	10.07	1.79	1.63	7	93
4	0	10.23	1.83	1.75	5	95
0	400	3.08	13.04	3.02	96	4
0.5	400	2.36	12.85	2.15	93	7
1	400	3.30	11.73	2.38	83	17
1.5	400	5.16	8.17	2.40	63	37
2	400	5.55	8.09	2.49	61	39
3	400	6.97	5.37	2.60	42	58
4	400	5.95	2.76	1.90	26	74

Table S3.3: Data statistics of SAXS curves shown in **Figure 3.2**.

[ADP•BeF_3] mM	[BCM] µM	Scaling Factor	R_g (reciprocal)	R_g (real)	r_{ave}	d_{max}
0	0	1.45	48.9	47.5	62.1	145
0.5	0	1.29	48.4	47.5	62.1	142
1	0	1.16	47.5	46.1	60.4	141
1.5	0	1.1	47.1	46.1	60.3	141
2	0	1.01	46.7	45.4	59.5	141
3	0	1.04	46.6	45.6	59.7	140
4	0	1	46.5	45.6	59.7	139
0	400	1.67	48.8	48.2	62.9	148
0.5	400	1.34	48.5	47.3	61.9	141
1	400	1.2	48.1	47	61.5	142
1.5	400	1.21	47.5	46.9	61.4	141
2	400	1.16	47.3	46.7	61	141
3	400	1.05	47.1	46.2	60.5	140
4	400	1	46.9	46	60.2	140

Table S3.4: FOXS fitting statistics of SAXS data represented in **Figure 3.3**.

[ADP•BeF_3] mM	[BCM] µM	Open Fit χ	Closed Fit χ	Ensemble χ	% Open (MES)	% Closed (MES)
1.5	0	22.41	4.03	3.22	10	90
1.5	40	19.09	5.53	3.05	20	80
1.5	100	15.75	7.50	3.62	30	70
1.5	200	12.52	9.24	3.27	42	58
1.5	400	8.12	15.09	3.84	67	33
1.5	800	5.68	16.67	3.54	79	21
1.5	1600	5.01	19.93	3.72	85	15
1.5	3200	4.11	20.21	3.51	90	10

Supplementary Table 5: Data statistics of SAXS curves shown in **Figure 3.3**.

[ADP•BeF_3] mM	[BCM] µM	Scaling Factor	R_g (reciprocal)	R_g (real)	r_{ave}	d_{max}
1.5	0	0.829	46.3	45	59	138
1.5	40	0.869	46.6	45.5	59.5	138
1.5	100	1.01	46.8	45.7	59.7	141
1.5	200	1.05	46.8	45.5	59.5	138
1.5	400	1	47.5	46.4	60.6	142
1.5	800	1.09	47.9	46.8	61.2	136
1.5	1600	0.964	47.9	46.6	61.0	136
1.5	3200	1	48.2	46.6	61.0	144

Figures

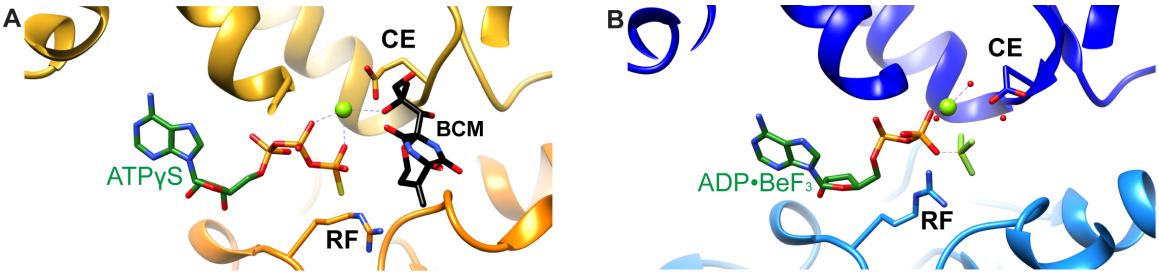

Figure S3.1: Binding sites of nucleotide and bicyclomycin in open- and closed-ring Rho structures; CE = Catalytic Glutamate, RF = Arginine Finger, BCM = bicyclomycin. **(A)** A representative interface from bicyclomycin-bound and open-ring Rho structure (PDB ID: 1XPO) shows that bicyclomycin binds proximal to the gamma phosphate of the non-hydrolyzable nucleotide analog ATPγS. Neither the catalytic glutamate nor the arginine finger are properly positioned to promote ATP hydrolysis. **(B)** A representative interface from a hydrolysis- and translocation-competent Rho structure (PDB ID: 3ICE) shows both the catalytic glutamate and the arginine finger are optimally positioned to promote ATP hydrolysis.

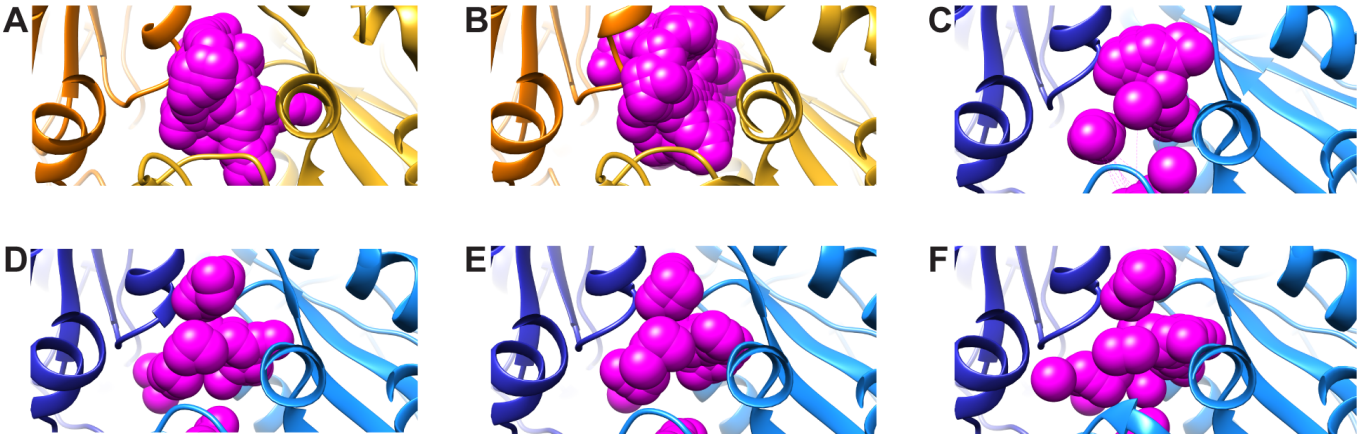

Figure S3.2: Visual representation of pocket sizes as determined by POVME (37). Each pink sphere represents 1 $Å^3$. Corresponding PDB codes: **(A)** 1XPO, **(B)** 1PVO, **(C)** 3ICE, **(D)** 3JJI, **(E),** 3JJL, **(F)** 3JJK. Yellow/gold – open-ring Rho; dark/light blue – closed-ring Rho.

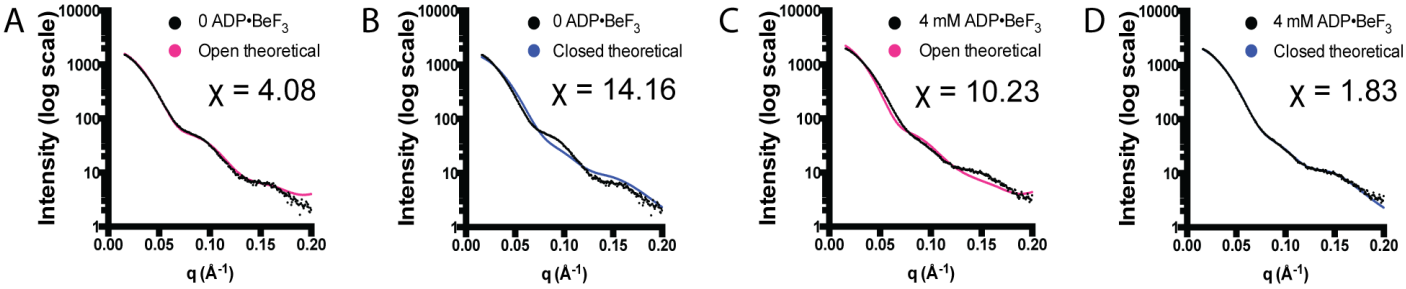

Figure S3.3: Comparison of experimental SAXS curves to theoretical curves calculated from crystallographic models (PDB ID for open: 1PV4, PDB ID for closed: 2JJI) using the FOXS server (31). **(A)** In the presence of RNA and in the absence of nucleotide, the observed curve is a close match to the theoretical curve calculated from an open-ring Rho structure. **(B)** The experimental SAXS curve observed in the presence of RNA and in the absence of nucleotide is a poor match to a closed-ring Rho structure. **(C)** In the presence of RNA and 4 mM ADP•BeF_3, the observed SAXS curve differs substantially from the theoretical open-ring curve. **(D)** The experimental curve observed in the presence of RNA and 4 mM ADP•BeF_3 is a close match to the theoretical closed-ring.

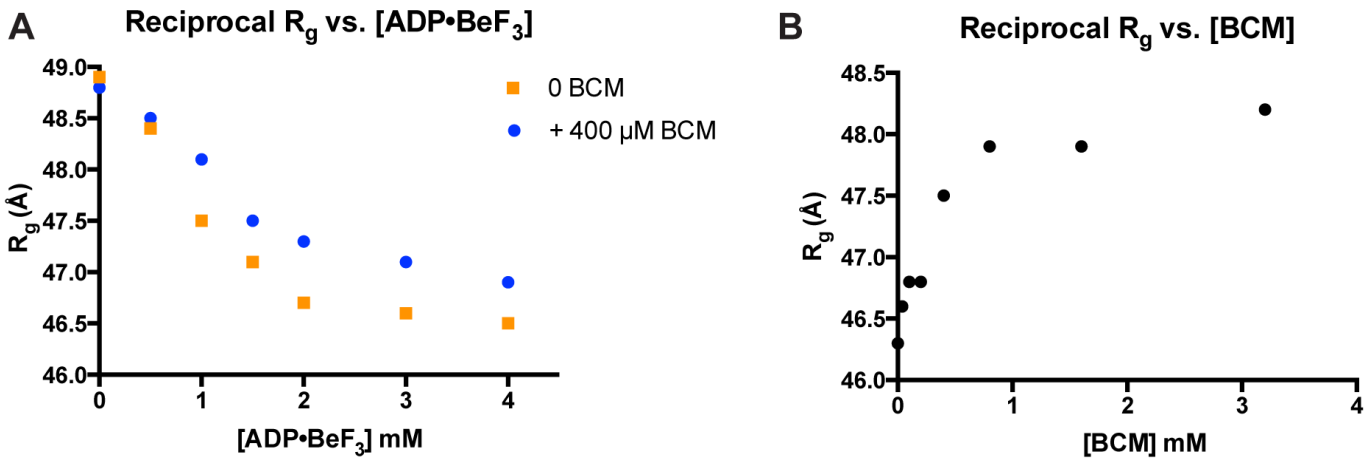

Figure S3.4: Reciprocal R_g values observed for SAXS data represented in **Figure 3.2** **(A)** and **Figure 3.3** **(B)**.

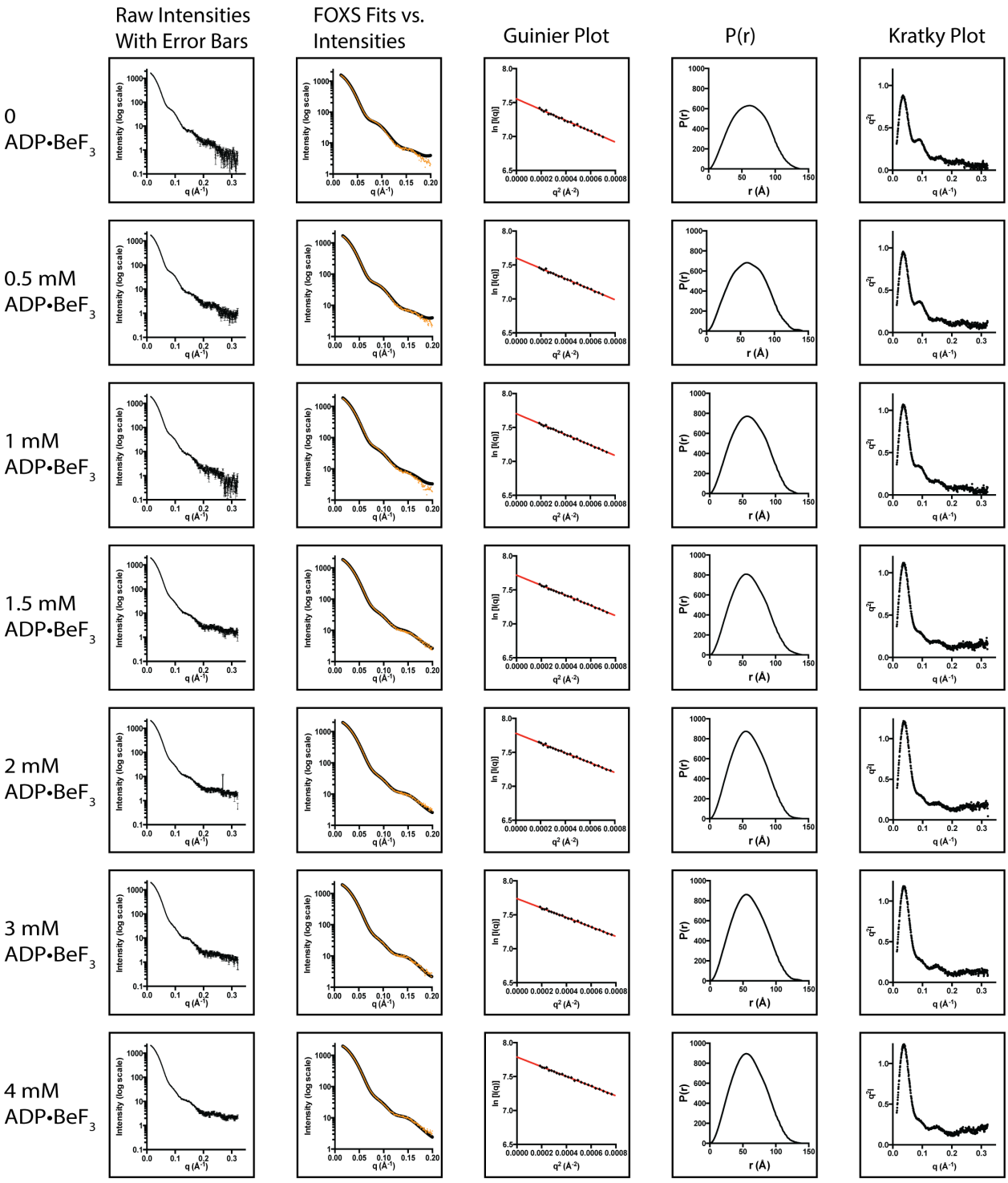

Fig. S3.5: Raw Intensitites with Error Bars, FOXS Fits vs. Intensities, Guinier Plots, P(r) Distributions and Kratky Plots for all data shown in **Figure 3.2a**.

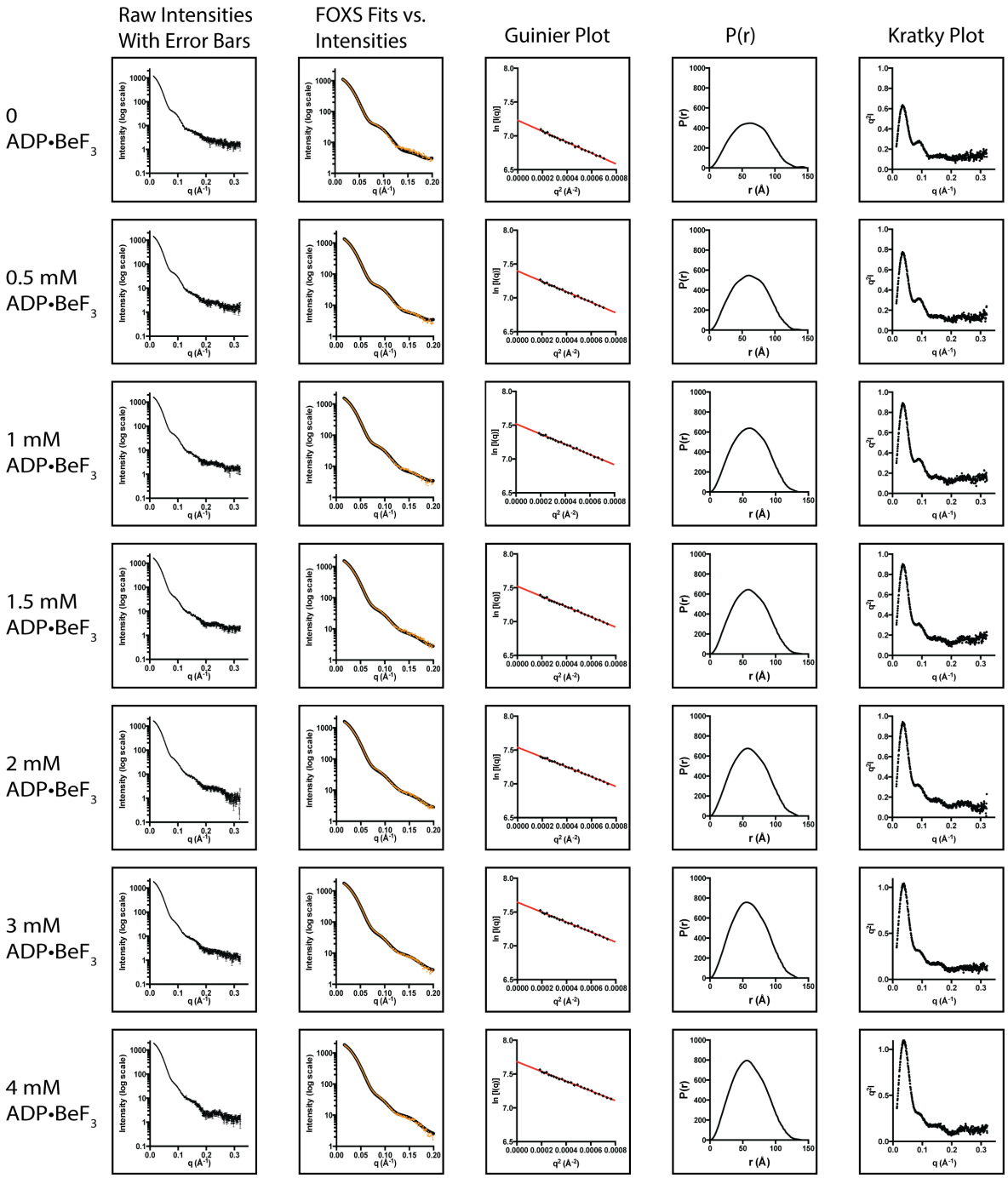

Figure S3.6: Raw Intensitites with Error Bars, FOXS Fits vs. Intensities, Guinier Plots, P(r) Distributions and Kratky Plots for all data shown in **Figure 3.2b**.

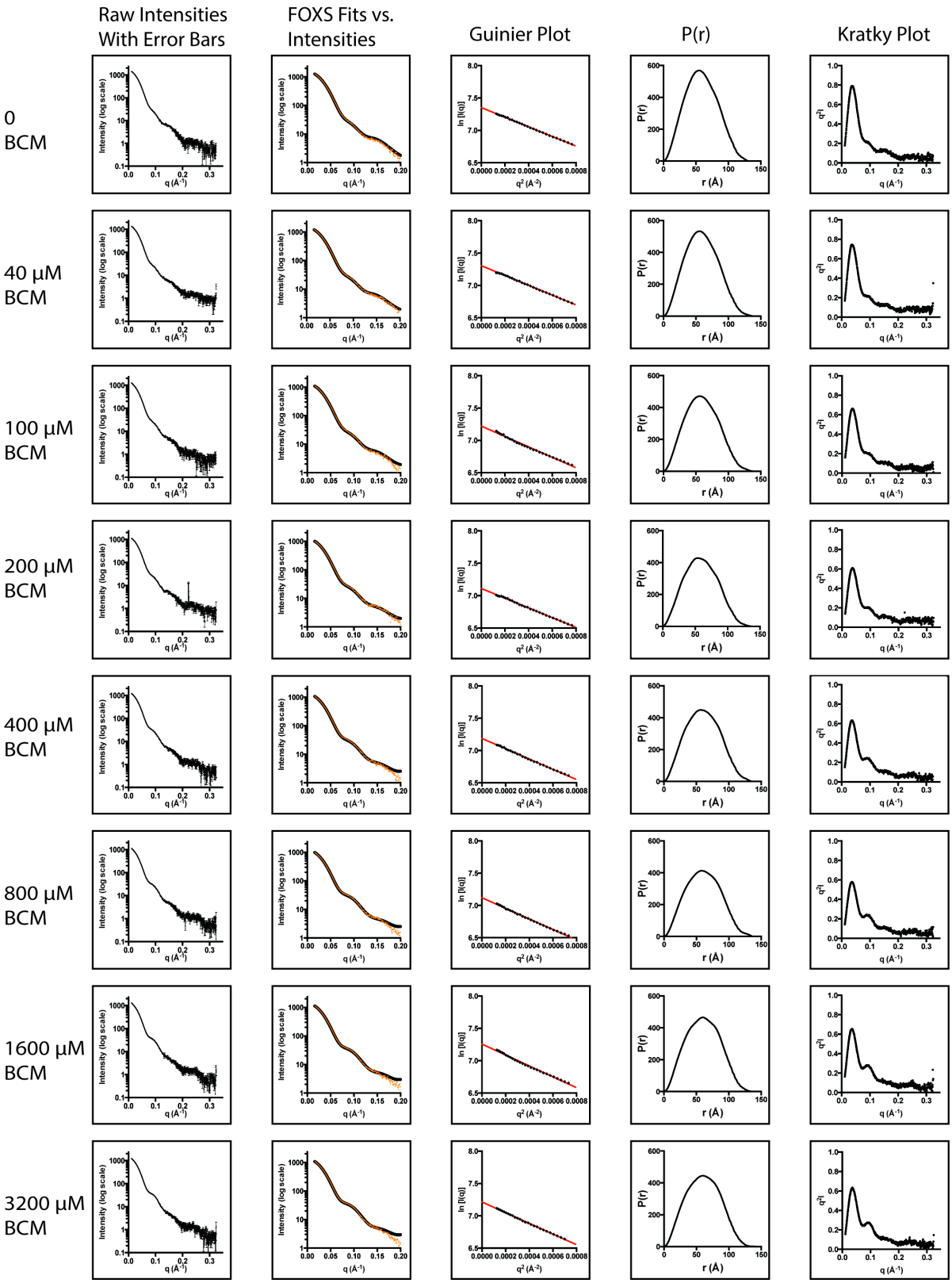

Figure S3.7: Raw Intensitites with Error Bars, FOXS Fits vs. Intensities, Guinier Plots, P(r) Distributions and Kratky Plots for all data shown in **Figure 3.3a**.

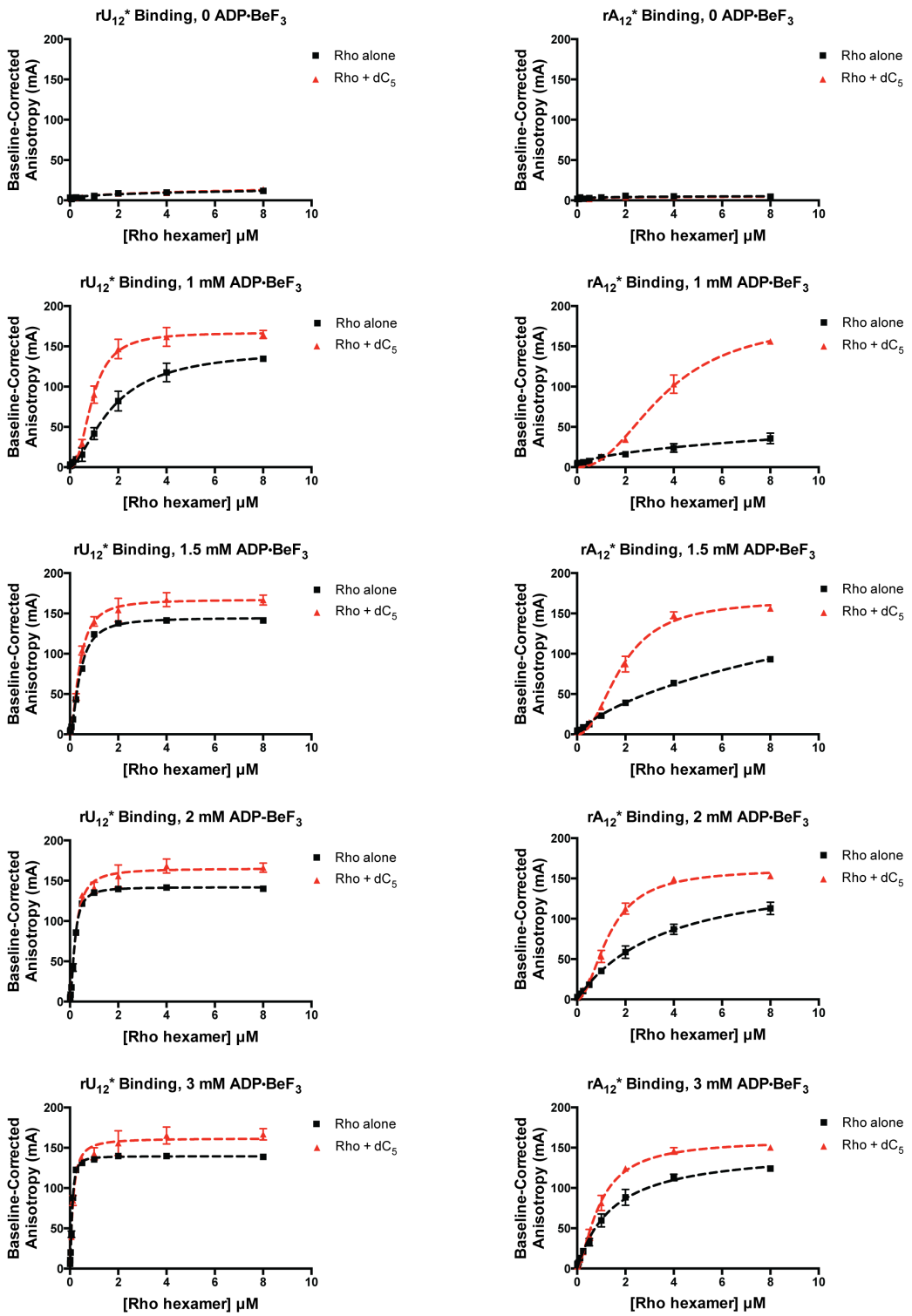

Figure S3.8: RNA binding by fluorescence anisotropy in the presence and absence of a short primary site ligand (dC_5). Little to no RNA binding is observed in the absence of nucleotide (0 ADP•BeF_3) with either rU_{12}* or rA_{12}*. RNA binding is assisted by dC_5 in cases where weak RNA binding affinity is observed due to low nucleotide concentration (1 mM ADP•BeF_3 with both rU_{12}* and rA_{12}*) or a sub-optimal RNA sequence (rA_{12}* curves with 1-3 mM ADP•BeF_3).

Chapter 4: Mechanism for the regulated control of transcription termination by a universal adapter protein

Introduction

The NusG and Spt5-family of proteins are the only transcription factors that are conserved across all living organisms (Tomar and Artsimovitch, 2013). The N-terminus of NusG/Spt5 proteins universally adopt a NGN (NusG N-terminal) domain that binds to RNA polymerase. These factors also contain variable, flexibly-attached C-terminal Kyrides-Onzonis-Woese ("KOW") domains, which in bacteria serve in part as a recruitment site for the Rho transcription terminator (Burmann et al., 2010). Although NusG is known to assist Rho-dependent termination of transcripts that bear non-optimal (i.e. cytosine-poor) *rut* sequences (Peters et al., 2012), the mechanism for this activity has remained unclear. To investigate the role of the NusG CTD in regulation of Rho-dependent termination, we determined a crystal structure of Rho in complex with the NusG CTD. Biochemical studies additionally revealed NusG promotes RNA- and ATP-dependent ring closure by Rho in the absence of any other factors. Please see Chapter 1 for an extensive introduction to this subject.

Results

Architecture of a Rho/NusG complex

To facilitate study of the interaction between Rho and NusG, we first determined which domains of NusG are necessary for this association (see **Figure 4.1A**). To this end, we cloned and purified full length NusG (residues 1-181, herein referred to as "NusG FL"), along with the isolated NTD (1-119) and CTD (125-181), all of which were appended with a cleavable His-MBP-TEV tag on their N-termini. Subsequent Ni^{2+} affinity-based pull-downs with purified *E. coli* Rho (obtained as described previously (Thomsen and Berger, 2009)) revealed that the NusG CTD was both necessary and sufficient to bind Rho (**Figure S4.1A**), and that the NusG NTD does not contribute to formation of the complex. We therefore proceeded with the isolated NusG CTD for all subsequent structural studies.

Although previous work by von Hippel and colleagues had suggested that only one NusG could bind to each Rho hexamer (Pasman and von Hippel, 2000), it was unclear how such an asymmetry could be enforced given that there are six potential binding sites for the protein on a Rho hexamer. Similarly, it was uncertain whether NusG's affinity for Rho could be affected by the conformational status of the Rho ring. During the course of crystal screening and optimization, I discovered that the key factors required for obtaining suitable crystals of a Rho/NusG complex incuded using a 18-fold excess of NusG CTD per Rho hexamer (i.e., 3 copies of NusG for every Rho monomer) and co-crystallizing the complex with ligands that favor formation of closed ring Rho in solution (specifically, a short, 12-mer poly-U RNA and the non-hydrolyzable nucleotide analog ADP-BeF_3, as described in Chapter 2). Addition and refinement of cryoprotectants eventually yielded crystals that grew in the space group P1 (see Methods for a detailed description of space group determination) and that diffracted to ~3.8 Å resolution following flash freezing in liquid nitrogen. Interestingly, a similar diffraction limit was observed with native (unfrozen) shots of Rho collected from a free-electron laser X-ray source (**Figure S4.1B**). This result indicates that the crystal form I obtains is more likely diffraction-limited by intrinsic disorder, rather than by defects that introduced from radiation damage or cryo-preservation.

Initial phases for determining the structure were obtained by molecular replacement (using only a Rho monomer as a search model). Inspection of the resulting solution indicated that the P1 unit cell contains two Rho hexamers. Density was clearly evident for RNA in the secondary site of each hexamer and for one ADP-BeF_3 moiety associated with every Rho monomer (**Figure S4.1C**) (the RNA was eventually modeled in three alternate positions, due to overlap of the Rho hexamer axis with crystallographic symmetry axes). Although initial maps indicated the likely presence of one NusG CTD per Rho monomer, unequivocal placement of NusG into the maps proved to be particularly challenging due to poor electron density for the domain (**Figure S4.1D**). A published NMR model of the NusG CTD (PDB ID: 2JVV) was eventually placed into density-modified simulated annealing composite omit maps by way of a 6D search in the apparent NusG density using the program ESSENS (**Figure 4.1B**, see Methods). To verify the orientation of the NusG CTD relative to Rho, single wavelength

anomalous dispersion (SAD) data were collected from crystals grown with selenomethionine derivatives of NusG, which revealed a close match between the expected location of a native methionine of NusG at position 133 and a peak that was evident in LLG anomalous difference maps (**Figure 4.1C**, see methods). To further verify this placement, I conducted another experiment using an engineered mutant of NusG bearing two methionines (the native M133 and an engineered one, Q177M); LLG anomalous difference maps again displayed parity between the expected locations of the two residues and the observed peaks (**Figure 4.1D**, see methods).

Following building and refinement, the Rho-NusG-rU_{12}-ADP•BeF_3 model converged with data at an R_{work} and R_{free} of 26.21% and 29.06%, respectively (**Table 4.1**). Six NusG CTDs could be seen to clearly associate with each Rho hexamer (see **Figure 4.1E**); 2Fo-Fc density and B-factors indicated that these correspond to three high and three low-occupancy interactions, consistent with the trimer-of-dimers configuration observed for Rho, which manifest in an alternating pattern around the ring (**Figure S4.1E**). Interestingly, a similar pattern of alternatingly high and low B-factors is in the Rho subunits and their associated ADP-BeF_3 moieties (**Figure S4.1F**). Inspection of the Rho/NusG interface (**Figure 4.1F**) reveals that NusG utilizes two loops, spanning residues 139-144 (herein referred to as "L1") and 164-167 ("L2"), which protrude off the β-barrel to engage Rho. The Rho portion of the interface involves the two α helices located on Rho's C-terminus that include residues 381-382 and 415-417. Similar to all other crystal forms of Rho observed to date (See Chapter 2 and (Skordalakes and Berger, 2003; Skordalakes et al., 2005; Thomsen and Berger, 2009)), the two most C-terminal residues (R418 and S419) of Rho could not be modeled.

Molecular dynamics simulations reveal details of the energetics that drive association of Rho and NusG

The poor quality of the maps in the region of the Rho/NusG interface led to ambiguity in the assignment of side chain rotamers. As a consequence, it was unclear which residues in the proximity of the interface might make meaningful contributions to the formation of the complex. I therefore turned to microsecond-scale Molecular Dynamics ("MD") simulations, in collaboration with Wen Ma and Klaus Schulten, as a means of exploring the energetics of this protein:protein interaction. The RMSD of NusG proved very stable (~5 Å) with respect to time over the course of the MD simulation (**Figure 4.2A**). Similarly encouraging was the observation that the surface area buried by the Rho/NusG interface, which is approximately 750 $Å^2$, also remained stable throughout the simulation (**Figure 4.2B**).

During the course of the MD simulation, three pairs of adjacent residues on NusG appeared to be particularly stably associated with Rho. The most stable of these contacts include I164 and F165 of NusG, which appear to nest into a hydrophobic pocket composed of I382, I386, F398, M415 and M416 on the C-terminal side of the Rho hexamer (see **Figure 4.2C** for a representative frame from the simulation). The D143 carbonyl group of NusG appears to vacillate between forming ionic contacts with the basic groups of Rho's K417 and R418 (the latter of which was not resolved in our structure), while F144 of NusG is flanked by I382 of Rho and the aliphatic base of

K417's side chain (**Figure 4.2D**). Finally, P140 and F141 of NusG appear to both flank W381 and I382 of Rho (**Figure 4.2E**).

Rho ring closure is promoted by association of the NusG CTD

Since Rho ring closure is promoted by association of cytosine-rich sequences with Rho's primary sites (see Chapter 3), and Rho-dependent termination of cytosine-poor *rut* sites (herein referred to as "weak" *rut* sites) is known to require NusG (Peters et al., 2012), we wondered whether NusG might assist Rho-dependent termination in part by promoting Rho ring closure. Such a mechanism help Rho overcome the ring-closing defect seen when a weak *rut* associates with Rho's primary sites (see Chapter 3), and is consistent with the observation that NusG is required for Rho to terminate transcription of the lacZ operon *in vitro* (Burns and Richardson, 1995). To test this hypothesis, I monitored the association of short, fluorescein-labeled RNA with Rho's secondary site to determine whether the presence or absence of NusG could affect the efficiency of Rho ring closure *in vitro* (see diagram in **Figure 4.4A**). Previous work has demonstrated that stable association of either fluorescein-labeled poly-U or poly-A RNAs (herein referred to as rU_{12}* or rA_{12}*, respectively) with Rho's secondary site is strictly dependent upon the inclusion of ligands (e.g., nucleotide) that enable isomerization of Rho to a closed-ring state, allowing Rho ring closure to be tracked by monitoring changes in Fluorescence Anisotropy ("FA") from the labeled RNA (see Chapter 3).

Inclusion of the NusG CTD (at a concentration equimolar to that of Rho hexamer) clearly promoted Rho's affinity for rA_{12}* at several different nucleotide concentrations over the observed affinity of rA_{12}* for Rho alone (evident at 1, 1.5 and 2 mM ADP-BeF_3; see **Figure S4.2**). The NusG CTD similarly stimulated the association of rU_{12}* with Rho at a modest nucleotide concentrations (1 mM) (**Figure S4.2**). The binding of rA_{12}* and rU_{12}* under these conditions was also facilitated by full-length NusG, but not by the NusG NTD, demonstrating that the NusG CTD is both necessary and sufficient to promote Rho ring closure (**Figure S4.2**). No binding of rU_{12}* or rA_{12}* was observed in the absence of nucleotide either with or without NusG (**Figure S4.2**), indicating that the NusG CTD on its own cannot drive nucleotide-independent ring closure.

Having observed that the NusG CTD can aid sealing of Rho ring over a broad swath of conditions, I surmised that high concentrations of NusG might drive Rho ring closure in the absence of any primary site ligands. By fixing the concentration of Rho (1 μM hexamer) and nucleotide (1 mM ADP-BeF_3) at a concentration insufficient to drive ring closure with rU_{12}* or rA_{12}*, and by then conducting titration experiments with my various NusG constructs, I observed that Rho ring closure can indeed be driven by high concentrations of either full-length NusG or the NusG CTD (**Figure 4.3B-C**). The observation that both native NusG and isolated CTD were similarly effective stimulants of Rho ring closure in the presence of either rU_{12}* ($K_{d,app}$ = 2.70 μM and 2.82 μM, respectively) or rA_{12}* ($K_{d,app}$ = 15.92 μM and 22.73 μM, respectively), along with the observation that NusG NTD had no visible effect on Rho ring closure at concentrations as high as 80 μM (**Figure 4.3B-C**, dark red circles), supports the notion that the NusG CTD is solely responsible for promoting Rho ring closure under these conditions. The

Hill coefficients observed for fits to the observed binding data range from ~1.5-1.7, implying that association of only one to two copies NusG are sufficient to drive ring closure by Rho in the absence of primary site ligands. Interestingly, the inclusion of a saturating concentration of a primary site ligand (dC_{15}) yielded fully closed Rho rings in which the NTD, CTD and FL NusG curves were indistinguishable from one another (**Figure 4.3B-C**, diamonds). This result suggests that NusG CTD-promoted Rho ring closure can be masked when primary sites are efficiently occupied.

To experimentally probe the contribution of individual residues to the Rho/NusG interaction, I next tested the ability of selected alanine mutants in the Rho-binding interface of full-length NusG to promote Rho ring closure. NusG-promoted Rho ring closure assays conducted with individual alanine mutants of residues 137-145 shows that mutation of residues G139, F141, D143 and F144 is particularly deleterious for Rho binding (**Figure 4.4D**). Similar testing of individual alanine substitutions within residues 163-167 of NusG suggests that mutation of either I164 or F165 to alanine is highly detrimental to association of the protein with Rho (**Figure 4.4F**). Conversion of the apparent equilibrium constants for the various NusG constructs into $\Delta G_{bind,app}$ values allowed for the calculation of binding free energy changes upon alanine substitution ($\Delta\Delta G_{bind,app}$), which were then mapped onto the structure of the Rho/NusG interface in a color-coded fashion (see **Figure 4.3E** and **Figure 4.3G**). This analysis shows that alanine substitution of NusG residues closest to the Rho/NusG interface yielded the most substantial loss of binding free energies, consistent with the modeling and placement of NusG in the structure.

Discussion

Despite decades of study, the molecular basis by which NusG promotes Rho-dependent termination has remained elusive. To better understand this interaction and the effect on NusG on Rho, I determined a co-crystal structure between the two proteins in the presence of RNA and an ATP mimetic. The structure reveals that NusG engage to the C-terminal face of Rho's motor domain through two loops that protrude off the core β-barrel formed by NusG's C-terminal domain (**Figure 4.1**). Molecular dynamics simulations of the Rho/NusG interface revealed several groupings of residues on NusG that appear to play a key role in driving the association of Rho and NusG (**Figure 4.2**). Alanine substitutions of residues on NusG proximal to the Rho interface allowed for the identification of six amino acids (G139, F141, D143, F144, I164 and F165) that provide particularly significant energetic contributions to the Rho/NusG interface. With the exception of G139 (which may serve a more indirect role in helping to position a loop for productive contacts with Rho), all of these residues were also observed to make stable contacts with Rho throughout the course of the MD simulation. Importantly, by monitoring the effect of NusG on Rho ring status *in vitro*, I uncovered a hitherto unforeseen role of NusG as a positive effector of Rho ring closure (**Figure 4.3B-C**), particularly in instances where the primary RNA binding is incompletely occupied. To our knowledge, this finding represents the first demonstration of a direct impact of NusG on Rho activity *in vitro* in the absence of RNA polymerase.

Numerous studies have implicated NusG as a factor that recruits Rho to an RNA polymerase that is destined for transcription termination (Mooney et al., 2009b; Nehrke et al., 1993; Sullivan and Gottesman, 1992). By combining the Rho/NusG structure with a variety of other published structures, I generated a tentative model for a ternary complex of two proteins and RNA polymerase (**Figure 4.4A-B**). The linker located between the NusG NTD and CTD is expected to comprise a flexible connection (Mooney et al., 2009b), providing freedom for Rho to adopt different orientations with respect to RNA polymerase. The NusG CTD has been proposed sample a sphere of sterically-allowed positions centered on the N-terminus of the linker, with a radius dictated by the length of the linker (**Figure 4.4A**, red dashed circle; model initially described by Cramer and colleagues (Martinez-Rucobo et al., 2011)). Association of Rho with an RNA polymerase complex through the NusG CTD would presumably allow Rho to pivot within a spherical region centered on the same location as that of the NusG CTD, but with a substantially larger radius owing to the size of the hexamer (**Figure 4.4A**, pink circle). Interestingly, within this larger sphere movement, Rho can sample regions in close proximity to the RNA exit channel of RNA polymerase. This close proximity could help drive the association of Rho with weak *rut* sites by increasing the effective local concentration of terminator relative to a nascent transcript.

Considering NusG as a factor that merely recruits Rho to weak *rut* sites fails to account for an observation demonstrating that Rho-dependent transcription termination on a weak *rut* cannot be driven *in vitro* simply by increasing the concentration of Rho in the absence of NusG (Burns and Richardson, 1995). This finding is noteworthy, as it suggests that NusG can accelerate at least one step in Rho's catalytic cycle. My observation that NusG directly promotes Rho ring closure around a substrate RNA

when Rho's primary site is incompletely occupied provides a molecular rationale for how NusG could accelerate passage through a kinetic barrier to promote efficient termination. Although the report that NusG does not impact Rho ATPase rate *in vitro* (Nehrke et al., 1993) might at first glance appear to contradict this model, I discovered that saturation of Rho's primary sites – which has been universally carried out in prior studies by others – can mask the effect of NusG on Rho ring closure (**Figure 4.3B-C**). Addition work will be needed to determine whether NusG promotes Rho ATPase activity (as well as ring closure) in conditions where Rho's primary sites would not be saturated with polypyrimidine nucleic acids.

The recent discovery that binding of pyrimidine-rich nucleic acids to Rho's primary sites promotes ring closure suggests that the probability that Rho acts upon a transcript that bears either a strong or weak *rut* sequence is controlled by at least two steps (see scheme in **Figure 4.4C**). In the first step, Rho engagement of a transcript by the binding of pyrimidine-cytosine (YC) motifs to Rho's primary sites would be governed by the thermodynamic favorability of Rho associating with appropriate dipyrimidine elements located within a given *rut*. In the second step, the rate of formation of a closed-ring and translocation-competent state would depend upon both the abundance and arrangement of YC motifs within the *rut*. This latter step may dictate how Rho discriminates between a strong vs. weak *rut* beyond what would be expected from differences in YC binding affinities alone. The binding of Rho to NusG may boost Rho's activity on weak *rut* sequences by raising the local concentration of Rho in the region proximal to the transcript, and may also increase the probability that Rho can form a closed-ring state in which Rho's primary sites are inefficiently occupied. Future studies will be needed to understand how the association of Rho with the NusG CTD is blocked by other transcription factors such as NusE, and how the exchange of these regulatory factors on the NusG CTD controls conditional Rho-dependent termination at both weak and cryptic *rut* sites.

Materials and Methods

Protein crystallization and data collection:

Closed-ring Rho was assembled as described previously (Thomsen and Berger, 2009), and Rho-NusG-rU_{12}-ADP•BeF_3 complexes (20 mg/mL final protein concentration) were assembled by combining closed-ring Rho complexes with a three-fold molar excess of NusG CTD per Rho monomer. Crystals were grown at 291K in Qiagen 15-well easyXtal plates by combining 2 µL of Rho-NusG-rU_{12}-ADP•BeF_3 with 2 µL of well solution (95 mM Potassium Glutamate, 50 mM Tris pH 8, 1 mM Spermine, 0.5% PEG 8k). Crystals were visible after an hour and grew to full size within one week. Crystals were harvested with Hampton CryoLoops after gradually introducing (through numerous, small pipetting steps) 30% 1,2-butanediol into the drop, and crystals were then flash frozen in liquid nitrogen. Diffraction data was collected at 90K at ALS BL 8.3.1.

Structure solution and refinement:

All data was processed using XDS (Kabsch, 2010). Ambiguity in space group assignment (parallel processing in P1, P3 and P6 yielded reasonable merging statistics in all three cases, but substantially higher R factors in subsequent refinement of data in P3 or P6 relative to P1) led us to assign the space group as P1 and flag an R-free set using thin resolution shells (Adams et al., 2010). Molecular replacement was conducted using PHENIX with a monomer from the $Rho^{PolyU\text{-}P21}$ structure as the search model (see Chapter 2). Although maps displayed obvious density for both RNA within Rho's central pore and ADP•BeF_3, density for NusG was initially quite poor (see **Figure S4.1**). Simulated-annealing composite omit maps were calculated in CNS (Brunger, 2007), FOMs were estimated using SigmaA (Winn et al., 2011), and density modified maps were generated with multi-domain NCS averaging using the program dmmulti (Cowtan and Main, 1993; Winn et al., 2011). Initial placement of NusG was initially achieved through fitting of a poly-alanine derivative of a published NusG CTD model (PDB ID 2JVV, (Mooney et al., 2009b)) into a region of the maps that bared clear resemblance to the NusG CTD by way of a 6D search with the program ESSENS (Kleywegt and Jones, 1997). NusG was subsequently replaced with a Rosetta energy-minimized derivative (generated with assistance from Roland Pache and Tanja Kortemme) that contained side chains, as refinement of a poly-alanine model consistently yielded ~1-2% higher R-factors. The final model was refined to an R_{work}/R_{free} of 26.21%/29.06%, and Ramachandran statistics were 97.5% favored, 2.5% allowed and 0.0% outliers as reported by MOLPROBITY (Davis et al., 2007).

Fluorescence Anisotropy (FA)-based RNA binding data collection and analysis:

Rho and NusG (NTD, CTD or FL) were concentrated to 60 and 10 mg mL^{-1} respectively, and dialyzed overnight into assay buffer (150 mM KCl, 20 mM HEPES-NaOH, 5 mM $MgCl_2$, 5% Glycerol, 0.5 mM TCEP pH 7.5). Oligonucleotides used as primary (dC_{15}) or secondary (rU_{12}* or rA_{12}*, which were 5′ 6-Fluorescein amidite-labeled, resuspended to 100 µM in MilliQ-purified water and stored at -80° C until use) site ligands were purchased from IDT. For assays, Rho was diluted with buffer and left to incubate with NusG constructs on ice for at least 1 hour. Rho was pre-incubated with fluorescein-labeled oligonucleotides, BSA and DTT (20 nM, 0.5 mg ml^{-1} and 5 mM final

concentrations respectively) in SAXS buffer on ice for at >30 min, and then combined with ADP•BeF3 in assay buffer for >30 min before reading. All FA measurements, data processing and generation of plots were conducted as described previously (Chapter 3).

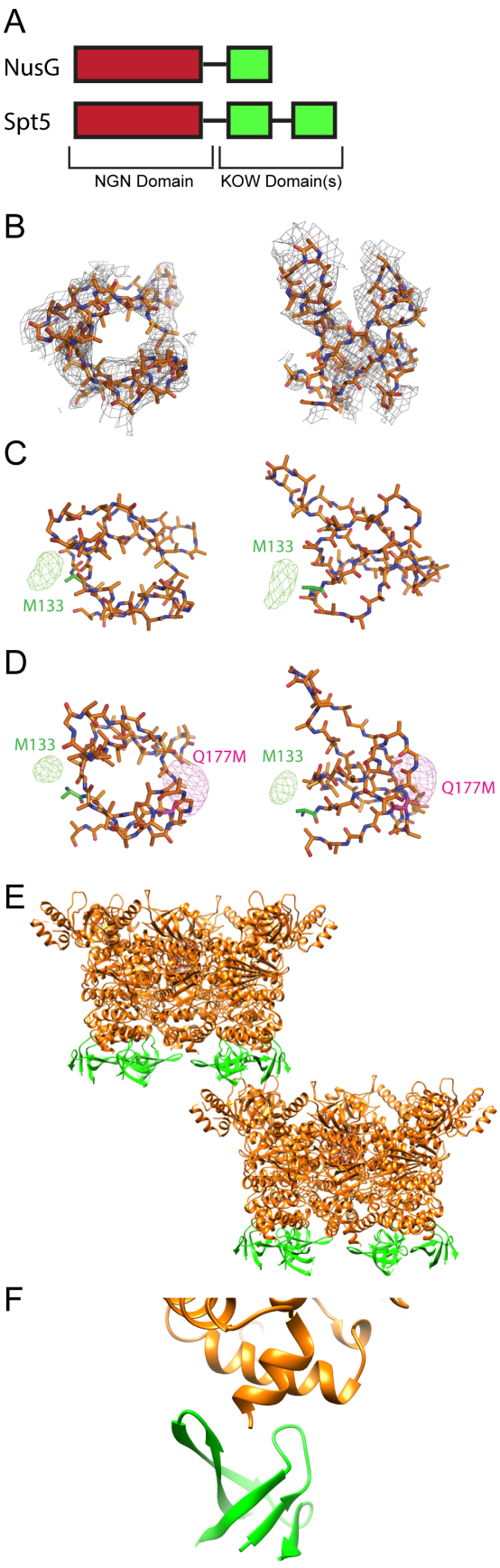
A
NusG
Spt5
NGN Domain
KOW Domain(s)
B
C
M133
M133
D
M133
Q177M
M133
Q177M
E
F

Figure 4.1: Crystal structure of closed-ring Rho in complex with NusG. **(A)** All NusG/Spt5-like proteins contain at least one NGN and one KOW domain. Some Eukaryotic Spt5s contain multiple KOW domains. Not pictured is the "mini" domain that is found within the NGN domain of some bacterial NusGs. **(B)** NusG was docked into simulated annealing composite omit maps using the solution from a 6D search of the maps with the program ESSENS. **(C)** LLG anomalous difference maps obtained from crystals grown with selenomethionine-substituted NusG CTD display a clear peak proximal to the native (M133) methionine. **(D)** LLG anomalous difference maps obtained from crystals grown with a selenomethionine-substituted mutant of NusG CTD (Q177M) display two clear peaks, one close the native (M133) methionine and another near the engineered site (Q177M). **(E)** The asymmetric unit of this Rho/NusG crystal form contains two closed-ring Rho hexamers and twelve copies of NusG (one associated with each Rho monomer). **(F)** Close-up of a Rho/NusG interface observed in this crystal form.

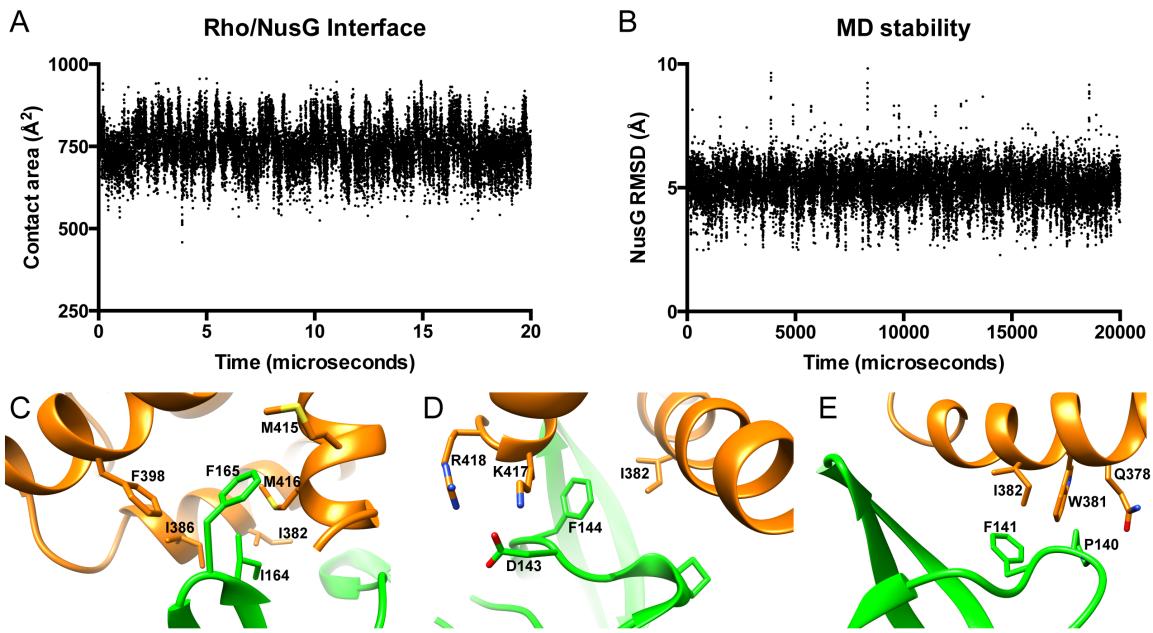

Figure 4.2: Molecular dynamics simulations of the Rho/NusG interface. **(A)** Contact area of the Rho/NusG interface is stable with respect to time throughout the simulation. **(B)** RMSD of NusG (Å) is stable with respect to time throughout the simulation. **(C)** A representative frame from the simulation shows I164 and F165 of NusG nest into a hydrophobic pocket on Rho. **(D)** A representative frame shows D143 and F144 of NusG contact Rho. **(E)** A representative frame shows P140 and F141 of NusG contact Rho.

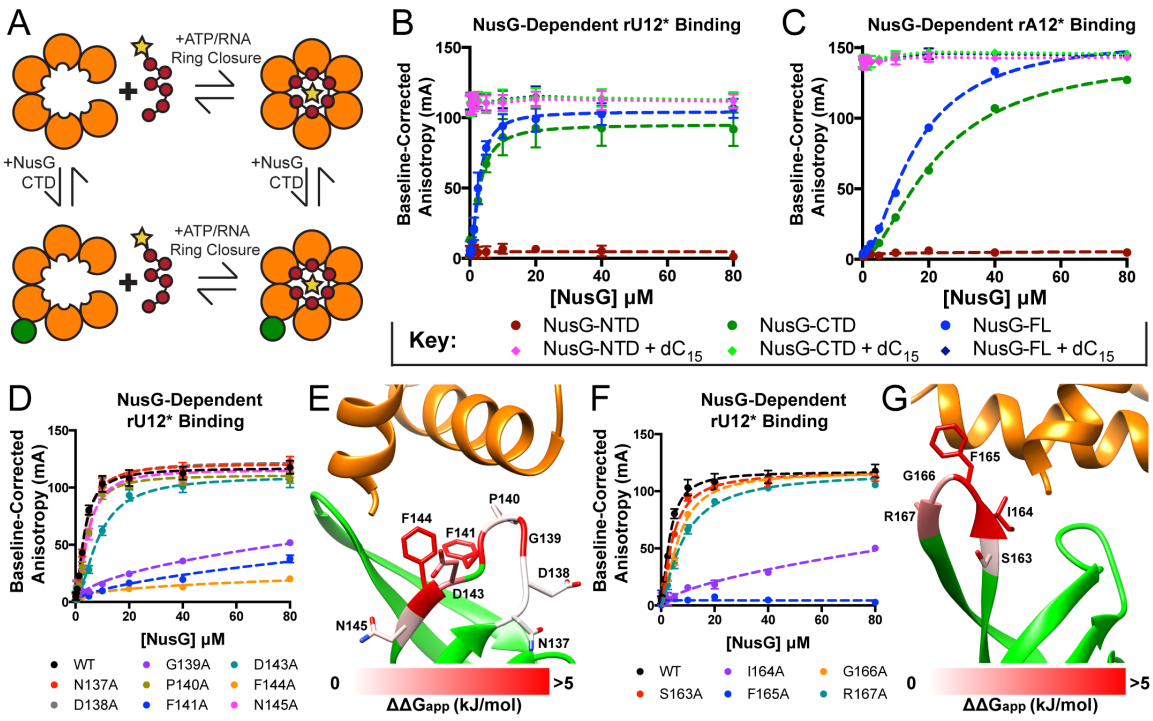

Figure 4.3: NusG promotes Rho ring closure. **(A)** Schematic of ring closure assay used to determine the impact of association of NusG on Rho ring closure. **(B-C)** High concentrations of NusG CTD (dark green circles) or FL (blue circles) are sufficient to drive binding of rU_{12}* or rA_{12}* to Rho. NusG NTD has no impact on rU^{12}* or rA^{12}* binding. The effect of NusG on Rho ring closure is masked by a high concentration of primary site ligands (pink, green and dark blue diamonds). **(D)** Individual alanine substitutions of NusG residues 137-145 reveal that mutations of F144, F141 or G139 to alanine are particularly deleterious for association of NusG with Rho. **(E)** Binding defects observed upon alanine substitution of NusG residues 137-145 plotted onto the Rho/NusG interface. White delineates no observed binding defect upon alanine substitution at that position, and darker red denotes a large defect. **(F)** Individual alanine substitutions of NusG residues 163-167 reveal that mutations of I164 or F165 are particularly deleterious for association of NusG with Rho. **(G)** Binding defects observed upon alanine substitution of NusG residues 163-167 plotted onto the Rho/NusG interface. White delineates no observed binding defect upon alanine substitution at that position, and darker red denotes a large defect.

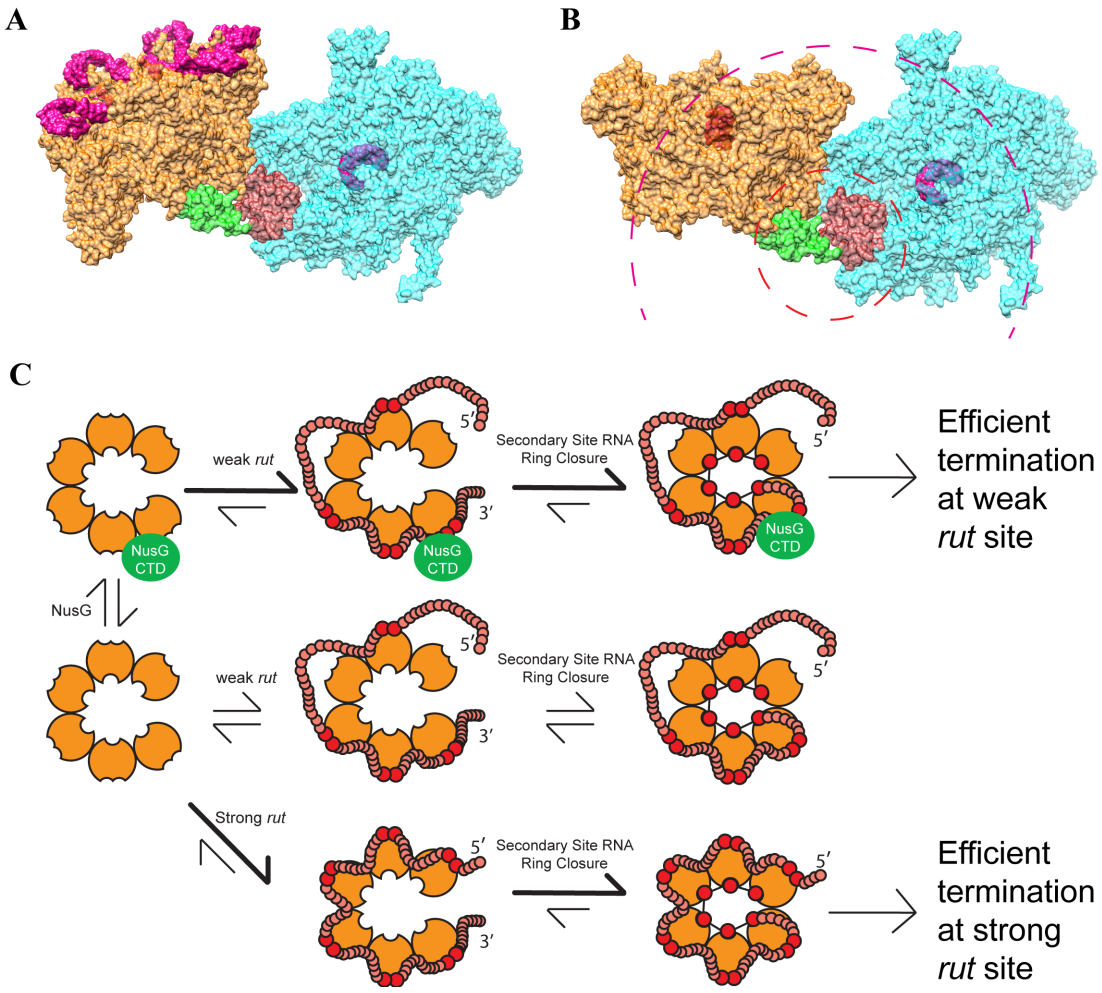

Figure 4.4: Model of how NusG promotes Rho-dependent termination. **(A)** A model based on superposition of a variety of published structures ((Knowlton et al., 2003; Martinez-Rucobo et al., 2011; Skordalakes and Berger, 2003; Weixlbaumer et al., 2013; Zuo et al., 2013)) in conjunction with the Rho/NusG structure described here allows for the modeling of a ternary complex of open-ring Rho (with a modeled *rut* shown in pink), NusG, and RNA polymerase. The nascent RNA strand is modeled in pink. **(B)** NusG would promote RNA-dependent ring closure by Rho, trapping the nascent strand in Rho's central pore and allowing Rho to pivot freely proximal to the RNA exit channel. RNA bound to Rho's secondary site is represented in red, and the nascent RNA within the polymerase is shown in pink. **(C)** Kinetic scheme depicting several possible pathways in Rho-dependent termination. In the case of a strong rut sequence (bottom), both *rut* binding and Rho ring closure would be fast, which would lead to a high probability of termination of transcription. In the case of a weak *rut* (middle), both *rut* binding and Rho ring closure would be expected to be slower than with a strong *rut*, leading to a lower probability of termination. A NusG-bound Rho (top) would be expected to both associate with and form a closed-ring state around a weak *rut* more quickly than in the absence of NusG, increasing the probability of termination relative to the probability in the absence of NusG.

Table 4.1. Data collection and refinement.

Data collection	
Space group	P1
Cell dimensions	
a, *b*, *c* (Å)	126.32, 126.35, 166.82
α, β, γ (°)	89.899, 89.942, 60.011
Wavelength	1.116
Resolution (Å)	40-3.40 (3.47-3.40)
$R_{,eas}$	12.9% (268.2%)
$I/\sigma I$	7.19 (0.46)
Completeness (%)	97.1 (89.7)
Redundancy	
Refinement	
Resolution (Å)	39.03-3.7
	(3.83-3.7)
No. reflections	93084 (9314)
R_{work} / R_{free}	26.22% / 29.09%
No. atoms	43668
Protein	43284
Ligand/ion	384
Water	
B-factors	203.09
Protein	203.51
Ligand/ion	156.56
Water	
R.M.S deviations	
Bond lengths (Å)	0.013
Bond angles (°)	1.28

Values in parenthesis correspond to the highest resolution bin

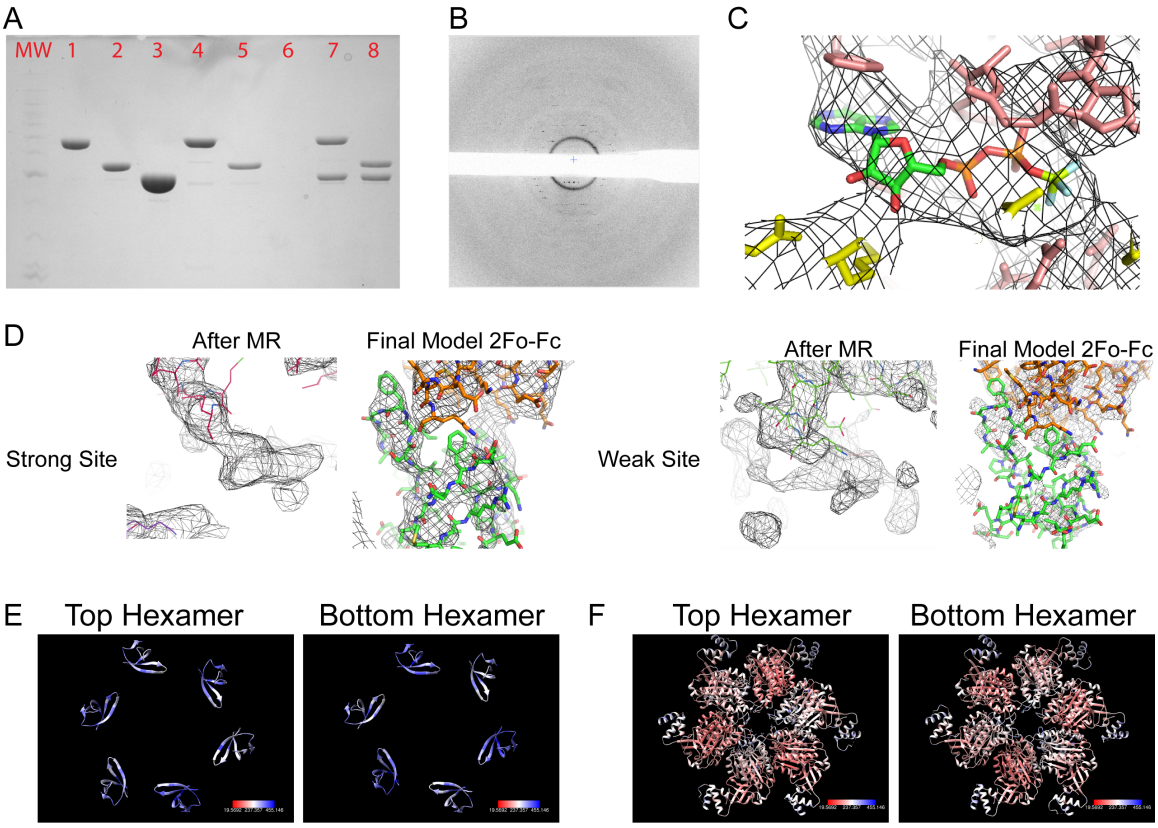

Figure S4.1: Validation of a Rho/NusG crystal structure. **(A)** Ni^{2+} Pull-downs with His-MBP-TEV-tagged NusG constructs reveal that the NusG CTD is necessary and sufficient to bind Rho. Lane 1: NusG FL, Lane 2: NusG CTD, Lane 3: Rho, Lane 4: Ni Elute from NusG FL pull-down, Lane 5: Ni Elute from NusG CTD pull-down, Lane 6: Ni Elute from Rho alone pull-down, Lane 7: Ni Elute from NusG FL + Rho pull-down, Lane 8: Ni Elute from NusG CTD + Rho pull-down. **(B)** Diffraction data collected at a free-electron X-ray source displays a high resolution limit of ~4Å. **(C)** Final $2F_o$-F_c density of a representative ADP•BeF_3 molecule associated with Rho. **(D)** FWT/PHWT observed after Molecular Replacement ("After MR") and final refined $2F_o$-F_c density of a representative NusG CTD from both a "strong" and a "weak" site. **(E)** NusG CTDs bound to both Rho hexamers display an alternating pattern of higher and lower B-factors. **(F)** Rho monomers and their associated ADP•BeF_3 moieties display a similar alternating pattern of higher and lower B-factors.

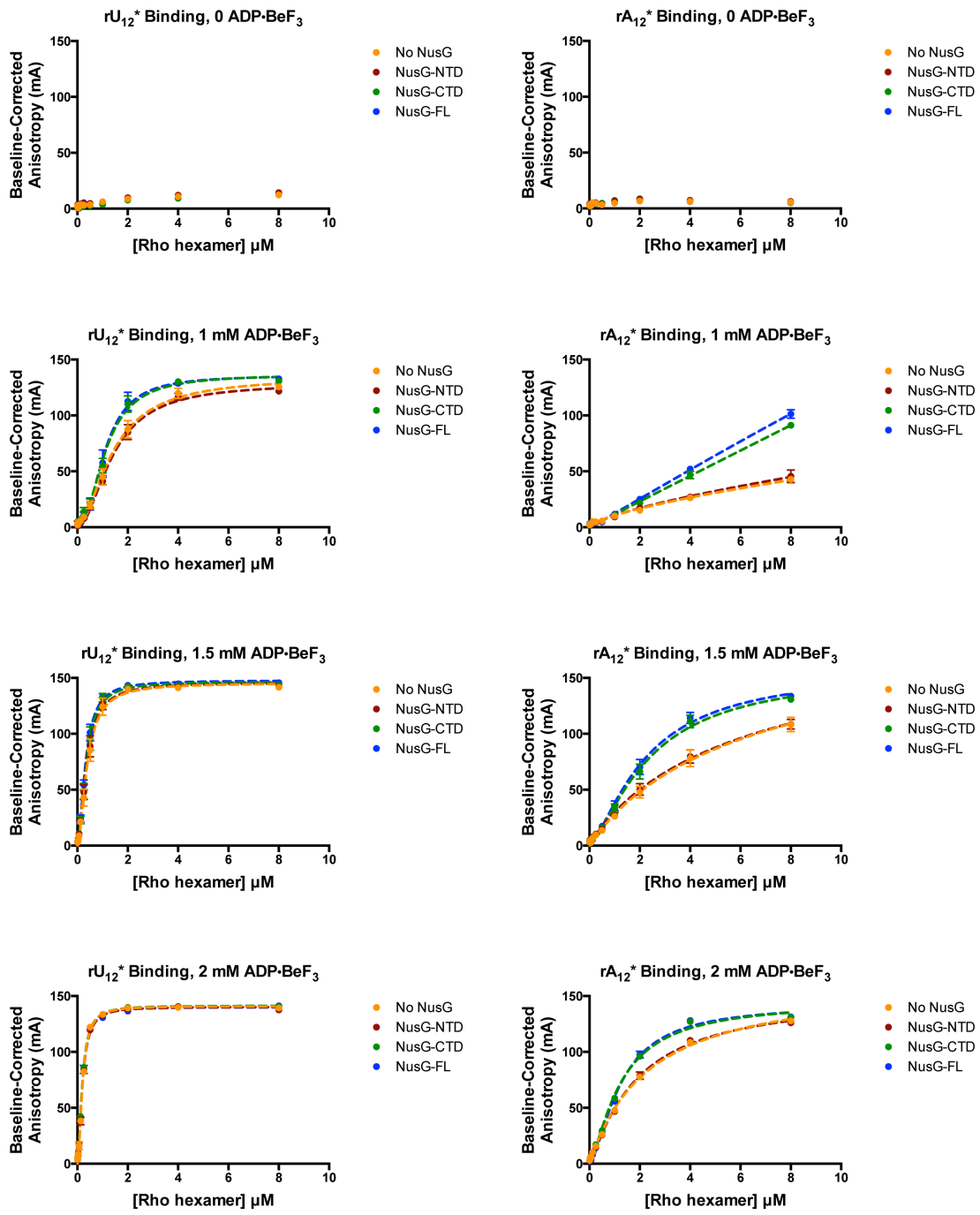

Figure S4.2: Rho's affinity for secondary-site RNAs as determined by fluorescence anisotropy in the presence or absence of different NusG constructs (FL, CTD or NTD). Little to no binding is observed with either rU_{12}* or rA_{12}* in the absence of ADP-BeF_3. The inclusion of the NusG CTD (either as the isolated CTD or with the FL construct) promotes rU_{12}* binding at 1 mM ADP-BeF_3, and rA_{12}* binding at 1, 1.5 and 2 mM ADP-BeF_3. None of the NusG constructs had any visible effect on rU_{12}* binding with 1.5 or 2 mM ADP-BeF_3, and the NusG NTD did not stimulate binding of rU_{12}* or rA_{12}* under any of the conditions that were tested.

Chapter 5: Future Directions and Final Remarks

Future Directions

It is now clear that the conformational control Rho activity is an extensively regulated aspect of bacterial transcription termination. The coupling between Rho ring closure and the binding of both ATP and RNA to Rho's secondary site (Chapter 2) is minimally required to form a catalytically-competent assembly. The promotion of Rho ring closure in response to binding of NusG (Chapter 4), or the association of a pyrimidine-rich *rut* sequence to Rho's primary sites (Chapter 3), serves as mechanism to fine-tune the RNA specificity of Rho-dependent termination. Further study will be required to determine whether the likelihood of Rho acting upon a particular transcript is solely determined by the density of YC motifs within a *rut* and/or by the spacing and arrangement of these elements. The discovery that bicyclomycin inhibits Rho ring closure (Chapter 3) raises the possibility that other known antagonists of Rho function, such as RARE and Hfq, may also act in part through modulation of Rho ring status. Going forward, new assays that allow for the time-resolved monitoring of Rho conformation, in conjunction with studies of the kinetics of other steps in Rho's catalytic cycle, are needed to determine whether the coordinated control of ring closure modulates what might otherwise be a rate-limiting step in Rho-dependent termination.

New in vitro assays to probe Rho ring state

Although the RNA binding assay described in Chapters 3 and 4 has proven to be a powerful tool for rapidly assessing the activity of ligand-dependent regulators of Rho's ring status, this assay also has several limitations. First, since the signal from the experiment depends upon changes in fluorescence anisotropy in response to the capture of an RNA strand within Rho's central pore, it cannot detect the effects of ligands that act on Rho conformation in an RNA-independent manner. Second, the assay relies upon a short oligonucleotide associating with the secondary site to track ring closure; some effectors of Rho ring state (e.g., the RARE element) appear to act only when they are localized in *cis* with other RNA motifs that engage the primary and/or secondary sites. Third, since the RNA-binding assay depends upon changes in fluorescence anisotropy as its readout, it cannot be used for single-molecule microscopy experiments, which would be useful for probing important facets of effector activity, such as the rate of interconversion between Rho ring states. Finally, since the assay is sensitive to the order of addition of ligands (presumably due to the formation of trapped, non-productive states upon the association of open-ring Rho with a non-hydrolyzable nucleotide analog), it is unable to provide insights into RNA binding energies or the rates of ring closure by Rho, and thus is useful primarily as a comparative endpoint assay.

Since bicyclomycin is now known to only bind to open-ring Rho, one alternative approach to interrogate Rho ring state would be to use a derivative of bicyclomycin with fluorescent properties that would enable direct monitoring of drug association. With such a compound in hand, one could allow a given effector to bind to Rho for a period of

time, and then determine the affinity of fluorescent bicyclomycin for the Rho complex of interest; a high affinity of fluorescent bicyclomycin for the sample of interest (relative to open and closed ring controls) would indicate that the Rho ring is predominantly open, whereas a low affinity would indicate a predominantly closed ring. A fluorescent derivative of bicyclomycin could be generated by attaching a fluorescein or phenazine group to the 5a position of bicyclomycin (Chapter 1, **Figure 1.3A**) (Brogan et al., 2003), and binding would then be monitored by changes in fluorescence anisotropy (or by FRET using a suitably labeled Rho construct). However, it may be difficult to obtain sufficient quantities of bicyclomycin for the synthesis of a fluorescent bicyclomycin derivative, especially since the published synthetic schemes for functionalization of the 5a-moiety with phenazine are hampered by multiple low yield steps (Brogan et al., 2003). It is also noteworthy that this assay could yield misleading results if an effector were to alter or interfere with bicyclomycin binding by direct competition. Although a fluorescent bicyclomycin-based ring closure assay could be a powerful way to probe Rho ring state in an RNA-independent manner, it would still bear all of the same limitations of an endpoint assay, and would likely not be of use for single molecule studies of Rho ring state.

The kinetics of Rho ring closure could conceivably be monitored by a time-resolved SAXS assay. In such an experiment, which would need to be conducted at a beamline with rapid mixing equipment (e.g. BL18ID at APS), the rate of ring closure could be measured with the natural substrate, ATP, rather than non-hydrolyzable nucleotide analogs. The use of ATP would allow for direct comparison of the SAXS data with rates obtained from pre-steady-state kinetic assays that have probe steps in Rho's ATPase cycle (Jeong et al., 2004), and could reveal whether ring closure by Rho is rate limiting compared to translocation. The most significant drawback to the SAXS assay is that it may not be useful for probing the kinetics of Rho ring closure in the presence of long RNAs that bear a full *rut* sequence, as the contribution of the unstructured regions of the RNAs may significantly complicate the SAXS profiles.

Lastly, it may be possible to leverage site-specific labeling schemes to introduce fluorescent probes that would be sensitive to conformational changes in the Rho ring. One way to accomplish this would be to identify a pair of residues that are close together in one ring state, and far apart in the other; isomerizations of Rho could then be monitored by tracking changes in FRET efficiency. Since *E. coli* Rho has no surface-exposed native cysteines, it may be possible to introduce the requisite probes by engineering in cysteines that would be labeled with a maleimide-conjugated dye such as AlexaFluor555, and then monitor isomerizations in Rho through changes in homo-FRET. An alternative readout to homo-FRET would be to modify the aforementioned pair to include one cysteine (labeled with maleimide chemistry) and one unnatural amino acid (e.g. one bearing an azide that could be labeled using click chemistry (Rostovtsev et al., 2002)) so that isomerizations of the Rho ring could be tracked through changes in FRET efficiency. This latter setup, although reliant upon what is likely to be a challenging labeling process to achieve the homogeneous introduction of FRET probes throughout a Rho hexamer, would be well suited for single molecule microscopy applications. Although an assay based on introduction of fluorescent labels

on Rho might be arduous to establish, it would be unburdened by the drawbacks intrinsic to the RNA binding-based ring closure assay.

Interrogating Rho's preference for specific arrangements of pyrimidine-rich elements in a rut sequence

From the findings described in Chapter 3, it is now apparent that Rho's primary RNA binding sites perform a sort of molecular arithmetic as a means to productively distinguish between different types of pyrimidine-rich sequences, with the rate of formation of a catalytically competent complex becoming more accelerated as Rho's OB folds become progressively occupied with di-pyrimidine segments. However, it remains unclear whether Rho might also have a preference for specific patterns of pyrimidine-rich sequences (e.g. whether primary site occupancy at subunits A/B/C is equivalent to subunits A/C/E). Knowledge of such a capability would be important for defining the sequence determinants that define whether a particular *rut* sequence is deemed suitable for independent *vs*. NusG –assisted Rho termination activity *in vivo*. One approach to answering these questions would be to use the anisotropy-based ring closure assay with separate oligonucleotides for the primary and secondary sites, and to compare the efficiency of Rho ring closure in response to long DNAs that contain different patterns of pyrimidine and purine elements (e.g. a DNA oligo bearing the sequence $C_{30}A_{45}$ vs. $C_{10}A_{15}C_{10}A_{15}C_{10}A_{15}$). Another approach would be to use any of the aforementioned assays that monitor ring status, and to compare the efficiency of Rho ring closure on long RNAs (likely made through *in vitro* transcription) that contain different patterns of pyrimidine-rich elements.

Determining how RARE inhibits Rho

It is currently unclear how the presence of a RARE motif adjacent to a rut site might inhibit the formation of a catalytically active Rho state (Sevostyanova and Groisman, 2015). One possibility is that RARE binds to Rho in a manner that impedes ring closure, either by acting as a steric block or by working through an allosteric mechanism. Alternatively, the binding of RARE to Rho may induce the formation of a closed-ring Rho state that is not catalytically competent. By using any of the *in vitro* ring closure assays described above, one could determine the extent to which long *rut* and RARE-containing RNAs promote Rho ring closure, and then compare this result to the ring closure efficiency observed with a similar RNA that bears a scrambled RARE sequence. This study would not only be useful for understanding whether the presence of a RARE sequence adjacent to a *rut* impacts Rho ring closure, but could also facilitate the determination of a crystal structure of a Rho:RARE complex by reducing the complexity of the requisite crystallographic screening.

Visualizing the Rho:Hfq interaction

The Rho:Hfq interaction is a particularly intriguing structural target due to the fact that both Hfq and Rho can exist as flat, six-fold symmetric rings (the latter is observed in some closed-ring Rho crystal forms). This structural congruency suggests that Hfq might bind coaxially with Rho's central pore and act as a steric block against the association of RNA or the translocation of an RNA strand through Rho's secondary site. This hypothesis, however, presents a conundrum – how could Rho access a closed-ring

state *in vivo* if a cofactor were sterically blocking an RNA from reaching its secondary (i.e., ATP-dependent) RNA-binding sites? Conversely, it is difficult to envision either how Hfq might engage an open-ring Rho (due to a mismatch between the 6-fold symmetry of Hfq and the quasi-6_5 symmetry of open-ring Rho), or how Hfq could bind closed-ring Rho in the act of translocation (as both the *rut* and the loop of RNA bridging Rho's primary and secondary sites would block association of Hfq with the N-terminus, and the 3'-region of the nascent RNA would prevent the association of Hfq with the C-terminal face). One possible solution to this question is that Hfq might template the formation of a closed-ring Rho in the absence of RNA; this model could be tested using any of the ring closure assays described above, as well as by standard structural approaches. Alternatively, Hfq might sequester Rho away from RNA. Given the complexities of the system, it may be prudent to determine whether the ring status of Rho controls (or is controlled by) Hfq before attempting to solve a co-crystal or cryo-EM structure of this complex.

Final Remarks

Rho has served as an archetype for understanding the mechanochemistry of other RecA and AAA+-type ring translocases. The finding that Rho's conformational state is controlled by a diverse set of intrinsic and extrinsic factors suggests that other ring translocases may be regulated through similar mechanisms. Indeed, ligand- and protein-dependent control of conformational changes has been observed with the bacterial (RecA-type) replicative helicase DnaB (Strycharska et al., 2013), the eukaryotic (AAA+-type) replicative helicase MCMs (Barry et al., 2007; Froelich et al., 2014), and the proteasomal AAA+-family unfoldase Rpt1-6 (Bashore et al., 2015). Future studies of other effectors of Rho activity will be needed to understand the full breadth of mechanisms that control Rho-dependent transcription termination. Similarly, analysis of effectors of other RecA and AAA+-type ring translocases will be needed to determine how broadly the regulatory principles uncovered through studies of Rho match those that govern other molecular motors.

References

Abrahams, J.P., Leslie, A.G., Lutter, R., and Walker, J.E. (1994). Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. Nature *370*, 621-628.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W*., et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica Section D, Biological crystallography *66*, 213-221.

Adelman, J.L., Jeong, Y.J., Liao, J.C., Patel, G., Kim, D.E., Oster, G., and Patel, S.S. (2006). Mechanochemistry of transcription termination factor Rho. Mol Cell *22*, 611-621.

Ahnert, P., Picha, K.M., and Patel, S.S. (2000). A ring-opening mechanism for DNA binding in the central channel of the T7 helicase-primase protein. The EMBO journal *19*, 3418-3427.

Albe, K.R., Butler, M.H., and Wright, B.E. (1990). Cellular concentrations of enzymes and their substrates. J Theor Biol *143*, 163-195.

Amatov, T., Pohl, R., Císařová, I., and Jahn, U. (2015). Synthesis of Bridged Diketopiperazines by Using the Persistent Radical Effect and a Formal Synthesis of Bicyclomycin. Angewandte Chemie International Edition *54*, 12153-12157.

Artsimovitch, I., and Landick, R. (2002). The transcriptional regulator RfaH stimulates RNA chain synthesis after recruitment to elongation complexes by the exposed nontemplate DNA strand. Cell *109*, 193-203.

Bailey, S., Eliason, W.K., and Steitz, T.A. (2007). Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. Science (New York, NY) *318*, 459-463.

Balasubramanian, K., and Stitt, B.L. (2010). Evidence for amino acid roles in the chemistry of ATP hydrolysis in Escherichia coli Rho. Journal of molecular biology *404*, 587-599.

Barry, E.R., McGeoch, A.T., Kelman, Z., and Bell, S.D. (2007). Archaeal MCM has separable processivity, substrate choice and helicase domains. Nucleic acids research *35*, 988-998.

Bashore, C., Dambacher, C.M., Goodall, E.A., Matyskiela, M.E., Lander, G.C., and Martin, A. (2015). Ubp6 deubiquitinase controls conformational dynamics and substrate degradation of the 26S proteasome. Nature structural & molecular biology *22*, 712-719.

Bear, D.G., Andrews, C.L., Singer, J.D., Morgan, W.D., Grant, R.A., von Hippel, P.H., and Platt, T. (1985). Escherichia coli transcription termination factor rho has a two-domain structure in its activated form. Proceedings of the National Academy of Sciences of the United States of America *82*, 1911-1915.

Bear, D.G., Hicks, P.S., Escudero, K.W., Andrews, C.L., McSwiggen, J.A., and von Hippel, P.H. (1988). Interactions of Escherichia coli transcription termination factor rho with RNA. II. Electron microscopy and nuclease protection experiments. Journal of molecular biology *199*, 623-635.

Belogurov, G.A., and Artsimovitch, I. (2015). Regulation of Transcript Elongation. Annual review of microbiology *69*, 49-69.

Belogurov, G.A., Mooney, R.A., Svetlov, V., Landick, R., and Artsimovitch, I. (2009). Functional specialization of transcription elongation factors. The EMBO journal *28*, 112-122.

Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J., and Rabinowitz, J.D. (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. Nat Chem Biol *5*, 593-599.

Bochman, M.L., and Schwacha, A. (2010). The Saccharomyces cerevisiae Mcm6/2 and Mcm5/3 ATPase active sites contribute to the function of the putative Mcm2-7 'gate'. Nucleic Acids Res *38*, 6078-6088.

Bogden, C.E., Fass, D., Bergman, N., Nichols, M.D., and Berger, J.M. (1999). The structural basis for terminator recognition by the Rho transcription termination factor. Molecular cell *3*, 487-493.

Boudvillain, M., Nollmann, M., and Margeat, E. (2010). Keeping up to speed with the transcription termination factor Rho motor. Transcr *1*, 70-75.

Brennan, C.A., Dombroski, A.J., and Platt, T. (1987). Transcription termination factor rho is an RNA-DNA helicase. Cell *48*, 945-952.

Brogan, A.P., Widger, W.R., and Kohn, H. (2003). Bicyclomycin fluorescent probes: synthesis and biochemical, biophysical, and biological properties. The Journal of organic chemistry *68*, 5575-5587.

Browning, D.F., and Busby, S.J. (2004). The regulation of bacterial transcription initiation. Nature reviews Microbiology *2*, 57-65.

Brunger, A.T. (2007). Version 1.2 of the Crystallography and NMR system. Nature protocols *2*, 2728-2733.

Burmann, B.M., Knauer, S.H., Sevostyanova, A., Schweimer, K., Mooney, R.A., Landick, R., Artsimovitch, I., and Rosch, P. (2012). An alpha helix to beta barrel domain switch transforms the transcription factor RfaH into a translation factor. Cell *150*, 291-303.
Burmann, B.M., Schweimer, K., Luo, X., Wahl, M.C., Stitt, B.L., Gottesman, M.E., and Rosch, P. (2010). A NusE:NusG complex links transcription and translation. Science (New York, NY) *328*, 501-504.
Burns, C.M., and Richardson, J.P. (1995). NusG is required to overcome a kinetic limitation to Rho function at an intragenic terminator. Proceedings of the National Academy of Sciences of the United States of America *92*, 4738-4742.
Capp, J.A., Hagarman, A., Richardson, D.C., and Oas, T.G. (2014). The statistical conformation of a highly flexible protein: small-angle X-ray scattering of S. aureus protein A. Structure (London, England : 1993) *22*, 1184-1195.
Cardinale, C.J., Washburn, R.S., Tadigotla, V.R., Brown, L.M., Gottesman, M.E., and Nudler, E. (2008). Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E. coli. Science (New York, NY) *320*, 935-938.
Chao, Y., and Vogel, J. (2010). The role of Hfq in bacterial pathogens. Current opinion in microbiology *13*, 24-33.
Chen, C.Y., and Richardson, J.P. (1987). Sequence elements essential for rho-dependent transcription termination at lambda tR1. J Biol Chem *262*, 11292-11299.
Chen, Z., Yang, H., and Pavletich, N.P. (2008). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature *453*, 489-484.
Chistol, G., Liu, S., Hetherington, C.L., Moffitt, J.R., Grimes, S., Jardine, P.J., and Bustamante, C. (2012). High degree of coordination and division of labor among subunits in a homomeric ring ATPase. Cell *151*, 1017-1028.
Cohn, M., Cohen, G.N., and Monod, J. (1953). [The specific inhibiting effect of methionine in the formation of methionine-synthase by Escherichia coli]. Comptes rendus hebdomadaires des seances de l'Academie des sciences *236*, 746-748.
Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M.R., and Berger, J.M. (2011). The structural basis for MCM2-7 helicase activation by GINS and Cdc45. Nature structural & molecular biology *18*, 471-477.
Costa, A., Renault, L., Swuec, P., Petojevic, T., Pesavento, J.J., Ilves, I., MacLellan-Gibson, K., Fleck, R.A., Botchan, M.R., and Berger, J.M. (2014). DNA binding polarity, dimerization, and ATPase ring remodeling in the CMG helicase of the eukaryotic replisome. eLife *3*, e03273.
Cowtan, K.D., and Main, P. (1993). Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. Acta crystallographica Section D, Biological crystallography *49*, 148-157.
d'Aubenton Carafa, Y., Brody, E., and Thermes, C. (1990). Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. J Mol Biol *216*, 835-858.
Davey, M.J., Jeruzalmi, D., Kuriyan, J., and O'Donnell, M. (2002). Motors and switches: AAA+ machines within the replisome. Nature reviews Molecular cell biology *3*, 826-835.
Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S.*, et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic acids research *35*, W375-383.
Delagoutte, E., and von Hippel, P.H. (2003). Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: Integration of helicases into cellular processes. Q Rev Biophys *36*, 1-69.
Deshaies, R.J. (2014). Proteotoxic crisis, the ubiquitin-proteasome system, and cancer therapy. BMC Biol *12*.
DiRita, V.J., and Mekalanos, J.J. (1989). Genetic regulation of bacterial virulence. Annual review of genetics *23*, 455-482.
Dolan, J.W., Marshall, N.F., and Richardson, J.P. (1990). Transcription termination factor rho has three distinct structural domains. The Journal of biological chemistry *265*, 5747-5754.
Dombroski, A.J., and Platt, T. (1988). Structure of rho factor: an RNA-binding domain and a separate region with strong similarity to proven ATP-binding domains. Proceedings of the National Academy of Sciences of the United States of America *85*, 2538-2542.
Durrant, J.D., de Oliveira, C.A., and McCammon, J.A. (2011). POVME: an algorithm for measuring binding-pocket volumes. J Mol Graph Model *29*, 773-776.

Dyer, K.N., Hammel, M., Rambo, R.P., Tsutakawa, S.E., Rodic, I., Classen, S., Tainer, J.A., and Hura, G.L. (2014). High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. Methods in molecular biology (Clifton, NJ) *1091*, 245-258.
Enemark, E.J., and Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. Nature *442*, 270-275.
Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., and Nudler, E. (2007). An allosteric path to transcription termination. Molecular cell *28*, 991-1001.
Epshtein, V., Dutta, D., Wade, J., and Nudler, E. (2010). An allosteric mechanism of Rho-dependent transcription termination. Nature *463*, 245-249.
Ericsson, C.D., DuPont, H.L., Galindo, E., Mathewson, J.J., Morgan, D.R., Wood, L.V., and Mendiola, J. (1985). Efficacy of bicozamycin in preventing traveler's diarrhea. Gastroenterology *88*, 473-477.
Etlinger, J.D., and Goldberg, A.L. (1977). A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. Proceedings of the National Academy of Sciences of the United States of America *74*, 54-58.
Faus, I., and Richardson, J.P. (1990). Structural and functional properties of the segments of lambda cro mRNA that interact with transcription termination factor Rho. J Mol Biol *212*, 53-66.
Flores, S., Echols, N., Milburn, D., Hespenheide, B., Keating, K., Lu, J., Wells, S., Yu, E.Z., Thorpe, M., and Gerstein, M. (2006). The Database of Macromolecular Motions: new features added at the decade mark. Nucleic Acids Res *34*, D296-301.
Frauenfelder, H., Petsko, G.A., and Tsernoglou, D. (1979). Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. Nature *280*, 558-563.
Froelich, C.A., Kang, S., Epling, L.B., Bell, S.P., and Enemark, E.J. (2014). A conserved MCM single-stranded DNA binding element is essential for replication initiation. eLife *3*, e01993.
Galluppi, G.R., and Richardson, J.P. (1980). ATP-induced changes in the binding of RNA synthesis termination protein Rho to RNA. J Mol Biol *138*, 513-539.
Geiselmann, J., and von Hippel, P.H. (1992). Functional interactions of ligand cofactors with Escherichia coli transcription termination factor rho. I. Binding of ATP. Protein Sci *1*, 850-860.
Gocheva, V., Le Gall, A., Boudvillain, M., Margeat, E., and Nollmann, M. (2015). Direct observation of the translocation mechanism of transcription termination factor Rho. Nucleic acids research *43*, 2367-2377.
Gogol, E.P., Seifried, S.E., and von Hippel, P.H. (1991). Structure and assembly of the Escherichia coli transcription termination factor rho and its interaction with RNA. I. Cryoelectron microscopic studies. Journal of molecular biology *221*, 1127-1138.
Gusarov, I., and Nudler, E. (1999). The mechanism of intrinsic transcription termination. Molecular cell *3*, 495-504.
Gutierrez, P., Kozlov, G., Gabrielli, L., Elias, D., Osborne, M.J., Gallouzi, I.E., and Gehring, K. (2007). Solution structure of YaeO, a Rho-specific inhibitor of transcription termination. The Journal of biological chemistry *282*, 23348-23353.
Hart, C.M., and Roberts, J.W. (1991). Rho-dependent transcription termination. Characterization of the requirement for cytidine in the nascent transcript. The Journal of biological chemistry *266*, 24140-24148.
Hirtreiter, A., Damsma, G.E., Cheung, A.C., Klose, D., Grohmann, D., Vojnic, E., Martin, A.C., Cramer, P., and Werner, F. (2010). Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. Nucleic acids research *38*, 4040-4051.
Hollands, K., Proshkin, S., Sklyarova, S., Epshtein, V., Mironov, A., Nudler, E., and Groisman, E.A. (2012). Riboswitch control of Rho-dependent transcription termination. Proceedings of the National Academy of Sciences of the United States of America *109*, 5376-5381.
Iseki, M., Miyoshi, T., Konomi, T., and Imanaka, H. (1980). Biosynthesis of bicyclomycin. II. Biosynthetic conditions and incorporation of radioactive precursors into bicyclomycin by washed mycelium. The Journal of antibiotics *33*, 488-493.
Itsathitphaisarn, O., Wing, R.A., Eliason, W.K., Wang, J., and Steitz, T.A. (2012). The Hexameric Helicase DnaB Adopts a Nonplanar Conformation during Translocation. Cell.
Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol *3*, 318-356.
Jeong, Y.J., Kim, D.E., and Patel, S.S. (2004). Nucleotide binding induces conformational changes in Escherichia coli transcription termination factor Rho. The Journal of biological chemistry *279*, 18370-18376.

Johnson, D.S., Bai, L., Smith, B.Y., Patel, S.S., and Wang, M.D. (2007). Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped T7 helicase. Cell *129*, 1299-1309.
Kabsch, W. (2010). XDS. Acta crystallographica Section D, Biological crystallography *66*, 125-132.
Kagawa, R., Montgomery, M.G., Braig, K., Leslie, A.G., and Walker, J.E. (2004). The structure of bovine F1-ATPase inhibited by ADP and beryllium fluoride. EMBO J *23*, 2734-2744.
Kalyani, B.S., Muteeb, G., Qayyum, M.Z., and Sen, R. (2011). Interaction with the nascent RNA is a prerequisite for the recruitment of Rho to the transcription elongation complex in vitro. J Mol Biol *413*, 548-560.
Kennell, D., and Riezman, H. (1977). Transcription and translation initiation frequencies of the Escherichia coli lac operon. J Mol Biol *114*, 1-21.
Kim, D.E., and Patel, S.S. (1999). The mechanism of ATP hydrolysis at the noncatalytic sites of the transcription termination factor Rho. J Biol Chem *274*, 32667-32671.
Kim, D.E., Shigesada, K., and Patel, S.S. (1999). Transcription termination factor Rho contains three noncatalytic nucleotide binding sites. J Biol Chem *274*, 11623-11628.
Kleywegt, G.J., and Jones, T.A. (1997). Template convolution to enhance or detect structural features in macromolecular electron-density maps. Acta crystallographica Section D, Biological crystallography *53*, 179-185.
Knowlton, J.R., Bubunenko, M., Andrykovitch, M., Guo, W., Routzahn, K.M., Waugh, D.S., Court, D.L., and Ji, X. (2003). A spring-loaded state of NusG in its functional cycle is suggested by X-ray crystallography and supported by site-directed mutants. Biochemistry *42*, 2275-2281.
Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003a). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J Appl Crystallogr *36*, 1277-1282.
Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003b). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography *36*, 1277-1282.
Koslover, D.J., Fazal, F.M., Mooney, R.A., Landick, R., and Block, S.M. (2012a). Binding and translocation of termination factor rho studied at the single-molecule level. J Mol Biol *423*, 664-676.
Koslover, D.J., Fazal, F.M., Mooney, R.A., Landick, R., and Block, S.M. (2012b). Binding and translocation of termination factor rho studied at the single-molecule level. Journal of molecular biology *423*, 664-676.
Kull, F.J., Sablin, E.P., Lau, R., Fletterick, R.J., and Vale, R.D. (1996). Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. Nature *380*, 550-555.
Lau, L.F., Roberts, J.W., and Wu, R. (1983). RNA polymerase pausing and transcript release at the lambda tR1 terminator in vitro. J Biol Chem *258*, 9391-9397.
Leela, J.K., Syeda, A.H., Anupama, K., and Gowrishankar, J. (2013). Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America *110*, 258-263.
Liao, J.-C., Jeong, Y.-J., Kim, D.-E., Patel, S.S., and Oster, G. (2005a). Mechanochemistry of T7 DNA Helicase. Journal of Molecular Biology *350*, 452-475.
Liao, J.C., Jeong, Y.J., Kim, D.E., Patel, S.S., and Oster, G. (2005b). Mechanochemistry of t7 DNA helicase. J Mol Biol *350*, 452-475.
Linderoth, N.A., and Calendar, R.L. (1991). The Psu protein of bacteriophage P4 is an antitermination factor for rho-dependent transcription termination. Journal of bacteriology *173*, 6722-6731.
Lowery-Goldhammer, C., and Richardson, J.P. (1974). An RNA-dependent nucleoside triphosphate phosphohydrolase (ATPase) associated with rho termination factor. Proc Natl Acad Sci U S A *71*, 2003-2007.
Lowry, O.H., Carter, J., Ward, J.B., and Glaser, L. (1971). The effect of carbon and nitrogen sources on the level of metabolic intermediates in Escherichia coli. J Biol Chem *246*, 6511-6521.
Lu, X.J., and Olson, W.K. (2003). 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. Nucleic Acids Res *31*, 5108-5121.
Lyubimov, A.Y., Costa, A., Bleichert, F., Botchan, M.R., and Berger, J.M. (2012). ATP-dependent conformational dynamics underlie the functional asymmetry of the replicative helicase from a minimalist eukaryote. Proceedings of the National Academy of Sciences of the United States of America *109*, 11999-12004.
Lyubimov, A.Y., Strycharska, M., and Berger, J.M. (2011). The nuts and bolts of ring-translocase structure and mechanism. Current opinion in structural biology *21*, 240-248.

Magyar, A., Zhang, X., Kohn, H., and Widger, W.R. (1996). The antibiotic bicyclomycin affects the secondary RNA binding site of Escherichia coli transcription termination factor Rho. The Journal of biological chemistry *271*, 25369-25374.
Marles-Wright, J., and Lewis, R.J. (2007). Stress responses of bacteria. Current opinion in structural biology *17*, 755-760.
Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C., and Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. The EMBO journal *30*, 1302-1310.
Matyskiela, M.E., and Martin, A. (2013). Design principles of a universal protein degradation machine. J Mol Biol *425*, 199-213.
McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. Journal of Applied Crystallography *40*, 658-674.
McSwiggen, J.A., Bear, D.G., and von Hippel, P.H. (1988). Interactions of Escherichia coli transcription termination factor rho with RNA. I. Binding stoichiometries and free energies. J Mol Biol *199*, 609-622.
Mironov, A.S., Gusarov, I., Rafikov, R., Lopez, L.E., Shatalin, K., Kreneva, R.A., Perumov, D.A., and Nudler, E. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. Cell *111*, 747-756.
Miwa, Y., Horiguchi, T., and Shigesada, K. (1995). Structural and functional dissections of transcription termination factor rho by random mutagenesis. J Mol Biol *254*, 815-837.
Miyamura, S., Ogasawara, N., Otsuka, H., Niwayama, S., and Tanaka, H. (1972). Antibiotic no. 5879, a new water-soluble antibiotic against gram-negative bacteria. The Journal of antibiotics *25*, 610-612.
Miyoshi, T., Miyairi, N., Aoki, H., Kosaka, M., and Sakai, H. (1972). Bicyclomycin, a new antibiotic. I. Taxonomy, isolation and characterization. The Journal of antibiotics *25*, 569-575.
Mooney, R.A., Davis, S.E., Peters, J.M., Rowland, J.L., Ansari, A.Z., and Landick, R. (2009a). Regulator trafficking on bacterial transcription units in vivo. Molecular cell *33*, 97-108.
Mooney, R.A., Schweimer, K., Rosch, P., Gottesman, M., and Landick, R. (2009b). Two structurally independent domains of E. coli NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. J Mol Biol *391*, 341-358.
Morgan, W.D., Bear, D.G., Litchman, B.L., and von Hippel, P.H. (1985). RNA sequence and secondary structure requirements for rho-dependent transcription termination. Nucleic acids research *13*, 3739-3754.
Morgan, W.D., Bear, D.G., and von Hippel, P.H. (1983). Rho-dependent termination of transcription. II. Kinetics of mRNA elongation during transcription from the bacteriophage lambda PR promoter. J Biol Chem *258*, 9565-9574.
Nehrke, K.W., Zalatan, F., and Platt, T. (1993). NusG alters rho-dependent termination of transcription in vitro independent of kinetic coupling. Gene expression *3*, 119-133.
Noji, H., Yasuda, R., Yoshida, M., and Kinosita, K., Jr. (1997). Direct observation of the rotation of F1-ATPase. Nature *386*, 299-302.
O'Shea, V.L., and Berger, J.M. (2014). Loading strategies of ring-shaped nucleic acid translocases and helicases. Current opinion in structural biology *25*, 16-24.
Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. In Methods in Enzymology, C.W. Carter, Jr., and R.M. Sweet, eds. (New York: Academic Press), pp. 307-326.
Painter, J., and Merritt, E.A. (2006). TLSMD web server for the gernation of multi-group TLS models. J Appl Cryst *39*, 109-111.
Pani, B., Banerjee, S., Chalissery, J., Muralimohan, A., Loganathan, R.M., Suganthan, R.B., and Sen, R. (2006). Mechanism of inhibition of Rho-dependent transcription termination by bacteriophage P4 protein Psu. The Journal of biological chemistry *281*, 26491-26500.
Park, H.G., Zhang, X., Moon, H.S., Zwiefka, A., Cox, K., Gaskell, S.J., Widger, W.R., and Kohn, H. (1995). Bicyclomycin and dihydrobicyclomycin inhibition kinetics of Escherichia coli rho-dependent transcription termination factor ATPase activity. Archives of biochemistry and biophysics *323*, 447-454.
Park, J.S., and Roberts, J.W. (2006). Role of DNA bubble rewinding in enzymatic transcription termination. Proceedings of the National Academy of Sciences of the United States of America *103*, 4870-4875.
Pasman, Z., and von Hippel, P.H. (2000). Regulation of rho-dependent transcription termination by NusG is specific to the Escherichia coli elongation complex. Biochemistry *39*, 5573-5585.

Patel, S.S. (2009). Structural biology: Steps in the right direction. Nature *462*, 581-583.
Patel, S.S., and Picha, K.M. (2000). Structure and function of hexameric helicases. Annu Rev Biochem *69*, 651-697.
Peters, J.M., Mooney, R.A., Grass, J.A., Jessen, E.D., Tran, F., and Landick, R. (2012). Rho and NusG suppress pervasive antisense transcription in Escherichia coli. Genes & development *26*, 2621-2633.
Peters, J.M., Mooney, R.A., Kuan, P.F., Rowland, J.L., Keles, S., and Landick, R. (2009). Rho directs widespread termination of intragenic and stable RNA transcription. Proceedings of the National Academy of Sciences of the United States of America *106*, 15406-15411.
Petoukhov, M.V., Konarev, P.V., Kikhney, A.G., and Svergun, D.I. (2007). ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. J Appl Cryst *40*, s223-s228.
Putnam, C.D., Hammel, M., Hura, G.L., and Tainer, J.A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Quarterly reviews of biophysics *40*, 191-285.
Rabhi, M., Espeli, O., Schwartz, A., Cayrol, B., Rahmouni, A.R., Arluison, V., and Boudvillain, M. (2011a). The Sm-like RNA chaperone Hfq mediates transcription antitermination at Rho-dependent terminators. The EMBO journal *30*, 2805-2816.
Rabhi, M., Gocheva, V., Jacquinot, F., Lee, A., Margeat, E., and Boudvillain, M. (2011b). Mutagenesis-based evidence for an asymmetric configuration of the ring-shaped transcription termination factor Rho. J Mol Biol *405*, 497-518.
Rambo, R.P. (2015). Resolving Individual Components in Protein-RNA Complexes Using Small-Angle X-ray Scattering Experiments. Methods in enzymology *558*, 363-390.
Rambo, R.P., and Tainer, J.A. (2013). Accurate assessment of mass, models and resolution by small-angle scattering. Nature *496*, 477-481.
Ray-Soni, A., Bellecourt, M.J., and Landick, R. (2016). Mechanisms of Bacterial Transcription Termination: All Good Things Must End. Annual review of biochemistry *85*, 319-347.
Rees, D.M., Montgomery, M.G., Leslie, A.G., and Walker, J.E. (2012). Structural evidence of a new catalytic intermediate in the pathway of ATP hydrolysis by F1-ATPase from bovine heart mitochondria. Proceedings of the National Academy of Sciences of the United States of America *109*, 11139-11143.
Richardson, J.P. (1982). Activation of rho protein ATPase requires simultaneous interaction at two kinds of nucleic acid-binding sites. J Biol Chem *257*, 5760-5766.
Richardson, J.P. (2002). Rho-dependent termination and ATPases in transcript termination. Biochim Biophys Acta *1577*, 251-260.
Richardson, J.P. (2003). Loading Rho to terminate transcription. Cell *114*, 157-159.
Roberts, J.W. (1969). Termination factor for RNA synthesis. Nature *224*, 1168-1174.
Rostovtsev, V.V., Green, L.G., Fokin, V.V., and Sharpless, K.B. (2002). A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective “Ligation” of Azides and Terminal Alkynes. Angewandte Chemie International Edition *41*, 2596-2599.
Rothenberg, E., Trakselis, M.A., Bell, S.D., and Ha, T. (2007). MCM forked substrate specificity involves dynamic interaction with the 5'-tail. The Journal of biological chemistry *282*, 34229-34234.
Ruteshouser, E.C., and Richardson, J.P. (1989). Identification and characterization of transcription termination sites in the Escherichia coli lacZ gene. J Mol Biol *208*, 23-43.
Sablin, E.P., Case, R.B., Dai, S.C., Hart, C.L., Ruby, A., Vale, R.D., and Fletterick, R.J. (1998). Direction determination in the minus-end-directed kinesin motor ncd. Nature *395*, 813-816.
Sanders, C.M., Kovalevskiy, O.V., Sizov, D., Lebedev, A.A., Isupov, M.N., and Antson, A.A. (2007). Papillomavirus E1 helicase assembly maintains an asymmetric state in the absence of DNA and nucleotide cofactors. Nucleic acids research *35*, 6451-6457.
Santangelo, T.J., and Roberts, J.W. (2004). Forward translocation is the natural pathway of RNA release at an intrinsic terminator. Molecular cell *14*, 117-126.
Sauer, R.T., and Baker, T.A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. Annual review of biochemistry *80*, 587-612.
Schmidt, H., Zalyte, R., Urnavicius, L., and Carter, A.P. (2014). Structure of human cytoplasmic dynein-2 primed for its power stroke. Nature.
Schneidman-Duhovny, D., Hammel, M., and Sali, A. (2010a). FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic acids research *38*, W540-544.
Schneidman-Duhovny, D., Hammel, M., and Sali, A. (2010b). FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic acids research *38*, W540-544.

Schneidman-Duhovny, D., Hammel, M., Tainer, J.A., and Sali, A. (2013). Accurate SAXS profile computation and its assessment by contrast variation experiments. Biophysical journal *105*, 962-974.
Schulz, S., Gietl, A., Smollett, K., Tinnefeld, P., Werner, F., and Grohmann, D. (2016). TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle. Proceedings of the National Academy of Sciences of the United States of America *113*, E1816-1825.
Schwartz, A., Margeat, E., Rahmouni, A.R., and Boudvillain, M. (2007a). Transcription termination factor rho can displace streptavidin from biotinylated RNA. The Journal of biological chemistry *282*, 31469-31476.
Schwartz, A., Rabhi, M., Jacquinot, F., Margeat, E., Rahmouni, A.R., and Boudvillain, M. (2009). A stepwise 2'-hydroxyl activation mechanism for the bacterial transcription termination factor Rho helicase. Nat Struct Mol Biol *16*, 1309-1316.
Schwartz, A., Walmacq, C., Rahmouni, A.R., and Boudvillain, M. (2007b). Noncanonical interactions in the management of RNA structural blocks by the transcription termination rho helicase. Biochemistry *46*, 9366-9379.
Sedlyarova, N., Shamovsky, I., Bharati, B.K., Epshtein, V., Chen, J., Gottesman, S., Schroeder, R., and Nudler, E. (2016). sRNA-Mediated Control of Transcription Termination in E. coli. Cell *167*, 111-121.e113.
Seifried, S.E., Easton, J.B., and von Hippel, P.H. (1992). ATPase activity of transcription-termination factor rho: functional dimer model. Proc Natl Acad Sci U S A *89*, 10454-10458.
Sevostyanova, A., Belogurov, G.A., Mooney, R.A., Landick, R., and Artsimovitch, I. (2011). The beta subunit gate loop is required for RNA polymerase modification by RfaH and NusG. Molecular cell *43*, 253-262.
Sevostyanova, A., and Groisman, E.A. (2015). An RNA motif advances transcription by preventing Rho-dependent termination. Proceedings of the National Academy of Sciences of the United States of America *112*, E6835-6843.
Shindyalov, I.N., and Bourne, P.E. (1998). Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng *11*, 739-747.
Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007). Structure and mechanism of helicases and nucleic acid translocases. Annual review of biochemistry *76*, 23-50.
Singleton, M.R., Sawaya, M.R., Ellenberger, T., and Wigley, D.B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. Cell *101*, 589-600.
Skordalakes, E., and Berger, J.M. (2003). Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. Cell *114*, 135-146.
Skordalakes, E., and Berger, J.M. (2006). Structural insights into RNA-dependent ring closure and ATPase activation by the Rho termination factor. Cell *127*, 553-564.
Skordalakes, E., Brogan, A.P., Park, B.S., Kohn, H., and Berger, J.M. (2005). Structural mechanism of inhibition of the Rho transcription termination factor by the antibiotic bicyclomycin. Structure (London, England : 1993) *13*, 99-109.
Sledjeski, D.D., Whitman, C., and Zhang, A. (2001). Hfq is necessary for regulation by the untranslated RNA DsrA. Journal of bacteriology *183*, 1997-2005.
Soares, E., Schwartz, A., Nollmann, M., Margeat, E., and Boudvillain, M. (2014). The RNA-mediated, asymmetric ring regulatory mechanism of the transcription termination Rho helicase decrypted by time-resolved nucleotide analog interference probing (trNAIP). Nucleic acids research *42*, 9270-9284.
Steinmetz, E.J., and Platt, T. (1994). Evidence supporting a tethered tracking model for helicase activity of Escherichia coli Rho factor. Proceedings of the National Academy of Sciences of the United States of America *91*, 1401-1405.
Stitt, B.L. (1988). Escherichia coli transcription termination protein rho has three hydrolytic sites for ATP. J Biol Chem *263*, 11130-11137.
Story, R.M., Weber, I.T., and Steitz, T.A. (1992). The structure of the E. coli recA protein monomer and polymer. Nature *355*, 318-325.
Strycharska, M.S., Arias-Palomo, E., Lyubimov, A.Y., Erzberger, J.P., O'Shea, V.L., Bustamante, C.J., and Berger, J.M. (2013). Nucleotide and partner-protein control of bacterial replicative helicase structure and function. Molecular cell *52*, 844-854.
Sullivan, S.L., and Gottesman, M.E. (1992). Requirement for E. coli NusG protein in factor-dependent transcription termination. Cell *68*, 989-994.

Svergun, D., Barberato, C., and Koch, M.H.J. (1995). CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. Journal of Applied Crystallography *28*, 768-773.
Svergun, D.I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr *25*, 495-503.
Terwilliger, T.C. (2000). Maximum-likelihood density modification. Acta Crystallogr D Biol Crystallogr *56*, 965-972.
Thomsen, N.D., and Berger, J.M. (2009). Running in reverse: the structural basis for translocation polarity in hexameric helicases. Cell *139*, 523-534.
Thomsen, N.D., and Berger, J.M. (2012). Crystallization and X-ray structure determination of an RNA-dependent hexameric helicase. Methods in enzymology *511*, 171-190.
Tomar, S.K., and Artsimovitch, I. (2013). NusG-Spt5 proteins-Universal tools for transcription modification and communication. Chemical reviews *113*, 8604-8619.
Valabhoju, V., Agrawal, S., and Sen, R. (2016). Molecular Basis of NusG-mediated Regulation of Rho-dependent Transcription Termination in Bacteria. The Journal of biological chemistry *291*, 22386-22403.
Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. The EMBO journal *1*, 945-951.
Wang, Y., and von Hippel, P.H. (1993). Escherichia coli transcription termination factor rho. I. ATPase activation by oligonucleotide cofactors. The Journal of biological chemistry *268*, 13940-13946.
Wei, R.R., and Richardson, J.P. (2001). Identification of an RNA-binding Site in the ATP binding domain of Escherichia coli Rho by H2O2/Fe-EDTA cleavage protection studies. J Biol Chem *276*, 28380-28387.
Weixlbaumer, A., Leon, K., Landick, R., and Darst, S.A. (2013). Structural basis of transcriptional pausing in bacteria. Cell *152*, 431-441.
Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta crystallographica Section D, Biological crystallography *67*, 235-242.
Wittinghofer, A., Scheffzek, K., and Ahmadian, M.R. (1997). The interaction of Ras with GTPase-activating proteins. FEBS letters *410*, 63-67.
Xu, Y., Johnson, J., Kohn, H., and Widger, W.R. (2003). ATP binding to Rho transcription termination factor. Mutant F355W ATP-induced fluorescence quenching reveals dynamic ATP binding. The Journal of biological chemistry *278*, 13719-13727.
Xu, Y., Kohn, H., and Widger, W.R. (2002). Mutations in the rho transcription termination factor that affect RNA tracking. The Journal of biological chemistry *277*, 30023-30030.
Yarnell, W.S., and Roberts, J.W. (1999). Mechanism of intrinsic transcription termination and antitermination. Science (New York, NY) *284*, 611-615.
Yasuda, R., Noji, H., Yoshida, M., Kinosita, K., Jr., and Itoh, H. (2001). Resolution of distinct rotational substeps by submillisecond kinetic analysis of F1-ATPase. Nature *410*, 898-904.
Yu, X., Horiguchi, T., Shigesada, K., and Egelman, E.H. (2000). Three-dimensional reconstruction of transcription termination factor rho: orientation of the N-terminal domain and visualization of an RNA-binding site. J Mol Biol *299*, 1279-1287.
Zhang, W., Dunkle, J.A., and Cate, J.H. (2009). Structures of the ribosome in intermediate states of ratcheting. Science *325*, 1014-1017.
Zuo, Y., Wang, Y., and Steitz, T.A. (2013). The mechanism of E. coli RNA polymerase regulation by ppGpp is suggested by the structure of their complex. Molecular cell *50*, 430-436.

Appendix

NusG Expression and Purification

Materials

- Wild type *E. coli* NusG (either N-terminus spanning residues 1-119, C-terminus 125-181, or full-length 1-181) cloned into MacroLab 1C LIC vector (N-terminal 6xHis-MBP-N10-TEV tag)
- Chemically competent BL21 (DE3) Codon+ RIL *E. coli* cells
- 1 LB-Kanamycin agar plate
- 6 L 2xYT media in baffled flasks
- 75 mL 2xYT in small (250 mL) baffled flask
- 6.1 mL of 1000x Kanamycin (50 mg/mL in milliQ-purified water)
- 0.72 g of IPTG
- 2 mL of 1000x Pepstatin A (2 mg/mL in 90% methanol / 10% acetic acid)
- 2 mL of 1000x Leupeptin (1 mg/mL in milliQ-purified water)
- 1 mL of 1M DTT
- 20 mL of 100x PMSF (100 mM, 1 mM final)
- 20x 15 mL Falcon tubes
- 8 vials of TEV protease (UCB MacroLab)
- Qsonica sonicator
- BioRad FPLC with attached 130 mL Sephacryl S-200 HR size exclusion column
- 2 5 mL HisTrap HP columns
- Rainin Rabbit Plus peristaltic pump (“bunny pump”) and tubing
- Sorvall RC-5B centrifuge
- SS34 centrifuge rotor and tubes
- Sorvall RC-3B plus centrifuge, rotor, and tubes
- Thermo ST40R tabletop centrifuge and rotor
- MilliQ-purified water
- 37° tabletop incubator
- Innova 44 shaking incubator
- 15 mL Slide-a-lyzer 3.5k MWCO dialysis cassette
- 2 10 mL syringes
- 3 mL syringe
- 2 22-gague needles
- 200 mL metal beaker
- Stir plate, stir bar and 1L beaker
- 2 Amicon Ultra 15 30k MWCO concentrator (Millipore)
- 2 Amicon Ultra 15 3k MWCO concentrators (Millipore)
- BioRad protein assay dye reagent (“Bradford reagent”)
- 2 L of Ni-A buffer:
  500 mM KCl, 50 mM Tris, 30 mM Imidazole,
  (1x PMSF/Leupeptin/Pepstatin A), 10% Glycerol, final pH to 7.5
- 500 mL of Ni-B buffer:
  500 mM KCl, 50 mM Tris, 300 mM Imidazole, 10% Glycerol, final pH to 7.5
- 500 mL of SEC buffer:
  500 mM KCl, 50 mM Tris, 10% Glycerol,1 mM DTT, final pH to 7.5

Methods

Day 1

- Transform chemically competent BL21 (DE3) Codon+ RIL cells using standard protocols, plate on LB-Kan plate, allow to grow overnight at 37° in tabletop incubator

Day 2

- Make, autoclave, and pre-warm 6 L of 2xYT media overnight at 37°
- Add 75 uL of 1000x Kanamycin to overnight flask, and inoculate with numerous (>50) colonies from LB-Kan plate. Allow to grow overnight in shaker at 37°

Day 3

- Add 1 mL of 1000x Kanamycin to each 2xYT
- Inoculate each of the 6 L of 2xYT with 10 mL from the overnight culture
- Grow 6 L cultures at 37° to an OD of 0.7-0.8
- Add 0.12 g of IPTG to each flask, allow to express overnight shaking at 18°

Day 4

- Pour 6 L of cultures into RC-5B centrifuge bottles
- Spin down cells in RC-5B pre-chilled to 4°, spin at 4000 x g for 20 minutes
- Wash both HisTrap HP columns (connected in series onto bunny pump) with 100 mL MilliQ-purified water, then 100 mL of Ni-A
- Resuspend cells into 50 mL of Ni-A, pour slurry into metal beaker, place beaker in middle of a tightly-packed ice bucket
- Lyse for 3 minutes total (15 seconds on / 59 seconds off) using Qsonica sonicator set to 75% power
- Pour lysate into SS34 centrifuge tubes, and spin at 17000 rpm for 30 minutes in RC-3B (pre-chilled to 4°)
- Load lysate onto equilibrated HisTrap HP columns (speed setting of “10.0”) with bunny pump
- Wash HisTrap columns with 200 mL of Ni-A buffer (switch to speed setting of “12.0” after 100 mL) then begin checking protein concentration lysate by adding 2 uL of flow-through to 48 uL of Bradford reagent. Wash until addition of flow-through to Bradford reagent does not yield a color change.
- Elute HisTrap columns with 30 mL of Ni-B buffer, concentrate to 10 mL volume in 2 Amicon Ultra 15 30k MWCO concentrators (5 mL in each) in tabletop centrifuge at 4,000x g – this concentrated sample is referred to as “Ni elute” herein
- Wash HisTrap columns with 100 mL of Ni-A buffer, cap top/bottom and store at 4° for use on Day 5
- Quantify concentration of Ni elute using Nanodrop (absorbance of 1.03, 0.234 and 0.753 at 1 cm path length ~= 1 mg/mL for NTD, CTD and FL constructs respectively)
- Check purity of Ni elute by SDS-PAGE (for WT constructs, there is invariably one band and this sample can be checked at the end of the prep instead)
- Add one tube of MacroLab TEV per 25 mg of protein, generally ~200 mg of protein / 8 tubes of TEV
- Carefully place Ni elute / TEV into dialysis cassette using syringe/needle
- Place dialysis cassette into a beaker containing a stir bar and 1L of Ni-A, place at 4° and allow to TEV cleave overnight on a stir plate (low speed)

- Begin equilibration of S200 column into SEC buffer using standard BioRad instrument protocol (130 mL, 0.5 mL/min)

Day 5

- After 14-16 hours of TEV cleavage, remove sample from dialysis cassette with syringe/needle
- Load sample onto 2x5 mL HisTrap column and follow with a 100 mL wash with Ni-A buffer, collect 5 mL fractions
- Check samples for protein using Bradford reagent (largest fraction of desired sample generally in fractions 2-6), concentrate protein-containing fractions to 2 mL volume in an Amicon Ultra 15 3k MWCO concentrator in tabletop centrifuge at 4000x g
- Run sample over S200 column, collect all peak fractions (generally ~12-15 mL total)
- Concentrate post-sizing sample to 10-15 mg/mL in an Amicon Ultra 15 3k MWCO concentrator in tabletop centrifuge at 4000x g. Check purity on SDS-PAGE gel. Dispense as 100 µL aliquots into 1.7 mL eppendorf tubes, flash freeze in liquid nitrogen, and store at -80° C until use.

Notes

- 3L of cells is sufficient for any native wild-type NusG prep (CTD or FL), and also for all but the most deleterious mutants (e.g. F165A). Yields for NusG-NTD or Selenomethionine-substituted CTD/FL preps are lower, and thus 6L is recommended for those preps.
- Supplement Ni-A and Ni-B buffers with 0.5 mM TCEP if prepping Selenomethionine-substituted or cysteine-containing constructs (check pH of buffers after adding TCEP)
- Especially when starting from frozen pellets (see stopping point in Day 4), it is fairly straightforward to do up to four different NusG preps in parallel as described on Day 4 / Day 5 steps.
- It is possible to do eight preps at once starting from pellets, but eight preps is easier to manage as two batches of four preps that are staggered throughout the “Day 4” protocol rather than one enormous run of eight in parallel (either way, eight preps will invariably lead to two successive 12+ hour-long prep days).
- Be sure to book at least two FPLCs for the requisite SEC runs if running numerous (4-8) preps at once, and consider using multi-injection SEC protocols to minimize hands-on time (e.g. Load 1st sample into loop, inject for 10 mL onto SEC column, switch to load and for 10 mL on SEC column, pause run and load 2nd sample into loop, 110 mL / collect fractions of 1st sample, inject and collect fractions of 2nd sample).

Preparation of Rho samples for SAXS data collection at ALS BL 12.3.1

Materials

- *E. coli* Rho (purified as described in Nathan Thomsen's dissertation)
- Eppendorf Xstream pipette and tips
- Full Skirt 96-well plates (VWR #10011-228)
- 25 mM ADP-BeF_3 (generated as described in Nathan Thomsen's dissertation)
- 2L of SAXS buffer:
  150 mM KCl, 50 mM HEPES, 5% Glycerol, 5 mM $MgCl_2$, 0.5 mM TCEP, final pH to 7.5
- 5 mL of 2x SAXS buffer (make up 100 mL at a time, store as 1 mL aliquots at -20° C)
- Thermo ST40R tabletop centrifuge and rotor
- Amicon Ultra 15 30k MWCO concentrator (Millipore)
- 10k MWCO slide-a-lyzer MINI dialysis float
- 250 nmole of rU_{12} RNA from IDT (lyophilized)
- Milliq-purified water
- Rainin Pipet-Lite Multi Pipette L12-20XLS+ (herein referenced as "multichannel pipette")
- Stir plate, stir bar and 1L beaker

Day 1

- Concentrate purified Rho to in an Amicon Ultra 15 30k MWCO concentrator to 30 mg/mL
- Pipette concentrated Rho into 10k MWCO slide-a-lyzer mini dialysis float
- Place dialysis float into a beaker containing a stir bar and 1L of SAXS buffer, place at 4° and allow to dialyze overnight on a stir plate (low speed)

Day 2

- Empty 1L of SAXS buffer from dialysis beaker and replace with a second liter of SAXS buffer; allow to dialyze for at least two hours
- Remove sample from dialysis float and quantify protein concentration based on A_{280} reading obtained using a Nanodrop. Keep sample on ice. Save second liter of SAXS buffer from beaker used for dialysis (herein referred to "dialysis buffer")
- Thaw aliquots of 2x SAXS buffer on ice
- Resuspend rU_{12} RNA to 2 mM final concentration (use stated quantity from synthesis and adjust based on concentration determined from A_{260} reading from a Nanodrop) in SAXS buffer used for dialysis
- Dilute 25 mM ADP-BeF_3 1:1 with 2x SAXS buffer
- Generate Rho stock solutions by diluting post dialysis sample with dialysis buffer to 3x the final desired concentrations (generally 5, 2.5, 1.25 and 0.6125 mg/mL final concentrations; make highest concentration stock first and then the rest from three successive 2-fold dilutions)
- Dilute RNA stock with dialysis buffer to concentrations that match 3x the final concentration of Rho hexamer (make the one that matches highest [Rho] first and then do three successive 2-fold dilutions)
- Dilute ADP-BeF_3+1x solutions with dialysis buffer to 3x the final desired concentrations
- Combine samples in these steps for each grouping (denoted A-F) of Rho/RNA/nucleotide (ordered as blank (**A**), high [Rho] (**B**), 2-fold dilution of high [Rho] (**C**), 4-fold dilution of high [Rho] (**D**), 8-fold dilution of high [Rho] (**E**), blank #2 (**F**)); dispense using Eppendorf Xstream pipette and appropriately sized tips
    - Step 1: Dispense 8 µL of dialysis buffer to **A** and **F**, 8 µL 3x Rho stocks to **B**-**E**, wait 30 minutes

    - Step 2: Dispense 8 µL of dialysis buffer to **A** and **F**, 8 µL 3x RNA stocks to **B**-**E**, wait 30 minutes
    - Step 3: Dispense 8 µL of 3x ADP-BeF_3 to **A**-**F**
- Mix all wells with 5 slow depressions with a 20 µL multichannel pipette set to 12 µL
- Spin down plate for 5 minutes at 4,000x g in table-top centrifuge pre-chilled to 4° C
- Either flash freeze plate in liquid nitrogen and store at -80° C until collection (at which point plate should be thawed in room temperature water and spun again at 4,000x g for 5 minutes), or immediately transport to ALS BL 12.3.1 for data collection

Notes:

- Follow all the directions on the Mail-in SAXS website (bl1231.als.lbl.gov/saxs_protocols)
- Do not concentrate the Rho sample after dialysis – this can yield sufficient buffer mismatches between the Rho-containing sample and the buffer blank to render the data unusable. If the Rho sample is not at sufficiently high concentration at the start of the Day 2 protocol then concentrate it again and start over with the dialysis steps.
- If one hopes to determine the impact of a ligand other than RNA or ADP-BeF_3 on Rho ring closure (e.g. protein cofactors, primary site nucleic acids, small molecules), the optimal way to do this is to include these at 3x the desired final concentration in the Rho master mixes. If the ligand is lyophilized then it should be brought up in the dialysis buffer; if it is a protein then it should be co-dialyzed with Rho throughout both dialysis steps. A Rho stock concentration higher than 30 mg/mL may be needed at the start. Alternatively, one can assemble Rho samples with four successive 6 µL pipetting steps rather than three, but the same rules apply with respect to using the dialysis buffer to resuspend and/or co-dialyze the ligand with Rho.
- SAXS is incredibly sensitive to buffer mismatches (especially with ligands that scatter strongly in a SAXS experiment, such as nucleic acids), so pipetting steps must be done with extreme care, especially when pipetting in the ADP-BeF3 master mixes. I generally touch the tip to the wall of the well I'm dispensing into just to be sure no residual amount adheres to the tip, and then gently tap the plate against a hard surface to push the drops to the bottom of the wells. Dialysis would be optimal, although it would involve the generation of very large quantities of hazardous waste and hence isn't possible with most experiments that contain millimolar concentrations of Be (especially if one desires to screen many conditions, which is a great use of the high throughput setup at 12.3.1) – hence the neurotic pipetting instead.

FA assay for monitoring RNA binding to Rho's secondary site

Materials

- *E. coli* Rho (purified as described in Nathan Thomsen's dissertation)
- 25 mM ADP-BeF_3 (generated as described in Nathan Thomsen's dissertation)
- Synergy Neo2 Plate Reader (BioTek)
- Thermo ST40R tabletop centrifuge and rotor
- 10 mg/mL BSA (pre-frozen aliquots)
- 1 aliquot of 5′ 6-Fluorescein amidite-labeled RNA (bring up to 100 mM in milli-Q water, store as 10 µL aliquots at -80° C)
- 1M DTT
- Milliq-purified water
- Rainin Pipet-Lite Multi Pipette L12-20XLS+ (herein referenced as "multichannel pipette")
- 2L of RNA binding assay buffer (herein referred to as "assay buffer"):
  150 mM KCl, 20 mM HEPES, 5% Glycerol, 5 mM $MgCl_2$, 0.5 mM TCEP, final pH to 7.5
- 2x assay buffer (make up 100 mL at a time, store as 1 mL aliquots at -20° C)
- Corning 384-well low volume plates (#3544)
- Stir plate, stir bar and 1L beaker
- 10k MWCO slide-a-lyzer MINI dialysis float
- Amicon Ultra 15 30k MWCO concentrator (Millipore)
- Benchtop microcentrifuge

Day 1

- Concentrate purified Rho to in an Amicon Ultra 15 30k MWCO concentrator to 30 mg/mL
- Pipette concentrated Rho into 10k MWCO slide-a-lyzer mini dialysis float
- Place dialysis float into a beaker containing a stir bar and 1L of assay buffer, place at 4° and allow to dialyze overnight on a stir plate (low speed)

Day 2

- Empty 1L of assay buffer from dialysis beaker and replace with a second liter of assay buffer; allow to dialyze for at least two hours
- Remove sample from dialysis float and quantify protein concentration based on A_{280} reading obtained using a Nanodrop. Keep sample on ice. Save second liter of assay buffer from beaker used for dialysis (herein referred to "dialysis buffer")
- Thaw aliquots of BSA, DTT, 2x assay buffer and fluorescein-labeled RNA on ice
- Spin down fluorescein-labeled RNA aliquot at 13,000x g in a benchtop microcentrifuge, and dilute to 1 mL with dialysis buffer (1 mM final concentration)
- Assemble a "3x RNA master mix" with the following composition:
  60 nM fluorescein-labeled RNA, 15 mM DTT (with equal volume of 2x assay buffer), 1.5 mg/mL BSA (with equal volume of 2x assay buffer), then fill to necessary volume with dialysis buffer
- Dilute Rho sample to 3x the highest concentration desired with dialysis buffer, place in 1st column of a 96-well plate
- Perform ten serial 2-fold dilutions of Rho with dialysis buffer (so columns 1-11 contain the dilution series of Rho) and include a no Rho control in column 12
- Add 16 µL of 3x RNA to the requisite number of wells in a 96-well plate
- Transfer 16 µL of 3x Rho stocks from dilution series to 3x RNA samples with a multichannel pipette, mix thoroughly, wait at least 30 minutes

- Dilute 25 mM ADP-BeF_3 with an equal volume of 2x assay buffer, and then dilute further to 3x the final desired concentration of nucleotide with dialysis buffer (herein described as 3x nucleotide). Transfer to wells 1-12 of a 96 well plate
- Add 16 µL of 3x nucleotide to the 3x Rho / 3x RNA of wells in a 96-well plate with a multichannel pipette and mix thoroughly. Set multichannel pipette to 20 µL and mix thoroughly again. Wait at least 30 minutes.
- Use the multichannel pipette to transfer both 20 µL replicates of every Rho/RNA/nucleotide mix to a Corning 384-well low volume plate ("assay plate") with replicates in adjacent wells (so each 12-well row fills a 24-well row on the assay plate). The scripts that will be used to process the data generated will assume that the no Rho control samples are in columns 23 and 24, so do the transfer accordingly. Be sure to only depress the pipette down to the first stop to avoid introducing bubbles.
- Spin the assay plate down in a tabletop centrifuge at 4000x g for 2 minutes – this step is necessary to remove any bubbles introduced in pipetting and to flatten out the meniscuses of every well
- Pre-warm the BioTek Neo2 plate reader to 30° C and follow on-screen prompts to switch in the proper filters for Fluorescein FA reading after loading the appropriate program. Select the appropriate well range and set the PMTs to 65/61. Set to read once per minute (or minimal interval allowed by instrument, generally ~1:15 if reading a full plate) for 30 minutes.
- Press start run, define file name and place plate on carriage. Press OK and allow reading for the full 30 minutes.

Data analysis

- Inspect plots of individual parallel and perpendicular values with respect to time. These values generally display a period of rapid change (likely indicative of temperature equilibration) that runs the first 5-10 minutes, and then a period of stability that is maintained through the end of the run. Any of the sets of readings in the stable period are suitable for further analysis; I generally use the reading from 10-15 minutes
- Click on Data Reduction and enable display of FA values in mA. Set time to a set of readings from stable period of parallel and perpendicular values, and click the excel button. Save the spreadsheet that pops up with an appropriate file name.
- Open the excel spreadsheet with FA values and copy/paste all the data (with row name, but without column names) into a plain text file (I find a text editor such as Komodo Edit is great for this step). Thus, if I had only taken readings from rows A and B, my file would look something like this (although "A" values and "B" values would each be on a single line):

```
A       171     174     164     165     138     131     93      89      68      65      57      59
        52      54      52      53      51      52      53      51      52      48      48      47
B       167     172     167     166     135     133     89      88      66      66      56      57
        53      54      52      51      50      51      52      51      52      49      49      45
```

- Save this data as a plain .txt file (e.g. assay_01.txt) and move it to the same directory as the "FAconvert.py" script (available at: https://github.com/jaglawson/platereader_code/)
- Replace the text defined as "inputfilename" with the name of your .txt file
- Open a terminal, navigate to the directory with the data and the FAconvert.py script
- Type the following command: "python FAconvert.py"
- Open the output files (defined by the variables "outputfilename" and "outputfilename2", default is to append "averaged_subtracted_" and "averaged_subtracted_tabbed_" to the input file name) generated by the script. Each pair of replicates (e.g. wells 1 and 2) have

now been averaged, and the baseline (defined as the average of 23 and 24) has now been subtracted from every one of these pairs of replicates. Thus, in the above example, the “averaged_subtracted_” output file from the above example would now look like this:

```
A
125.0
117.0
87.0
43.5
19.0
10.5
5.5
5.0
4.0
4.5
2.5
B
122.5
119.5
87.0
41.5
19.0
9.5
6.5
4.5
3.5
4.5
3.5
```

- The “averaged_subtracted_tabbed_” output file would look like this:

```
A       B
125.0   122.5
117.0   119.5
87.0    87.0
43.5    41.5
19.0    19.0
10.5    9.5
5.5     6.5
5.0     4.5
4.0     3.5
4.5     4.5
2.5     3.5
```

- I personally find the “averaged_subtracted_tabbed_” format easier to handle for copy/pasting into a data processing program such as Prism, but I also generate the “averaged_subtracted_” file since it’s a format that would be more conducive to further scripting of this analysis
- Open Prism and click “enter __ replicate values side-by-side subcolumns” and set __ to read the number of replicates you expect to have for this particular experiment (note, each pair of duplicates from a particular master mix above will only count as one replicate and has already been averaged at this point)
- Copy/paste the data into Prism as Y values, replace the former row titles (e.g. “A” or “B”) with a meaningful descriptor, and input the concentrations of Rho hexamer in the dilution series as X values
- Press the “Analyze” button, select nonlinear curve fit and select all data series you want to fit. Select “One site -- specific binding with Hill Slope”,
- Save the output Kd values from the fits, along with the plots as their own separate file

Notes:

- Annoyingly, the final $K_{d,app}$ values from this assay are somewhat sensitive to the incubation times between addition of Rho to RNA until reading, and even more sensitive to the time from nucleotide to Rho/RNA until starting the read. I generally set a timer for myself and keep those incubation times fixed as 30 minutes of Rho/RNA, 30 minutes of Rho/RNA/ATP, and then 10 minutes for transfer and starting the read. Strict adherence to this timeline generates highly repeatable results with respect to the raw data and the resulting Kd,app fits.
- If one hopes to determine the impact of a ligand other than RNA or ADP-BeF_3 on Rho ring closure (e.g. protein cofactors, primary site nucleic acids, small molecules), the optimal way to do this is to include these at 3x the desired final concentration in the Rho master mixes. If the ligand is lyophilized (small molecule or RNA) then it should be brought up in the dialysis buffer; if it is a protein then it should be co-dialyzed with Rho throughout both dialysis steps. A Rho stock concentration higher than 30 mg/mL may be needed at the start. Alternatively, one can assemble samples for FA reading with four successive 13 µL pipetting steps rather than three 16 µL steps, but the same rules apply with respect to using the dialysis buffer to resuspend and/or co-dialyze the ligand with Rho.
- In a manner somewhat similar to SAXS (but not as severe), FA readings can be sensitive to mismatches in buffer conditions, especially glycerol. Be mindful of this fact when adapting this protocol to fit your needs.